

How do I conduct an Item Analysis on a Test?

What is an item analysis and what can I learn from it?

How do I interpret a Standard Training Activity Support System (STASS) Test Analysis Report?

How do I conduct an item analysis on a test?

What is an item analysis and what can I learn from it?

Through an item analysis, each item on a test and the entire test as a whole unit are evaluated based on the student responses to the items. As a result, it is possible to determine which items are poorly written and which items do not completely assess their learning objective. If an item is poorly written, a student may be able to select the correct response without knowing the material.

How do I interpret a Standard Training Activity Support System (STASS) Test Analysis Report?

A Standard Training Activity Support System (STASS) Test Analysis Report (CSM0510R) is generated by some commands for each test that is given. The information on this report is useful in conducting an item analysis. For each item, the report details:

- The amount of students that selected each alternative.
- The alternatives selected by the students in the upper 27%.
- The alternatives selected by the students in the lower 27%.
- The correct alternative.
- The difficulty index.

The following is an abbreviated example of a STASS Test Analysis Report.

STANDARD TRAINING ACTIVITY SUPPORT SYSTEM TEST ANALYSIS REPORT									
CIN	COURSE SHORT TITLE								
POS	RANGE	A	B	C	D	OMIT	DIFF	QUESTION	
1.	TOTAL	0	14*	0	0	0	1.00	B0044 01 02 02 003 7	
	UPPER	0	3	0	0	0			
	LOWER	0	3	0	0	0			
2.	TOTAL	0	0	0	13*	1	0.93	B0044 01 02 02 003 4	
	UPPER	0	0	0	3	0			
	LOWER	0	0	0	3	0			
3.	TOTAL	0	2	9*	3	0	0.64	B0044 01 02 02 005 1	
	UPPER	0	0	3	0	0			
	LOWER	0	1	1	1	0			

Note: Only items 1 -3 are shown. The rest of the items for the test are not displayed.

The **POS** or position number refers to the sequential number (or position) of the item on the test.

The **RANGE** identifies the part of the sample for which the information is displayed. The three ranges reported are total students (TOTAL), the upper 27% (UPPER), and the lower 27% (LOWER). The TOTAL includes the UPPER, the LOWER, and the middle 46% of the students taking the test.

- For item #1 in the example above, 14 (0+14+0+0) students answered the item. Of the 14, three students were in the upper range, and three students were in the lower range.
- For item #2 in the example above, 13 (0+0+0+13) students answered the item. Of the 13, three students were in the upper range, and two students were in the lower range.
- For item #3 in the example above, 14 (0+2+9+3) students answered the item. Of the 14, three students were in the upper range and three students were in the lower range.

A, B, C, and **D** represent the four alternatives for a multiple-choice item. Each column lists the number of students who selected that alternative for the item. An asterisk (*) indicates the correct alternative.

For item #3 in the above example, nine students selected alternative C, the correct alternative. Two students incorrectly selected alternative B and three students incorrectly selected alternative D. All of the students in the upper 27% of the test, and one student in the lower 27% selected alternative C. However, one student in the lower 27% of the test incorrectly selected alternative B and another lower 27% student incorrectly selected alternative D.

OMIT is the number of students who did not respond to (omitted) the item.

For items #1 and #3 in the above example, all of the students answered the item. However, the TOTAL row for item #2 shows that one student did not answer this item. This same student is also in the lower 27%, because there is a 1 in the LOWER row.

DIFF is the difficulty index of the item. The difficulty index refers to the proportion of students who answered the item correctly. The difficulty index is calculated by dividing the number of students who answered the item correctly by the total number of students taking the test. The difficulty index ranges from 0.00 (no one answered item correctly) to 1.00 (everyone answered item correctly). Thus, the **larger** the difficulty index, the **easier** the item. The acceptable range of difficulty for technical training is .50 to .90. Sometimes a difficulty of 1.00 may be desirable, such as in the area of safety, where it is crucial that everyone knows the information.

- For item #1 in the example above, all 14 students selected the correct alternative. Thus 14 divided by 14 is 1.00, the difficulty index for item #1.
- For item #3 in the example above, nine students answered the item correctly. Thus nine divided by 14 is .64, the difficulty index for item #3.

QUESTION is the item number of the item from the test bank.

(Source: CNET/USN. **CSM 0510R TEST ANALYSIS REPORT**. Retrieved August, 17, 2001 from http://www.cnet.navy.mil/netpdtc/intrpd/stasshelp/csm_rpt/csm0510r.htm.)

How do I conduct an item analysis on a test?

There are four major steps in conducting an item analysis.

