



IBLCE[®]

*International Board of
Lactation Consultant Examiners*

Guidelines for Writing Test Items for the IBCLC Certification Program: how to write questions for the IBCLC exam

Written by:
Leon Gross, PhD
Psychometric Consultant to IBLCE

Preface

Being successful, both professionally and personally, involves making decisions from a limited, well-defined set of choices. Which of these four job offers should I accept? Should I return to school for an advanced degree? Which of these five houses would be best for my family, considering cost, size, location, and school district? Should I (we) have another child?

These are more than important decisions – they are potentially life-altering events. We take these decisions seriously, as we should, and often find them agonizingly difficult to make. Former American baseball star and self-appointed humorist-philosopher Yogi Berra has said, “When you come to a fork in the road, take it!” Whilst funny, his advice is of no assistance. Just as the fork in the road provides the limited choices that confront the decision-maker – so difficult to resolve that Yogi makes a cognitively paralytic non-decision – making these major decisions is not facilitated by knowing the limited multiple choices from which we have to choose.

Why is this relevant to writing test items for the IBLCE certification exam? Very simply, this discussion is to make the point that a multiple-choice certification test is not an artifact or a dilutively convenient technique for assessing knowledge. Rather, multiple-choice testing is a valid and reliable methodology for determining how much aspiring professionals know, and whether that amount of knowledge is – or should be – considered sufficient for certification. In addition, multiple-choice tests are very capable of assessing not only the recall of knowledge, but more importantly for a health care profession, the *application* of knowledge.

This manual is written for the thousands of IBLCE certified lactation consultants who wish to develop and submit multiple-choice test items for use on the IBLCE certification test. Writing items involves knowing what to do, as well as what *not* to do. Both aspects are discussed here.

IBLCE sincerely hopes that you are interested in developing and submitting items, and that you will use this manual to help guide your item creations. Expect to find this effort more challenging than you anticipated, but also more fulfilling. It is also a wonderful way to give back to the profession. The Board of Directors of IBLCE thanks you for your effort.

Chapter 1: Anatomy of an Item

Consider the following item.

Multiple-choice items should contain:

- * a. at least three options
- b. at least one correct answer
- c. one answer that is obviously wrong
- d. one blank to be filled in

The portion of the item that poses the question (printed here in blue) is the *stem*. An item stem can be written as an incomplete statement, as shown in the above example, or as a complete question (Example: *Multiple-choice items should contain which of the following?*). Either style is acceptable. However, because stems are often written as incomplete statements, the unit is referred to as a multiple-choice item, rather than a question.

The choices are known as responses, answers, options, alternatives, or choices. The incorrect or wrong answers are known as *distractors* (printed here in orange). The right answer (printed here in green) is known as the *keyed* or *correct response*.

Every item should have 3-5 options (i.e., 2-4 distractors), only one of which is correct. Most items will have four options (i.e., three distractors). Three options (i.e., two distractors) are acceptable in limited circumstances when the options form a complete set. For example, the effect of a medication on lactation would be for it to increase, decrease, or have no effect on lactation. Again, most items will contain four options – three distractors and one correct response.

Chapter 2: Begin With An Idea – Developing The Item Stem

First, identify an important idea or concept that you wish to test. The item stem will reference this concept either directly or indirectly. Also, the stem usually determines the area of the examination grid (i.e., the scientific discipline and chronological period) in which the item should be classified.

The concept referenced by the stem should *not* be “nice to know” but rather important or consequential. Articulate for yourself why this concept is important for lactation consultants to know. What are the consequences for the mom, baby, or other aspect of breastfeeding if this concept is not known or is not understood? How does the knowledge of this concept enhance the skill and/or effectiveness of lactation consultants? When you can resolve these issues for yourself, you are ready to prepare the item stem.

In presenting the content issue to examinees, the stem is the most important part of an item. Therefore, the stem must be written clearly and contain sufficient information so that examinees know what type of response is expected. A properly written stem should enable examinees to answer the item, correctly or incorrectly, before reviewing the choices. In addition, the stem should not contain any blank to be filled in because such blanks sometimes impede clarity and disrupt continuity of the examinee’s thinking.

