ITEM WRITER'S MANUAL

FOR THE

**AMERICAN HOLISTIC NURSES
CREDENTIALING CORPORATION
(AHNCC)**

Center for Nursing Education and Testing
35 Journal Square, Suite 901
Jersey City, NJ 07306

Item Writer's Manual for
AMERICAN HOLISTIC NURSES CREDENTIALING CORPORATION

Table of Contents

Item Writer's Manual for
Certification Examinations

Introduction

This item-writing manual has been prepared to assist you in writing test questions for certification examinations in nursing and related fields. While the section on Development of the Examination addresses the process of creating a certification examination, the principles of item-writing in later sections of the manual can be used in a wide variety of situations where nursing knowledge is measured. The principles apply to teacher-made tests in nursing education programs, continuing education programs, and staff development programs. We hope that you will enjoy learning the process of item writing–a unique process that combines creativity with adherence to rules and editorial standards–and that you will find it a useful skill throughout your professional career.

Development of the Certification Examination

Purpose. The purpose of a nursing certification examination is to evaluate the extent to which registered nurses have attained the knowledge and skills deemed necessary for competent practice in a particular nursing specialty. Certification indicates a higher degree of professional competence than the minimal requirements for licensure as a registered nurse. In a specialty practice, nurses combine elements of basic nursing knowledge with knowledge in the specialty to deliver care/services to patients/clients. The certification examination is designed to measure the attainment of that knowledge at a defined level of competence.

Defining the Domain of Practice. The initial step in the process of developing a certification examination is to define the practice domain. Important building blocks for the development of a certification examination include the specialty's published *Standards of Practice* and statement of the scope of the practice. Most specialties develop a *Core Curriculum* that includes the essential body of knowledge that practitioners need to practice competently. This is also a valuable aid in defining the domain of practice. Most importantly, the process requires the participation of expert practitioners in the specialty area. This group, often called the Certification Task Force, has the responsibility of ensuring that the examination will be a valid, reliable measurement instrument. This is accomplished by performing a formal analysis of practice (practice or job analysis).

Validity. Validity is the most important consideration in test development and utilization. For a certification examination to accomplish its intended purpose, the examination must assess the performance that the scores are intended to represent. Simply stated, validity addresses the question, "Are we testing what we think we're testing? That is, are we testing the essential knowledge required to practice competently in this specialty practice?" Test validation is the process of accumulating evidence to support the assertion that the examination is truly measuring what it intends to measure and, therefore, that scores on the examination provide meaningful information.

The most important method of attaining that evidence is to determine what practitioners in the specialty are actually doing through a practice analysis, usually a survey of nurses in the specialty. It is essential that the content domain to be covered by a certification examination is clearly defined and explained in terms of the importance of the content for competent performance in the nursing specialty.

Reliability. The other important concept in testing has to do with the consistency of the instrument (the test) in measuring the candidates' ability level. As an analogy, consider the bathroom scale. A reliable scale will give the same measurement (weight) if a person gets on and off the scale several times within a few minutes. An unreliable scale may give quite different readings. In testing, the consistency of performance is measured several ways–most often by looking at the internal consistency of the examination. That is, if a person performs well on the odd-numbered items, the person should also perform well on the even-numbered items in the test. Correlation coefficients are calculated to determine if a group of persons perform consistently throughout the examination. Reliability, an index of how consistently–or accurately--the examination measures the candidates' knowledge, is necessary for the test to be valid.

Practice Analysis. The purpose of a practice analysis is to define the content domain that will be measured by an examination. The term "practice analysis," rather than the traditional term "job analysis," is used since job analysis suggests that the requirements of a specific job are being defined, while certification is concerned with a broader practice in a wide range of settings and a diverse population of patients/clients. Practitioners in the specialty, however, operate from the same general knowledge base and use similar practice methodologies.

The development of a certification examination requires accurate information about the kinds of client problems encountered in practice and about the knowledge, skills, and judgment required to handle these problems. Practice-analysis studies must be performed periodically by certification bodies in order to ensure that the examination continues to reflect current practice. A number of variables can influence the evolving practice of a nursing specialty. Among these are changing trends in the provision of care to clients, new knowledge and technology, the educational preparation of practitioners, and evolving standards of practice or new roles in the specialty.

Throughout the process of test development, the procedures described in the *Standards for educational and psychological testing* (AERA, APA, & NCME, 2014) are followed. Other standards and guidelines used to ensure a legally defensible certification examination include the *Accreditation Standards of the Accreditation Board for Specialty Nursing Certification* (ABSNC, 2019), and the *Principles of fairness: An examining guide for certification boards* (CLEAR, 2015).

Logical analysis. While there are a variety of approaches to practice analysis, logical analysis of the nature and requirements of a practice is the oldest approach and is, in fact, a key component of *any* practice analysis methodology. It includes the involvement of a panel of experts (Certification Task Force) who have extensive experience and knowledge of practice. Prior to meeting together, these experts review available data concerning the practice and practitioners. The

Task Force, together with test development staff, meet to define the nature of the practice in terms of the needs of the clients served, the nursing activities performed for the clients, and the knowledge needed to provide these services. From this logical analysis, a theoretical model of practice emerges and the knowledge, skills, and abilities (KSAs) of the practice are delineated within this framework. This model provides the basis of the test specifications (test blueprint). The Task Force may then assign weights to the various parts of the blueprint.

