Tool 29-1. Test-Writing Checklist

The following checklist provides the basis for documenting each of the steps of the CRT process and is linked to the CRT development model in Figure 29-1.

What follows is a list of basic questions you should consider addressing as you document your test development process. (Note: The authors are not attorneys and do not claim these questions to be legal requirements. Consult with your organization's legal staff if you have any concerns.)

Purpose of the Test	
What is the need for the test?	
How was the need determined?	
Who sponsored the effort? Why were they the sponsor(s)?	
Analyze Job Content	
What job/duty/task is being analyzed?	
If a duty or a task, of what job(s) is it a part?	
What is the hierarchical relationship of job/duty/task elements? Include a copy of the analysis.	
Who performed the job analysis?	
Establish Content Validity of Objectives	
For what job/duty/task were the objectives derived?	
In what course(s) are they covered?	
What are the names, titles, and credentials of the subject matter experts who validated the objectives?	
When was the validation performed? Include a copy of the objectives and the job analysis.	
Create Cognitive Items/Rating Instruments	
For what job and objectives were the items/rating instruments created?	
For what course is the test intended?	
Who wrote the items/rating instruments?	
What are the titles and credentials of those who wrote the items/rating instruments?	
When were the items/rating instruments written?	
Include a copy of the items/rating instruments.	

Establish Content Validity of Items and Instruments

For what job and objectives were the items/rating instruments created?	
For what course is the test intended?	
Who validated the items/rating instruments?	
What are the titles and credentials of those who validated the items/rating instruments?	
When were the items/rating instruments validated?	
Include a copy of the form on which subject matter experts indicated the match between the items/rating instruments and the objectives or the job elements.	

Initial Test Pilot

When was the test pilot conducted?	
Where was the pilot conducted?	
Who conducted the pilot?	
Who were the sample test takers?	
How were the sample test takers chosen?	
What were the pertinent characteristics of the sample test takers?	
What was the process used to conduct the pilot?	
What changes were made to the test as a result of the pilot?	

Item Analysis

Who were the test takers whose test data were used in the item analysis?	
What were the relevant characteristics of this sample of test takers?	
How were the sample test takers chosen?	
When were the item analysis data collected? By whom? Where?	
What program was used to analyze the test data?	
What were the results of the item analysis?	
What changes were made to the test as a result of the item analysis?	
Include a copy of the item analysis printout.	

Parallel Forms and Item Banks

For what job/objectives were the parallel forms or item banks created?	
How many parallel forms were created? When were they created?	
How was a decision made about the number of parallel forms to be created?	
Who made this decision?	
What software was used to create the item banks? Who chose the software and why was it chosen?	
What are the sizes and structure of the test item banks?	
How are items selected from the banks?	
What process was used to create the parallel forms?	
What is the equivalence reliability of the parallel forms? How was it established?	

Cut-Off Scores

How is the test scored?	
What is the cut-off score for the test?	
When was the cut-off score determined?	
What procedure(s) was (were) used to set the cut-off score?	
Include all data pertinent to the procedure(s) used, including a description of test takers and how they were chosen, judges and their credentials, item difficulty estimates, stakeholders and their suggested cut-off scores.	
Who decided what the cut-off score would be? What are their titles and credentials for making this decision?	

Reliability of Cognitive/Performance Tests

What procedures were used to establish the reliability of the test/raters?	
When were the data collected upon which the reliability was calculated?	
What is the reliability of the test/raters?	
Include all pertinent data used in the calculation of the reliability coefficient.	

Report Scores

To whom are scores reported?	
In what form are the scores distributed?	
Are the scores accompanied by course means? Medians? Other descriptive statistics?	
Are reported scores composites of sub-scores? If so, what are they and are they differentially weighted? If so, how? How were the weights determined?	
What guidance regarding the use of the scores is provided with them?	