1. [Identify the test conditions.](#)
Test conditions may include situations/conditions, which may have influenced the results of the test.
2. [Conduct an item analysis by interpreting the results from a Standard Training Activity Support System \(STASS\) Test Analysis Report.](#)
You will answer nine questions that will assist you in finding patterns in the data. These patterns will help you evaluate each item on the test.
3. [Interpret the results from the item analysis.](#)
From the nine potential patterns identified in step 1, you will identify three major patterns.
4. [Identify areas for course review and student performance from the interpretation of the item analysis.](#)
From the three major patterns, you will identify opportunities for course review and student performance improvements.

1. Identify the test conditions.

Test conditions may include situations and conditions that may have influenced the results of the test. In this section, provide a summary of the administration of the specified test. Please indicate location of the test administration, the number of people tested and any unusual situations and conditions that occurred that may have interfered with the administration of the test. These unusual situations and conditions may or may not have caused interruptions or cancellations to the test. Such occurrences may include:

- Emergency situations requiring the evacuation of the testing room.
- Outside disruptions.
- Unauthorized personnel entering the testing room during the test.
- A test taker being disruptive during the test.
- A test taker becoming ill or otherwise having to leave the room during the test.
- A test taker deviating from acceptable conduct during the test.

Test conditions are important because they may influence the results of the test. As a result, the item analysis may be slightly affected but results in an “inaccurate” test. It is important to constantly review the test conditions as you are interpreting the item analysis.

2. Conduct the item analysis.

When conducting the item analysis in “Charting a Course”, answer the nine questions to guide the evaluation of the STASS Test Analysis Report. These questions are related to the test and the test items. The answers to these questions serve as the starting point for the analysis and revision of tests and their items.

- Question 1. [Is there a pattern of correct answers throughout the test?](#)
Question 2. [Which items, if any, were missing by all or almost all of the students?](#)
Question 3. [Which items, if any, were omitted by a high percentage of students?](#)
Question 4. [Which items, if any, have an incorrect alternative chosen by a high percentage of high-achieving students?](#)
Question 5. [Which items, if any, have a difficulty index smaller \(closer to 0.00\) than the other items for this objective?](#)
Question 6. [Which items, if any, were answered correctly by more low-achieving students than high-achieving students?](#)
Question 7. [Which items, if any, were missed by none or almost none of the students?](#)

- Question 8. [Which items, if any, have incorrect alternatives that were NOT selected?](#)
 Question 9. [Which items, if any, have a difficulty index larger \(closer to 1.00\) than the other items for this objective?](#)

Question 1. Is there a pattern of correct answers throughout the test?

Look at the correct answers to make sure that they do not form a predictable pattern. For instance, when using an answer sheet that will be optically marked or graded, do the answers form diagonal lines, straight lines, and/or curves?

What Might This Mean?	Possible Solution
The students may be able to figure out this pattern. If this is the case, the students may be tested on their ability to figure out the pattern, rather than their understanding of the material.	Rearrange the order of the alternatives and/or the items. This should remove any patterns, but it is important to look at the answer sheet to identify any new patterns that may result from the rearrangement.

Question 2. Which items, if any, were missed by all or almost all of the students?

Look at the STASS Test Analysis Report and look for items in which the difficulty index is very small (closer to 0.00) and/or the correct alternative for the item was rarely selected.

Example:

POS	RANGE	A	B	C	D	OMIT	DIFF	QUESTION
2.	TOTAL	2	2	6	1*	3	0.07	Z0044 01 02 02 003 4
	UPPER	1	0	2	1	2		
	LOWER	1	2	1	0	1		

In the example above, the DIFF index is 0.07 and only 1 student selected alternative D, the correct alternative.

What Might This Mean?	Possible Solution
<p>This can happen for multiple reasons:</p> <ul style="list-style-type: none"> • The item may have been miskeyed. • There may have been more than one correct answer. • There may not be a correct answer. • This material may not have been covered in the instruction. • The item may have been written at a higher K-level than the objective being tested. • The item does not assess the objective. • The instructional material may not have adequately addressed this objective. • The students may not have had the prerequisite skills for this instruction. 	<p>The responses to the other questions and reviewing the item with content experts will help you pinpoint the reason or reason(s) why all or almost all the students missed the item. Revision of the item may be necessary. The next section (3. Interpreting the Item Analysis) will guide you through this review process.</p>

<ul style="list-style-type: none"> • The students may not have had adequate practice. • The instruction may have been too far removed in time from the test. • The item may have been written at too high a reading level. • The wording on the item stem may have been confusing. • The item alternatives may have been confusing. • The students did not have enough time to complete the item. • The difficulty index for the learning objective is closer to 0.00. 	
---	--

Question 3: Which items, if any, were omitted by a high percentage of students?

Look at the STASS Test Analysis Report and look for items in which the OMIT column has more than one or two in the TOTAL row. Additionally, if a large percentage of OMITs fall in the UPPER range, this may further illustrate that a problem exists.