Empirical analysis. Following the logical analysis, a larger-scale study of practice is usually performed to include the participation of a much larger group of nurses in the specialty practice. This study is performed to ensure that important aspects of practice are not overlooked and to determine if the weights assigned to various parts of the test blueprint correspond to actual practice. A survey methodology is most often used for the empirical study, since it involves the participation of actual practitioners and allows for a comprehensive description of the practice area. It is also relatively inexpensive when compared to other empirical approaches to practice analysis, such as direct observation of practice or the collection and analysis of critical incidents.

Test specifications (blueprint). The analysis of the survey results are used to define the test specifications (blueprint). The blueprint provides an outline of the domain of practice (i.e., major topics) to be addressed in the examination. Often, in nursing certification examinations, the blueprint is a two-dimensional table that specifies two critically important categories: (1) the type of client (or patient problem/case), and (2) the major activities performed by nurses in the specialty practice. The blueprint also incorporates the list of activity statements in the empirical survey that were found to be important in actual practice. (The blueprint for your certification examination is included at the end of this manual in Appendix C.)

Item writing. In addition to adhering to the test blueprint, a certification examination's validity is largely dependent on the quality of the test questions (items). Item-writing for a practice profession poses challenges related to differences in practice in different areas of the country. In addition, professional practice is constantly changing, requiring on-going validation of items. Further, to test specialty nursing knowledge, items should require the examinee to analyze and apply information in practice situations, rather than to simply recall facts. Writing these kinds of items requires considerable item-writing skill and practice. Therefore, much of the expense and effort of a certification program is aimed at writing, editing, verifying, and pretesting items.

For initial item-writing of a certification examination, teams of experts from various practice settings and geographic areas are brought together to be taught how to write items and to generate items according to the test specifications, using the specific activity statements that were found to be important in the survey of practice. As items are written, they are documented (validated), using current, authoritative, nursing and medical resources, including internet resources. A large pool of items is written, since it is expected that about one-third of the items will be lost. That is, many of the items will prove to be too easy or too difficult, or will have other flaws.

Pilot-testing. Following the item-writing sessions, the test questions are edited and reviewed. Further documentation (validation) is performed, as needed. The items are assembled into several pilot test forms to be administered to as many volunteer examinees as possible–usually at a national meeting of the specialty association. The items in the pilot test are reviewed by the Certification Task Force, and additional revisions are made, as needed, prior to pilot testing. Persons who participate in the pilot testing are given feedback about their overall test performance–the percent of questions answered correctly and the average percent correct of all who took that form of the pilot test. Following statistical analysis of the results of the pilot administration, the items for the final examination are selected. The Certification Task Force, which by now has probably evolved into the Certification Board, reviews the items selected for the certification examination and makes revisions, as needed. Now the actual test is ready to be administered.

Setting the passing standard. To set the passing standard, a panel of experts, who may or may not be members of the Certification Board, perform a standard-setting procedure known as a modified Angoff procedure to establish the passing criterion. This procedure involves examining each item on the examination and estimating the probability that a candidate who minimally passes the examination would answer the question correctly. In the modified procedure, judges are given the opportunity to discuss their initial choices of probability, consider actual test data, and then modify their initial decisions, if they wish. Prior to performing the procedure, the panel of experts is provided training and practice with items that are not on the actual examination.

On-going test development activities. Following the initial administration of the actual certification examination, the Certification Board reviews the test and item analysis statistics for each question. This process continues on a regular (usually annual) basis. On-going item writing is needed to build the item bank and to ensure that questions are not compromised by too-frequent exposure on multiple forms of the examination. Before an item-writing session, the item bank is reviewed to identify areas in which items are needed. Nurses in the specialty who have expertise in the areas needed are invited to participate in item-writing sessions.

Avoiding item bias

Bias is the presence of some characteristic of an item that results in differential performance for individuals of the same ability but from different groups, e.g., ethnic, age, cultural, gender identity, or religious groups. Since important decisions are based on test scores, it is critical to avoid bias, which may unfairly influence examinees' scores.

To avoid bias and ensure sensitivity and fairness, items are reviewed at several points during the test development process for references to age, gender, race, religion, disability, or any possibly offensive terminology. Bias review begins with the item writers who should consider the following principles throughout the item-writing process:

1.　　Avoid stereotypes. No population group should be depicted as either being inferior or superior. For example, older people should not be characterized as feeble,

dependent, and incompetent. Other examples include stereotypical occupations (e.g., Chinese launderer) or stereotypical situations (e.g., objective male and intuitive female).

2.   Avoid inflammatory material, such as incest or violence, except where such material is both relevant and essential to meet test specifications. In nursing examinations, it may be necessary to ask questions about material that might have a negative emotional impact for some population subgroups, e.g., questions about patients undergoing sexual reassignment surgery. These questions should always be handled carefully and sensitively.

3.   Avoid using diagnoses as adjectives. For example, instead of "the diabetic," refer to "the person who has diabetes mellitus."

4.   Avoid having any population dominant on the examination. Use ethnicity only when necessary. Avoid male and female identification when possible. If gender is important to identify, be sure there is a balance of male and female throughout the examination.

Another consideration of bias has to do with the level of the content of test questions. That is, questions written for a basic nursing student should not include content taught at the graduate level. Similarly, questions written for a certification examination should not include advanced practice content unless the examination is intended to measure advanced practice in the specialty area.

A sample Item Bias Review Form is available to use during review of items to help guard against stereotyping or inadequate representation of various sub–groups, referred to on the form as "designated sub-groups of interest (DSI)." The sample form is in Appendix B.