For example, consider the following poorly written item.

Canberra is:

- a. located in Tasmania
- * b. the capital of Australia
- c. known for its tall mountains
- d. experiencing a significant decrease in crime

The stem is poorly written. After reading it, examinees do not know what the item is asking them to answer about Canberra. Is the item asking about Canberra’s geography (a), demography (b), topography (c), or sociology (d)?

Examinees must read each of the choices in order to know what the item is asking. This is similar to asking someone to make a decision, but not informing that person what the decision is about. Another serious flaw is the lack of homogeneity among the choices, as noted above. Four items could be created from this one item, each posing a different question about Canberra. Consider the examples that follow.

Canberra is the capital of:

- a. New Zealand
- * b. Australia
- c. Indonesia
- d. Malaysia

Which of the following cities is the capital of Australia?

- a. Sydney
- * b. Canberra
- c. Melbourne
- d. Brisbane

The Australian Capital Territory in which Canberra is located is carved out of which of the following states?

- * a. New South Wales
- b. Queensland
- c. Tasmania
- d. Victoria

Each of these item stems has sufficient focus and information to prompt an answer. These items become easier to understand with regard to intent, but do *not* become easier to answer correctly. *All items should have sufficient focus and information in the stem to prompt an answer.*

The preceding items test recall of knowledge. Health professions certification examinations should also test *application* of knowledge, which is considered when conceptualizing the item and developing the stem. This issue is discussed later in Chapter 4.

Chapter 3: Identify Misconceptions – Listing The Distractors

For the multiple reasons discussed in Chapter 2, the item stem is the most important part of an item. However, listing well-conceived choices is also critical for an item to be valid. An item stem may be well-written and pose an important clinical problem solving situation. However, a single poorly conceived distractor may invalidate the item.

Despite being a standard psychometric term, “distractor” is a misnomer. Incorrect responses are *not* intended to distract examinees from the correct response. If wrong answers functioned to distract, they would be misleading and “tricking” examinees. That is *not* their purpose, and items should *not* be written to mislead or trick.

The preceding chapter noted that an item stem should have sufficient focus and information to prompt an answer from the examinees. Distractors should be written to represent the compelling, rational, logical *wrong* answers that examinees would provide, if there were no choices from which to choose. These logical wrong answers are the misconceptions that examinees have. Listing them as distractors neither misleads, nor does it make the item easy to correctly answer. Their inclusion makes a multiple-choice item more like an open-ended question, but with the numerous advantages associated with computer scoring and rigorous statistical analysis.

Consider the item from Chapter 2 that asks the state from which the Australian Capital Territory that includes Canberra is carved. Each of the four choices is an Australian state. Some examinees would be expected, logically, to write Queensland, Tasmania, or Victoria, if they had no choices from which to choose. Another version of the item asked which city is Australia’s capital. Each of the three distractors – Sydney, Melbourne, and Brisbane – is larger than Canberra, and a reasonable, logical *wrong* answer.

Consider the following simple arithmetic item.

$$10 \times 5 =$$

- a. 48
- b. 49
- * c. 50
- d. 51

Initially, the distractors might seem to be compelling, but they are not. There is no computationally logical way to obtain any of the distractors. Although they are numerically close to the correct response, they are logically and cognitively extremely distant. More logical, compelling but incorrect choices would include 2, 5, and 15, which represent calculation errors based on division (a), subtraction (b), and addition (d), respectively, rather than multiplication.

Consider the same choices, but change the stem to ask: “How many states comprise the United States?” In this version of the item, the original distractors of 48, 49, and 51 are very logical and compelling because of the varying number of states during various eras, the existence of 48 contiguous states, and the periodic effort to include Washington, DC as a 51st state. In this version of the item, numerical proximity and cognitive proximity coincide.

