

Criteria Checklist for Test Authors

Test: _____

Author: _____

Date: _____

Notes: Test authors can refer to this checklist as they write items, and then later as the author and other subject matter experts review the draft. The first three sets of criteria are generally applicable to all tests, while the last six sets apply to tests containing certain item types. Refer to previous pages in this document for explanations of some of the criteria.

Major Category/Criterion	OK?
1. Overall test	
a. Objectives for the test, and how the results will be used, are clearly defined.	<input type="checkbox"/>
b. Test blueprint has been written, and it maps the content for the test in terms of topics and levels.	<input type="checkbox"/>
c. Items in the test provide adequate coverage with respect to the blueprint (important areas not overlooked; extraneous areas not included).	<input type="checkbox"/>
d. If section scores are reported, there are at least 8 items for each section.	<input type="checkbox"/>
e. Test as a whole is a good measure of the knowledge and skills the training program (or other developmental activity) was intended to enhance.	<input type="checkbox"/>
2. Instructions	
a. Clear instructions for the overall test (e.g., how responses are recorded, penalty or not for guessing, time allowed).	<input type="checkbox"/>
b. Where necessary, clear instructions for certain item types (e.g., "mark all those that are correct" for multiple-select items, whether or not response options can be used more than once in a matching item).	<input type="checkbox"/>
c. It is clear to the test taker whether the focus is on selecting the <i>correct</i> answer or the <i>best</i> answer.	<input type="checkbox"/>
3. Criteria applicable to more than one item type	
a. Each item is focused on only <i>one</i> thought (knowledge or skill area).	<input type="checkbox"/>
b. Item is relevant and important – something people should know versus being able to look up; not trivial.	<input type="checkbox"/>
c. Difficulty level of the item is appropriate for the test takers.	<input type="checkbox"/>
d. Stem is a complete thought (i.e., the task is clear without having to look at the response options).	<input type="checkbox"/>

Major Category/Criterion	OK?
e. The bulk of the text is in the stem; the response options are as short as possible.	<input type="checkbox"/>
f. Syntax is clear and unambiguous; wording is concise and straightforward.	<input type="checkbox"/>
g. Low level of reading difficulty; unnecessarily big or unusual words are avoided.	<input type="checkbox"/>
h. Syntax doesn't preclude any of the response options; all are grammatically correct.	<input type="checkbox"/>
i. All distracters are plausible, and could be selected by less competent test takers.	<input type="checkbox"/>
j. Response options are independent (e.g., one isn't a subset of another, no opposites).	<input type="checkbox"/>
k. Response options are in a logical order (e.g., by size, alphabetical).	<input type="checkbox"/>
l. Correct response options are randomized with respect to position among distracters.	<input type="checkbox"/>
m. Items are independent of each other (e.g., content of one item doesn't help the test taker get some other item right).	<input type="checkbox"/>
n. There is a mix of items of different difficulty levels (some easier, some harder).	<input type="checkbox"/>
o. Within sections, or within items clustered by item type, the easier items come first.	<input type="checkbox"/>
p. Definitive terms like <i>always</i> , <i>all</i> , and <i>never</i> are avoided.	<input type="checkbox"/>
q. Where appropriate, items are clustered by type; don't have to frequently change types (e.g., T-F items appear together).	<input type="checkbox"/>
r. Correct response option and distracters are similar in terms of length, sophistication, etc.	<input type="checkbox"/>
s. Response options don't all repeat words that could be included in the stem.	<input type="checkbox"/>
t. Positive wording used when possible (e.g., use of NOT, EXCEPT, etc. minimized); negative words in all caps or underlined when they must be used.	<input type="checkbox"/>

Major Category/Criterion	OK?
u. No double negatives (ever).	<input type="checkbox"/>
v. No irrelevant clues as to the correct answer in stem or response options.	<input type="checkbox"/>
w. Items are not split between pages.	<input type="checkbox"/>
4. Multiple choice	
a. Only one correct answer; all distracters are definitely wrong (or definitely not as good as the correct response).	<input type="checkbox"/>
b. Either 4 or 5 response options (with fewer than 4, guessing is a problem; more than 5 adds little value and increases test time).	<input type="checkbox"/>
c. "All of the above" is avoided, and "None of the above" is used sparingly (e.g., when something is calculated).	<input type="checkbox"/>
d. No two response options have the same meaning (the savvy test taker could eliminate both).	<input type="checkbox"/>
e. "A and B," "A, B and D" response option logic is used only when you've given up trying to write a better item.	<input type="checkbox"/>
5. True-false	
a. Statement is worded precisely so it is unequivocally (but not obviously) true or false.	<input type="checkbox"/>
b. About 60% of the statements are false (to counterbalance the fact that people guessing are more likely to select true).	<input type="checkbox"/>
c. Vague terms that are open to interpretation (e.g., <i>frequently</i> and <i>sometimes</i>) are avoided.	<input type="checkbox"/>
d. True statements are no longer, no more specific, and no more correct sounding than false statements.	<input type="checkbox"/>
6. Multiple select	
a. At least 4, but not more than 10, response options.	<input type="checkbox"/>
b. At least one of the response options is correct (none correct would be tricky).	<input type="checkbox"/>

Major Category/Criterion	OK?
7. Matching	
a. All premises and response options are members of the same set (a similar category or topic).	<input type="checkbox"/>
b. Each premise has one, and <i>only one</i> , correct response option associated with it.	<input type="checkbox"/>
c. The longer text items (definitions, concepts, etc.) are premises on the left; the response options on the right are as brief as possible.	<input type="checkbox"/>
d. Irrelevant association cues between premises and options are eliminated.	<input type="checkbox"/>
e. Where appropriate, the number of premises and response options are not equal (e.g., more options than premises, or options can be used more than once).	<input type="checkbox"/>
8. Ranking	
a. At least 4, but not more than 10, response options.	<input type="checkbox"/>
b. There is unequivocally a correct order to the response options.	<input type="checkbox"/>
9. Fill in the blank	
a. Only one or two blanks for each item.	<input type="checkbox"/>
b. Correct responses are single words, very short phrases, or values.	<input type="checkbox"/>
c. A priori decisions are made about which responses will be scored as correct (e.g., misspelled words; will \$1,800, \$1800, and \$1,800.00 all be correct?).	<input type="checkbox"/>