Writing Good Multiple-Choice Test Questions

The Test Blueprint

Good tests do not arise spontaneously from item writers. They are carefully planned. A test has a blueprint, or test plan, which provides the test specifications for item writers. Usually, there is also a list of key content to be covered by the test, called a "test content outline." For certification examinations, there is often a list of "activities performed" in the particular specialty area. All of these serve as item-writing guides.

Prior to the item writing session, the current pool of items is reviewed and compared to the blueprint categories, activity lists, and/or content outlines. This review reveals areas of content that need additional items, and provides guidance to the item writers about specific topics that need to be addressed.

Characteristics of Multiple-Choice Items

A multiple-choice item consists of a question/problem and a list of several suggested solutions. The question/problem, which may be stated as a direct question or as an incomplete statement, is called the *stem* of the item. The list of suggested solutions may include words, numbers, symbols, or phrases and are called options. The examinee is typically requested to read the stem and the list of *options* and to select the one correct, or best, alternative. The correct option in each item is called the correct answer or *key*, and the remaining options are called *distractors*. The incorrect answers receive their name from their intended purpose--to distract those examinees who do not know the correct answer.

How does one decide whether to develop an item as a closed question, ending in a question mark, or as an open item, structured as an incomplete sentence that is completed by one of the options? The answer depends on which method most clearly asks the question that the examinee must answer. Although a direct, closed question is easier to write and is more likely to frame a clear question, an incomplete statement can be more concise and, if skillfully phrased, can also ask clear question. A common practice is to begin with a direct, closed question, and to change the form to an incomplete (open) statement only when the clarity of the problem can be maintained and greater conciseness (and readability) achieved.

Recall of Knowledge

Items at the level of recall of knowledge ask the examinee to remember specific facts, common terms, basic concepts, and principles. In these items, there is only one possible correct answer. All other alternatives are absolutely wrong.

Direct Question:

Which of these organisms causes Lyme disease?

1. *Coxiella burnetii.*
2. *Borrelia burgdorferi.*
3. *Cryptococcus neoformans.*
4. *Blastomyces dermatitidis.*

Key: 2

Indirect question:

The causative organism of Lyme disease is

1. *Coxiella burnetii.*
2. *Borrelia burgdorferi.*
3. *Cryptococcus neoformans.*
4. *Blastomyces dermatitidis.*

Key: 2

In nursing, as in other applied sciences, not all knowledge can be stated so precisely that there is only one absolutely correct response. In fact, when we get beyond the simple aspects of factual knowledge, answers of varying degrees of acceptability are the rule, rather than the exception. Usually several possible reasons or possible procedures may be possible. Measures of achievement in these situations become a matter of selecting the best answer. This type of question is especially useful for measuring learning outcomes that require the understanding, application, or interpretation of factual information. In these situations, care must be taken to be certain that the best answer is agreed upon by experts, as well as documented with current references, so the answer can be defended as clearly the best.

Example:

Which of these factors is ***most*** likely to be responsible
for the increase in cases of Lyme disease in Connecticut?

1. An increase in the deer population.
2. A decrease in childhood immunization programs.
3. An increase in the number of domestic cats per population.
4. A decrease in the number of persons covered by medical insurance.

Key: 1

The best-answer type of multiple-choice item tends to be more difficult than the correct-answer type, partly due to the finer discriminations called for and partly due to the fact that such items measure more complex reasoning.

Knowledge of principles is also an important factor in nursing knowledge. Multiple-choice items can be constructed to measure knowledge of principles in a way similar to those designed to measure facts.

Example:

> Which of these principles of infection is applicable to the transmission of Lyme disease?
>
> 1. The highly infective blood and purulent exudates of infected patients transmit the causative organism.
> 2. Inhalation of the causative organism by immunocompromised individuals leads to the disease.
> 3. The causative organism is spread through the bite of a tick that is carried by a host animal.
> 4. The causative organism is transmitted from person-to-person through sharing of contaminated clothing, hairbrushes, or combs.
>
> Key: 3

Measuring Understanding and Application of Knowledge

In measuring nursing knowledge, it is usually desirable to go beyond factual knowledge to measure various aspects of understanding and application of that knowledge in clinical situations. It is, however, important to keep in mind that such items go beyond factual recall of knowledge only if the applications and interpretations are new to the examinees. That is, specific applications or interpretations of knowledge can be taught directly to learners, as any other fact is taught. When this is done, and the test items contain the same problem situations and solutions used in teaching (or used in the core curriculum for the certification examination), it is obvious that the examinees can be given credit for no more than the mere retention of factual knowledge.

To measure understanding and application, the patient situation must be somewhat different than those used in class (or in reference material). The objective of the item writer is to determine if an examinee can apply knowledge and make clinical decisions in new patient situations. Therefore, it is important that item writers of nursing examinations be seasoned practitioners who have had many clinical experiences and can imagine other scenarios that would require the examinee to apply knowledge and make clinical judgments.

To write items at the application level, first consider the content/concept to be tested. Then think of a typical situation where the nurse needs to use that knowledge in patient care. Briefly describe the patient situation: age, diagnosis, and pertinent complications or psychosocial problems, medications, etc. Do not include irrelevant information, unless the purpose of the question is to distinguish between relevant and irrelevant data.

A helpful hint for writing application items is to keep the nurse and the client in the stem of the question. This helps the item writer keep the focus on how a nurse uses specialized knowledge in the care of the client, rather than focusing on recall of a fact.