There are three distractor formats to avoid. These are “none of the above,” “all of the above,” and various other response combinations (e.g., “a and c above”). These distractor formats are psychometrically weak and therefore, are not used on IBLCE examinations. Be sure that the items you submit do *not* contain any of these option formats.

Chapter 4: Testing *Application Of Knowledge*

Items that test recall of knowledge are easier to develop than items that test application of knowledge. It is especially easy to write an item asking about a factual statement on a particular page of a particular book. Recall items, especially “book look-up” items, are also easier to reference. Although being easier to reference is a virtue, it is *not* by itself sufficient to make an item valid.

Recall items should test the important facts and principles that form the foundation of lactation consulting. Items that are merely nice to know or test knowledge of definitions are below entry level, and are *not* appropriate for the examination, despite the ease with which they can be referenced.

Although recall items are appropriate, certification tests need a substantial number of application items. Each day in the professional life of a lactation consultant involves working with moms and/or babies in novel situations that require thinking skills such as assessment, correlation, judgment, evaluation, analysis, and problem solving. As these are the skills that primarily determine competency in lactation consulting, the examination must test these skills in order to have adequate validity.

The most direct approach to writing application items is to think about situations that you have encountered professionally. What decision did you have to render? What additional information did you need? What was the best advice to offer the mom? Was the baby’s low weight gain indicative of a serious problem?

These and many other questions can be based on a clinical situation. You may write several items per situation, but with two caveats. First, IBLCE will use only one of the items on any one examination. That, however, is fine, as the database of test items must be many times larger than the length of the 200-item test, and the multiple items expand the database. Second, the items that are based on the same situation should be independent of each other.

After completing the questions for a particular situation, review the encounter to determine if it can be “cloned.” Can one or more characteristics about the mom (e.g., age, number of prior children, breastfeeding history, and/or onset and severity of reported problem), baby (e.g., age and/or onset and severity of reported problem), or any other “variable” in the situation be changed? Although modifying one or more variables somewhat fictionalizes the situation, if the modifications are realistic and within presentation limits, the situation retains its realism and basis for application items. The resultant items may need to vary from the original situation, and although the distractors and correct response may vary, the items retain their validity.

Another useful way of creating a clinical situation is to begin the stem with “You are” This beginning indicates something about being contacted by a mom (“You are contacted by a mom who is concerned about ... ”), working with a mom (“You are discussing a mom’s breastfeeding schedule because her ...”), or another type of encounter. In beginning the stem with “You are,” your story will flow. However, you should continually do a reality check to be sure that the situation is focused and of entry-level difficulty.

Items based on clinical situations often culminate in asking which of the following (WOTF) should you (i.e., the lactation consultant) do *first*, WOTF is *most* important (or *least* important), WOTF should you *not* do, or many other clinical decisions. Be sure that these questions are clinically distinguishable and appropriate. For example, if two or more choices are important to do, asking which one should be done *first* may not be appropriate. If asking for the *most* important action to take or advice to give, be sure that the criteria upon which *most* is based, is clear. For example, “WOTF is the *most* important food that the mom should add to her diet?” should include a qualifier (e.g., “in order to eliminate her Vitamin C deficiency”).

The distinction between recall and application items has a distinguished history in psychometrics. It was originally developed by Benjamin Bloom in a broader format to differentiate six levels of educational objectives. Five of the six levels represented different levels of applying knowledge. As the most important benefit of a taxonomic model for testing is to elevate items to a level above recall, the taxonomy is often presented in two levels: recall and application.

Chapter 5: Image-Based Items

Nearly 40% of the items in an IBLCE examination contain an image in the stem. Most of these images are digital color photographs depicting a breastfeeding problem relating to the baby or child (e.g., cleft palate), mother (e.g., mastitis), or mother-baby interaction (e.g., positioning). Image-based items assess application of knowledge, but have particular clinical relevance, since they present common breastfeeding problems as they are manifested to a lactation consultant.