Examples:

When teaching persons in a high-risk area about how to prevent Lyme disease, the nurse should include which of these instructions?

1. "Keep away from stagnant water, where mosquitos might breed."
2. "Wear long pants, with the cuffs tucked into your socks."
3. "Use an antibacterial soap to bathe after you've been in the woods."
4. "Avoid contact with anyone who is being treated for the disease."

Key: 2

A four-month-old infant has a fiberglass hip-spica cast applied for the treatment of developmental dysplasia of the hip. While the cast is drying, the nursing diagnosis that should be given the *highest* priority is high risk for

1. pain related to musculoskeletal injury.
2. sensorimotor activity deficit related to limited physical activity.
3. impaired skin integrity related to irritation from cast.
4. altered peripheral tissue perfusion related to cast constriction.

Key: 4

A patient had a total hip replacement using a posterior approach on the previous day. The nurse makes all of the following observations of the patient. Which one requires *immediate* intervention?

1. The wound drainage system (Hemovac) contains about 20 mL of dark red drainage.
2. Anti-emboli hose are on both lower extremities.
3. The hip is in an adducted position.
4. The head of the bed is elevated 30 degrees.

Key: 3

Analysis of Clinical Situation

Items requiring analysis of data present a considerable amount of patient information. They may ask the examinee to draw a conclusion, make a nursing diagnosis, or identify what additional information should be obtained.

Examples:

A 23-year-old construction worker is brought to the Emergency Department following a 15-foot fall at work. The worker complains of pain in the pelvic area and right thigh. X-rays reveal right superior and inferior pubic rami fractures and a midshaft fracture of the right femur. No other fractures are seen. Because of the nature of the patient's injuries, the nurse should be aware that the patient is at high risk for which of these complications?

1. Pulmonary contusion and pneumothorax.
2. Myocardial contusion and dysrhythmia.
3. Burst fracture of spine and paraplegia.
4. Urethral injury and ruptured bladder.

Key: 4

A 30-year-old patient undergoes intramedullary fixation of a midshaft femur fracture. Three days later, the patient becomes restless, agitated and complains of shortness of breath. The nurse observes a petechial rash on the patient's chest. It is *most* likely that the patient is experiencing which of these complications?

1. Fat embolism syndrome.
2. Pulmonary embolism.
3. Cerebral edema.
4. Myocardial ischemia.

Key: 1

A 72-year-old male underwent a right hip arthroplasty 24 hours ago. The patient is receiving epidural analgesia for pain management and complains that he is no longer getting adequate pain relief. Using a 10-point rating scale, the nurse determines that the patient's pain has gone from 3 to 8. The nurse should seek the answer to which of these questions *initially*?

1. Does the right leg show signs of infection?
2. Is there excessive drainage on the patient's hip dressing?
3. Has the epidural catheter become dislodged?
4. Does the patient have a low pain threshold?

Key: 3

Regardless of the item format, the most important question for the item writer to keep in mind is, "What is the concept that I am trying to test?" By starting with the concept to be tested, the item writer will be able to maintain a focus for the item. It is usually helpful to first jot down the stem of the question, then write the correct answer, and finally the distractors. It is usually necessary to go back and revise the stem, to ensure the distractors' plausibility. Remember to document (i.e., validate) that the correct answer is indeed the one and only correct answer presented, using current, authoritative sources.

Writing Test Questions in Case Sets

Case sets of test questions are a group of items pertaining to a single patient or situation. The case begins with an introductory paragraph, called the "initial expand." Several questions are usually asked about the initial situation. Then, another expand may modify the situation, as time passes or healing (or a complication) occurs. Case sets have the advantage of seeming more like a real situation that a nurse would be dealing with over time.

To develop a case set, first consider the content/concepts to be tested, and think of a typical patient situation. Briefly describe the situation: age of employee, diagnosis, any pertinent complications or psychosocial problems, medications, etc. This initial "story" is the initial *expand*. Keep it simple.

Then list the concepts that you want to test in this case set of questions. Put them in order from first to last, in a logical sequence. Write the stem of the first question and the answer to the question. Determine if you need to modify the stem to reach the answer. Then write the distractors (wrong answers). Reconsider the stem. Does it have to be further modified?

Continue to write questions in this manner, aiming for 4-5 questions per case. You may need to add additional expands as the case changes; e.g., "Mrs. X is discharged and returns to the clinic two weeks later for follow-up." You can follow a case over a long period of time, e.g., "Two years later, Mrs. X says she wants to become pregnant."

In developing a case set, consider:

1. Is the problem that is presented in the case situation appropriate for the test?
2. Does the initial expand provide sufficient information for the items in the case set?
   Conversely, is there extraneous information that should be deleted from the expand?
3. Do the items flow in a logical order? Does the case seem like a "real" situation to you?
4. Is each item in the case set independent of the other items? The answer to an item should *not* depend on the answer to another item in the case.
5. Are there words in the expands, stems, or options that are clues to the correct answer? Or are there false clues that suggest that a distractor is the correct answer? Items should *not* be designed to trick the examinees.

An example of a case set of items about the nursing care for a dermatology patient is given on the following pages. Case sets lend themselves well to nursing situations, since usually there are a number of important concepts that need to be tested.

SAMPLE CASE SET OF QUESTIONS

Mr. Godfrey Brown, 62 years old, calls the dermatologist's office to make an appointment because of a recurring infection of the toenails of both feet.

EXPAND - a paragraph initiating a case situation that gives information needed to answer questions about the situation. Additional expands are used as the situation changes and new information is required.

1. During the initial phone conversation, Mr. Brown tells the nurse, "That darned fungus infection of my toenails is back after three years of pretty much being under control. I've had this problem on and off for 20 years, and nothing seems to get rid of it!" Which of these responses by the nurse would be *best*?

STEM - the initial part of the item in which the problem is stated. The stem may be presented as a question, direction, or an incomplete statement.

1. "You have a condition that is almost impossible to treat effectively."

OPTIONS - all of the response choices in an item.

2. "A new surgical procedure may be able to help you."

3. "There are new medications available that are quite successful in treating your problem."

4. "You must have neglected measures to prevent this kind of infection from recurring."

Key: 3
NP: Implementation
Codes: A/2, Infection

KEY - the option that is the correct answer. Incorrect answers are called DISTRACTORS.

NP - Nursing process phase

CODES - identifies area of dermatology certification exam blueprint tested in this item

---

Mr. Brown keeps his scheduled appointment and is seen by the dermatologist.

---

2. Mr. Brown's diagnostic tests include a potassium hydroxide (KOH) preparation and a fungal culture. Both tests are performed because the KOH preparation alone *cannot*

    1. indicate if fungi are present.

    2. yield results in less than 10 days.

    3. be performed with a high degree of accuracy in an office setting.

    4. identify the specific fungal pathogen.

Key: 4
NP: Planning
Codes: A/3, Infection

The diagnostic tests confirm the diagnosis of dermatophytic onychomycosis. The dermatologist plans to prescribe terbinafine hydrochloride (Lamisil) 250 mg daily for 12 weeks.

---

3. The nurse should give Mr. Brown which of these instructions about taking Lamisil?

    1. "Be sure to take the medication with food."

    2. "You must take the medication for the whole twelve weeks."

    3. "If you miss a dose of medication, double the next dose."

    4. "Stay out of direct sunlight while you are taking this medication."

Key: 2
NP: Implementation
Codes: A/1, Infection

4. Mr. Brown is given a prescription for Lamisil. He asks the nurse, "When will my toenails start to look normal again?" The nurse's response should be based on the understanding that after initiating Lamisil therapy, the nails will probably look normal

    1. in about two weeks.

    2. in about one month.

    3. when the nails grow out completely.

    4. when the medication regimen is completed.

Key: 3
NP: Implementation
Codes: A/2, Infection

5. The nurse teaches Mr. Brown measures to prevent the recurrence of his infection. Which of these statements, if made by Mr. Brown, would indicate that learning had occurred?

    1. "I plan to use an antifungal powder in my shoes weekly."

    2. "I'll go barefoot whenever I can at home and outside."

    3. "I'll select shoes that fit snugly."

    4. "I plan to get new socks made of rayon or polyester."

Key: 1
NP: Evaluation
Codes: A/2, Infection

Item-Writing Principles

The Stem*

Goal: State the examinee's task clearly and concisely.

1. Give examinee enough information to answer the question.

    Poor:

    Hyperthyroidism is

    1. more commonly found in men than women.
    2. related in some way to pre-existing severe stress.
    3. primarily a disease of persons over 50.
    4. generally accompanied by intolerance to cold.

    Better:

    A client is observed for symptoms of hyperthyroidism, which include

    1. lethargy.
    2. bradycardia.
    3. brittle nails.
    4. heat intolerance.

2. Include only relevant information in the stem (unless the item is testing the ability to distinguish between relevant and irrelevant information).

    Poor:

    Mrs. Mary Jones, 35 years old, goes to the physician's office because of symptoms suggestive of hyperthyroidism. She is married and has three children. She is employed on a part-time basis as a cook in a restaurant close to her home. If Mrs. Jones has hyperthyroidism, she is likely to have which of these symptoms?

    1. Lethargy.
    2. Bradycardia.
    3. Brittle nails.
    4. Heat intolerance.

    Better:

    A client is being observed for symptoms of hyperthyroidism, which include

    1. lethargy.
    2. bradycardia.
    3. brittle nails.
    4. heat intolerance.

*The examples of poor items are taken from actual teacher-made tests. They often have multiple problems, including poor grammar, ambiguity, no correct answer, or multiple correct answers.

3. Avoid stems that ask for a series of multiple true-false options.

   Poor:

   Which of the following is true about the middle years?

   1. It encompasses ages 19 to 30.
   2. It is the most conflict-free period of life.
   3. It is characterized by dramatic changes in our sense of values.
   4. It is marked by a conflict between intimacy and values.

   Better:

   According to Erikson, the middle adult years are characterized by the conflict between

   1. intimacy and isolation.
   2. generativity and stagnation.
   3. integrity and despair.
   4. industry and inferiority.

4. Put a word(s) that is repeated in each of the options in the stem.

   Poor:

   What is the primary reason for maintaining the patient's skin integrity?

   1. The skin is the covering for the underlying structures.
   2. The skin is the body's main defense against the invasions of microorganisms.
   3. The skin stores water for the body.
   4. The skin can help control body temperature.

   Better:

   The primary reason for maintaining a patient's skin integrity is that the skin

   1. covers the body's underlying structures.
   2. provides the body's main defense against infection.
   3. stores water for the body.
   4. controls the body's temperature.

   Even Better:

   The skin serves which of these primary functions for the body?

   1. It covers underlying structures.
   2. It prevents infection.
   3. It stores water.
   4. It controls temperature.

   As the item is made clearer, it becomes apparent there are multiple correct answers.