Consider the wise saying, “A picture is worth a thousand words.” For an image-based item, the photo comprises most of the stem, and only complementary information (e.g., age of baby) may be necessary. Typically, the stem culminates in asking what condition is represented, what problem is depicted, what is causing the problem, what should be done about the problem, or how successful an intervention is likely to be. The options should be written as they are for other items; that is, by listing the incorrect responses that candidates would write if they had no options from which to choose.

All of the guidelines that apply to item writing are applicable to writing image-based items. This includes providing references, as discussed in the next chapter. However, image quality is an additional important factor. Images that accompany items should be well lit, photographed with high resolution, and of the overall quality that would be expected of a technical or medical publication. Also, submitted images should be accompanied by a completed consent form ([click here](#)). IBLCE provides Continuing Education Recognition Points (CERPs) for each image that meets technical specifications and is accompanied by a usable item and associated references.

Chapter 6: Item References

All test items should be referenced, ideally from literature published within the past five years. One challenge is that references are more elusive in documenting application items than recall items. Usually, documentation of the correct response in application items is general, whilst recall items have more specific references. The clinical situation upon which an application item is based is likely to be referenced by the depicted condition or problem in general, rather than by the specifics of the particular clinical situation. This is acceptable. Also, the allowance of a more general documentation facilitates referencing a cloned clinical situation, as long as the same condition or problem is being depicted.

There is one additional consideration in referencing items. The correctness and validity of an item may depend not only on being supported by 1-3 references, but also, by there being no controversy or support for one of the distractors in the literature. Item authors are not expected to exhaustively research the literature for certainty regarding an item's correctness; however, item authors should have a general awareness of controversy and conflicting professional philosophies. Test items should not favor one philosophy over another. If testing a controversial issue is important, it is best for the item to ask examinees about the *nature* of the controversy, rather than assuming a partisan position.

Chapter 7: Editorial and Technical Considerations

Nearly every item that is submitted is edited by IBLCE staff and psychometric consultants for consistency in format, enhanced text, and linguistic considerations prior to translation into more than 15 languages. Therefore, it is not necessary for you to serve as both an item author and editor. However, there are some stylistic and structure issues to consider in item development. These issues are discussed below with examples.

Limit the use of negatives to when necessary. Negative stems are more likely to be misread than positive stems, and are more likely to create translation difficulties. Consider the following example.

Foods that are high in vitamin B include all of the following *except*:

- * a. citrus fruits
- b. whole grains
- c. fish
- d. dairy

It is not necessary for this item to be negatively worded. A better item follows.

Foods that are high in vitamin B include:

- a. citrus fruits
- * b. whole grains
- c. tomatoes
- d. liver

A negative statement can sometimes be phrased in a positive way. For example, consider the question, “WOTF should the mom *not* do?” This would be stated better by asking, “WOTF should the mom avoid doing?”

Double negatives should *never* be used. They are unnecessary and are frequently confusing. For example, an item stating that “moms’ concern about an adequate milk supply is *not uncommon*,” is better stated as “many moms (or “some moms”) are concerned about having an adequate milk supply.”

The choices should be reasonably comparable in the length and amount of detail. Consider the following example.

When taking antibiotic medications, a new mom should avoid:

- a. going outside
- b. lifting the baby
- c. being with friends
- * d. worrying about the presence of the drug in her milk, because the small amount that is transmitted is generally less harmful to the baby than artificial milk

The correct response to this item is obvious, based on its relative length and level of detail. The item author regarded the item's concept is vital, but the item would fail to measure knowledge of the concept because nearly all examinees would answer it correctly, even if they did not know the concept. The item should be rewritten. A better version follows.

A mom who is taking antibiotic medication has been told by her physician that she should continue breastfeeding. However, she is not sure whether to continue. You should tell her that the antibiotic will:

- a. prevent the baby from getting sick
- b. increase her milk supply
- c. increase the baby's energy level
- * d. not harm the baby

Finally, the IBLCE Board of Directors thanks you again for developing and submitting test items. The Board is confident you will agree that stepping through this process is gratifying and a wonderful way of "giving back" to the profession.