5. Use technical terms when appropriate, but keep the reading level low (since the purpose of your tests are not to measure reading ability). To keep the reading level low, use common terms (e.g., harmful instead of deleterious).

15

6. Avoid words like "adequate," "acceptable," and "moderate" because they mean different things to different people.

7. Avoid abbreviations unless they are common.

8. Underline, italicize, or bold face key words, especially if they are negative (*not, least, except, opposite*).

9. Use a negatively worded stem sparingly and only if absolutely necessary.

    Examples of appropriately written "negative" stems:

    The physician writes the following postoperative orders for an adult who just had a partial gastric resection..
    Which one should the nurse **question**?

    1. Morphine sulfate by PCA 1 mg q 6 min, with a 40 mg lockout q4h.
    2. Attach gastric sump tube to continuous suction at 130 mm Hg.
    3. Intravenous infusion of 5% dextrose in saline at 125 mL/hr.
    4. Monitor bowel sounds for return of peristalsis.

    An 18-year-old male patient arrives at the outpatient surgical center for a scheduled arthroscopy. As the nurse is preparing the patient for surgery, he makes all of the following statements. Which one should be reported to the surgeon?

    1. "I took two Tylenol pills for pain in my knee last night."
    2. "My knee was really swollen yesterday."
    3. "I didn't sleep well last night."
    4. "My brother broke out with chickenpox this morning."

10. Use correct grammar, spelling, and punctuation.

    A. The stem should contain a subject and a verb.

    B. If the stem of the item is open, each option should

        1. form a complete sentence.
        2. start with a small letter.
        3. end with a period.
        4. be grammatically consistent with the stem.

    C. If the stem is closed, the stem should end with a question mark and each option should start with a capital letter.

    D. Use punctuation (commas, semicolons, and periods) to help clarify long and complex stems but try to keep sentences short.

Poor:

The following is true regarding the use of disposable gloves:

1. apply gloves using sterile technique
2. you must wear the exact size glove that fits your hand
3. to be used when hands will be in contact with heavy secretions or excretions (fecal incontinence, oral care)
4. apply right hand first - then left hand - making sure to touch only the cuff

11. Check that the stem of one item does not give away (clue) the correct response in another item. For example, the following three items appeared in this order on an actual teacher-made test:

7. All the following adhere to the principles of medical asepsis, **except**

   1. proper handwashing.
   2. wearing a clean uniform.
   3. use of transfer forceps to move supplies.
   4. use of examination gloves when handling body fluids.

8. After the hands are washed with a germicidal agent such as chlorhexidine, it would be **most** accurate to say that the hands have been

   1. rendered free of all organisms.
   2. sterilized.
   3. rendered free of all bacteria.
   4. cleansed of most bacteria on the skin surface.

9. One very simple nursing procedure that is considered the **most** important in preventing the spread of infectious diseases is

   1. isolation of all patients suspected of having an infection.
   2. wearing latex gloves when performing nursing procedures for surgical patients.
   3. washing the hands thoroughly before and after each contact with a patient.
   4. sterilizing the hands with a strong germicide at least once a day.

12. Avoid dependent items.

Poor:

Which of these medications is classified as a sodium-glucose co-transport inhibitor?

1. Sitagliptin (Januvia).
2. Canaglifozin (Invokana).
3. Dulaglutide (Trulicity).
4. Acarbose (Precose).

A side effect of the medication is

1. gastrointestinal discomfort.
2. injection site reaction.
3. urinary tract infection.
4. increased risk of peritonitis.

The Options

1. Make sure that there is one, and only one, correct or best answer.

2. Keep all of the options approximately equal in length.

3. Make all of the distractors plausible and attractive.

4. Avoid the use of all-of-the above or none-of-the-above.

    Poor:

    Which of the following situations could be offensive to a person who is ill?

    1. A nurse wearing strong cologne.
    2. A nurse who has just smoked a cigarette.
    3. A soiled dressing that has not been removed from the room.
    4. All of the above.

5. Use between three and five options for each item. The fewer the options, the greater the ability to answer correctly by guessing.

6. Develop options that are homogeneous.

    Poor:

    Which of these foods contains the highest amount of vitamin C?

    1. An orange.
    2. Pork.
    3. Beef.
    4. Legumes.

    Better options:

    1. An orange.
    2. A peach
    3. An apple
    4. A banana.

7. Avoid overly wordy options that become difficult and confusing to read.

8. Use qualifying terms related to amounts (e.g., *small, medium, large*), frequency (*always, never, often, rarely, usually*), recency (*earlier, later*) with caution.

9. Use words like "best," "most," "essential," "first," and "priority" in the stem *only* when the options indicate varying degrees of plausibility or applicability.

Poor:

When a client is receiving heparin sodium, it would be **most** important to have which of these drugs readily available?

1. Protamine sulfate.
2. Atropine sulfate.
3. Calcium gluconate.
4. Phytonadione (Mephyton).

Better:

Because a client is receiving heparin sodium, it would be important to have which of these drugs readily available?

1. Protamine sulfate.
2. Atropine sulfate.
3. Calcium gluconate.
4. Phytonadione (Mephyton).

10. When possible, present options in some logical order (e.g., chronological order, smallest to largest).

Poor:

A client is to receive 3,000 mL of fluid intravenously in 24 hours. His intravenous drop factor is 15 gtt/mL. His infusion should be regulated to deliver about how many of these drops per minute?

1. 31
2. 20
3. 53
4. 42

Better order of options:

1. 20
2. 31
3. 42
4. 53

11. If two options form a logical pair, make a pair of the other two options.

Poor:

A client is admitted to the hospital because of acute pancreatitis. The nurse should expect the client to have which of these blood test results?

1. Increased amylase.
2. Decreased amylase.
3. Increased calcium.
4. Decreased sugar.

Better options:

1. Increased amylase.
2. Decreased amylase.
3. Increased creatinine.
4. Decreased creatinine.

12. If distractors have more than one component, make all components wrong (Don't have distractors that are partially correct).

 Poor: A female client is being observed for symptoms of hyperthyroidism, which include

   1. lethargy, thin hair, and increased appetite.
   2. bradycardia, lethargy, and palpitations.
   3. brittle nails, thin hair, and lethargy.
   4. heat intolerance, amenorrhea, and tachycardia.

 Better:

 A female client is being observed for symptoms of hyperthyroidism, which include

   1. lethargy and thin hair.
   2. bradycardia and slow speech.
   3. brittle nails and masklike facies.
   4. heat intolerance and amenorrhea.

13. Make all options mutually exclusive. Numerical options should not overlap.

 Poor:

 The normal pH of arterial blood is between

   1. 7.25 - 7.35.
   2. 7.35 - 7.45.
   3. 7.45 - 7.55.
   4. 7.55 - 7.65.

 Better options:

   1. 7.20 - 7.30.
   2. 7.35 - 7.45.
   3. 7.50 - 7.60.
   4. 7.65 - 7.75.

14. Avoid irrelevant clues such as grammatical structure, well-known word associations, or connections between the stem and correct answer.

 Poor:

 The fruit that contains the highest amount of vitamin C is an

   1. orange.
   2. peach.
   3. apple.
   4. banana.

 Better:

 The fruit that contains the highest amount of vitamin C is a(n)

   1. orange.
   2. peach.
   3. apple.
   4. banana.

15. Avoid language that may offend or exclude a particular group of individuals.

16. Randomly distribute the correct answer among the alternative positions throughout the test, having approximately the same proportions of 1's, 2's, 3's, 4's as the correct answer.

Item Template

\#

_____ . _____

_____

_____

_____

_____

1. _____

_____

2. _____

_____

3. _____

_____

4. _____

_____

Key: _____

Blueprint Codes: _____

Reference: _____

_____

_____

References: Test Construction

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: AERA.

Accreditation Board for Specialty Nursing Certification. (2019). *Accreditation standards*. Mt. Laurel, NJ: Author.

Council on Licensure, Enforcement & Regulation. (2015). *Principles of fairness: An examination guide for credentialing & registration boards*. Lexington, KY: Author.

Lane, S., Raymond, M., & Haladnya, T.M. (Eds.) (2015). *Handbook of Test Development*, 2nd ed. New York: Routledge.

McDonald, M.E. (2018). *The Nurse Educator's Guide to Assessing Learning Outcomes*. 4th ed. Boston: Jones & Bartlett Publishers.

National Council of State Boards of Nursing. (2019). *NCLEX-RN Examination: Test Plan for the National Council Licensure Examination for Registered Nurses*, Effective April 1, 2019 - March 31, 2022. Chicago: NCSBN.

National Council of State Boards of Nursing. (2020). *NCLEX-PN Examination: Test Plan for the National Council Licensure Examination for Practical Nurses*, Effective April 1, 2020 -March 31, 2023. Chicago: NCSBN.

Oermann, M.H., & Gaberson, K.B. (2019). *Evaluation and testing in nursing education*. (6th ed.) New York: Springer.

.

APPENDIX A

Item Stems

<div align="center">**EXAMPLES OF TEST ITEM STEMS FOR AREAS OF THE NURSING PROCESS**</div>

I.   Assessment

To help establish a diagnosis of "x," the nurse should ask the patient/significant other which of these questions?

Which of the following information given by a patient indicates that s/he is at imminent risk for developing "x"?

When doing a physical assessment of a patient, the nurse should expect which of these findings related to "x"?

When a patient complains of "x," the nurse should expect him to describe "x" in which of these ways?

The nurse should assess a patient for symptoms of "x," which include

Considering a patient's diagnosis, the patient is most likely to have which of these symptoms?

A patient complains of "x." To help establish a diagnosis of _____, the nurse should ask which of these questions (or ..., the nurse should collect which of the following information)?

The physician orders all of the following for a particular patient. Which one should the nurse *question*?

II.   Analysis

A patient complains of "x, y, and z." Based on these findings, the nurse should suspect the patient has which of these conditions?

A patient makes all of the following comments. Which one definitely needs further discussion?

A patient complains about "x." To respond to the patient's concerns, the nurse should make which of these interpretations of his statements (or behavior)?

The nurse obtains all of the following data about "x." Which one should be reported?

Because of the nature of a patient's symptoms (i.e., "a, b, and c"), the nurse should ask which of these questions?

Because of a patient's diagnosis (or condition or signs and symptoms), which of these nursing diagnoses would be appropriate?

III.   Planning

Which of these nursing diagnoses should receive priority in a patient's care?

Given some data, which of these goals is appropriate in a patient's care?

Given some data, which of these goals should receive priority in a patient's care?

Because a patient has "x," his care plan should include which of these measures/notations?

To meet the goal of "x," which of these measures is indicated?

In preparing a patient for surgery, the teaching plan should include which of the following information?

The purpose of a measure/treatment is to

Because a patient has "x," the patient should be referred to which of these agencies/community resources?

Which of these actions could the RN delegate to a technician?

A registered nurse is assigned to care for the following patients. Which patient can be delegated to an LPN/LVN?

IV.   Implementing

A patient has "x." To promote the patient's comfort/optimal functioning, which of these actions should the nurse take? first? initially?

Because a patient has "x," which of these instructions should the patient be given (or ..., which of the following information should the patient receive)?

Which of these steps in a procedure or technique is correct? is incorrect?

A patient has "x." Because of this finding, which of these actions should the nurse take?

Which of these approaches would likely/most likely encourage the patient to comply with "x"?
A staff member takes all of the following actions (or includes all of the following measures) in a patient's care. Which one requires further discussion?

V.   Evaluating

A patient has received "x" (e.g., medication, treatment, nursing measure). The "x" has been effective if the patient has which of these findings (or, ... if which of these laboratory/blood values is increased or decreased)?

Which of these findings/observations/assessments of a patient receiving "x" would indicate that "x" is ineffective (or, ... would indicate that the patient's goal/therapy needs to be changed)?

A patient has been given instructions about "x." Which of these comments/statements by the patient indicates that s/he understands the instructions (or, ... indicates the need for further teaching)?

A patient has received "x" test. All of the following actions/steps were done. Which one indicates that the test was done correctly (or incorrectly)?

A patient should be observed for side effects of "x" drug, which include

APPENDIX B

Sample: Item Bias Review Form

Sample: Item Bias Review Form

Stereotyping and Inadequate Representation

Reviewer's Name: _____

Test: _____ Date: _____

Sex, Ethnic, Cultural, Religion & Class Bias	Item Number								
Does the test item…									
1. contain material that is inflammatory, controversial or emotionally charged for designated sub-groups of interest (DSI)?									
2. contain language or material that is demeaning or offensive to DSI?									
3. portray members of DSI in situations that do not involve authority or leadership?									
4. depict members of DSI as having stereotyped character-istics?									
5. depict members of DSI as having stereotyped occupa-tions?									
6. contain biased or offensive artwork?									
7. contain content that is different or unfamiliar to some DSI?									
8. measure knowledge expected of a nurse who has been practicing as a holistic nurse for a minimum of one year with 45 hrs of continuing education?									

Sex, Ethnic, Cultural, Religion & Class Bias	Item Number									
Does the test item…										
9. reflect information and/or skills that may not be expected to be within the educational or experiential background of all examinees?										
10. contain information that could benefit examinees of some DSI?										
11. contain some characteristic and/or feature that will lead DSI to answer the item correctly or incorrectly for the wrong reason?										
12. contain words that have different or unfamiliar mean-ings for different DSI?										
13. Is the test item free of group-specific language, vocabulary, or reference pronouns?										
14. Will any of the distractors be especially attractive to mem-bers of DSI for cultural reasons?										
15. Does the explanation concern-ing the nature of the task required tend to differ-entially confuse members or DSI?										
16. Does the test item contain clues that could facilitate the performance of members of some DSI?										
17. Does the format or structure of the test item present greater problems for nurses from some backgrounds than for others?										

Overall Judgments

Reviewer's Name: _____

Test: _____ Date: _____

Overall Judgments	Yes	No	Unsure
1. Does the test cover topics of interest and relevance to DSI where possible?			
2. Does the test as a whole represent DSI positively in nonstereotyped ways and settings?			
3. Does the test as a whole represent DSI in proportion to their incidence in the population?			
4. Is the test content balanced in terms of being equally familiar to all DSI?			

APPENDIX C

AHNCC Test Specifications
(Blueprints)

BLUEPRINT FOR THE EXAMINATIONS

The *HN-BC® Examination,* with a maximum of 170 items, is organized by Core Value and has the following approximate weightings in each area:

1. Holistic Philosophy, Theories, and Ethics (20%)
2. Holistic Nurse Self-Reflection, Self-Development, and Self- Care (9%)
3. Holistic Caring Process (37%)
4. Holistic Communication, Therapeutic Relationships, Healing Environments, and Cultural Care (23%)
5. Holistic Education and Research (11%)

The *HNB-BC® Examination,* with a maximum of 190 items, is organized by Core Value and has the following approximate weightings in each area:

1. Holistic Philosophy, Theories, and Ethics (20%)
2. Holistic Nurse Self-Reflection, Self-Development, and Self- Care (9%)
3. Holistic Caring Process (37%)
4. Holistic Communication, Therapeutic Relationships, Healing Environments, and Cultural Care (23%)
5. Holistic Education and Research (11%)

The *AHN-BC® Examination,* with a maximum of 150 items, is organized by Core Value and has the following approximate weightings in each area:

1. Holistic Philosophy, Theories, and Ethics (12%)
2. Holistic Nurse Self-Reflection, Self-Development, and Self- Care (3 %)
3. Holistic Caring Process (43%)
4. Holistic Communication, Therapeutic Relationships, Healing Environments, and Cultural Care (25%)
5. Holistic Education and Research (17%)

The *APHN-BC® Examination,* with a maximum of 250 items, is organized by Core Value and has the following approximate weightings in each area:

1. Holistic Philosophy, Theories, and Ethics (12%)
2. Holistic Nurse Self-Reflection, Self-Development, and Self- Care (3 %)
3. Holistic Caring Process (43%)
4. Holistic Communication, Therapeutic Relationships, Healing Environments, and Cultural Care (25%)
5. Holistic Education and Research (17%)