Example:

POS	RANGE	A	B	C	D	OMIT	DIFF	QUESTION
2.	TOTAL	2	2	6	1*	3	0.07	Z0044 01 02 02 003 4
	UPPER	1	0	2	1	2		
	LOWER	1	2	1	0	1		

In the example above, the OMIT column has 3 in the total row indicating that three of the 14 persons taking the test omitted this question. Out of the three persons, two persons were in the UPPER range.

What Might This Mean?	Possible Solution
<p>This can happen for multiple reasons:</p> <ul style="list-style-type: none"> • This material may not have been covered in the instruction. • The item may have been written at a higher K-level than the objective being tested. • The item does not assess the objective. • The instructional material may not have adequately addressed this objective. • The students may not have had the prerequisite skills for this instruction. • The instruction may have been too far removed in time 	<p>The responses to the other questions and reviewing the item with content experts will help you pinpoint the reason or reason(s) why all or almost all the students omitted this item. The next section (3. Interpreting the Item Analysis) will guide you through this review process.</p>

<p>from the test.</p> <ul style="list-style-type: none"> • The students may not have had adequate practice. • The item may have been written at too high a reading level. • The wording on the item stem may have been confusing. • The item alternatives may have been confusing. • There may not be a correct answer. • There was not enough time to answer the item. • The difficulty index for the learning objective may be closer to 1.00. 	
---	--

Question 4: Which items, if any, have an incorrect alternative chosen by a high percentage of high-achieving students?

Look at the STASS Test Analysis Report and look for items in which an incorrect alternative has been overly selected by the students in the upper 27%. Approximately the same number of students should have selected each incorrect alternative. If this is not the case, then there may be a problem.

Example:

POS	RANGE	A	B	C	D	OMIT	DIFF	QUESTION
2.	TOTAL	2	2	7	1*	2	0.07	Z0044 01 02 02 002 4
	UPPER	1	0	4	1	0		
	LOWER	1	1	1	0	2		

In the example above, four of those in the UPPER range responded to option C an incorrect response.

What Does This Mean?	Possible Solution
<p>This can happen for multiple reasons:</p> <ul style="list-style-type: none"> • The item may have been miskeyed. • There may have been more than one correct answer. • The instructional material may not have adequately addressed this objective. • The students may not have had adequate practice. • The item may have been written at too high a reading level. • The instruction may have been too far removed in time from the test. 	<p>The responses to the other questions and reviewing the item with content experts will help you pinpoint the reason or reason(s) why all or almost all the high-achieving students choose a particular incorrect alternative. Revision of the item may be necessary. The next section (3. Interpreting the Item Analysis) will guide you through this review process.</p>

<ul style="list-style-type: none"> • The wording on the item stem may have been confusing. • The item alternatives may have been confusing. 	
---	--

Question 5: Which items, if any, have a difficulty index smaller (closer to 0.00) than the other items for this objective?

Look at the STASS Test Analysis Report and compare the difficulty indexes of items that measure the same objective. The difficulty indexes should be approximately equal. Look for items in which the difficulty index is closer to 0.00 than any other item for the objective.

Example:

POS	RANGE	A	B	C	D	OMIT	DIFF	QUESTION
3.	TOTAL	2	2	3	13*	0	0.65	Z0044 01 02 02 003 3
	UPPER	1	0	0	4	0		
	LOWER	1	1	1	2	0		
4.	TOTAL	2	2	4*	10	2	0.20	Z0044 01 02 02 003 4
	UPPER	1	1	2	1	0		
	LOWER	1	0	1	3	1		
5.	TOTAL	12*	2	4	1	1	0.60	Z0044 01 02 02 003 5
	UPPER	2	0	2	1	0		
	LOWER	1	2	1	0	1		
6.	TOTAL	14*	2	2	1	1	0.70	Z0044 01 02 02 003 6
	UPPER	3	0	2	0	0		
	LOWER	1	2	1	1	0		

The above example shows four items on a test that come from the same objective. Item 4 has a difficulty index (DIFF = 0.20) that is smaller than the other items for the objective.

What Might This Mean?	Possible Solution
<p>This can happen for multiple reasons:</p> <ul style="list-style-type: none"> • The item may have been miskeyed. • There may have been more than one correct answer. • There may not be a correct answer. • The item may have been written at a higher K-level than the objective being tested. • The item may be written at a different level than the other items. • The instructional material may not have adequately addressed this objective. • The students may not have had the prerequisite skills for this instruction. • The students may not have had adequate practice. 	<p>The responses to the other questions and reviewing the item with content experts will help you pinpoint the reason or reason(s) why the difficulty index of this item is smaller than the other items for the objective. The next section (3. Interpreting the Item Analysis) will guide you through this review process.</p>

<ul style="list-style-type: none"> • The item may have been written at too high a reading level. • The wording on the item stem may have been confusing. • The item alternatives may have been confusing. 	
--	--

Question 6: Which items, if any, were answered correctly by more low-achieving students than high-achieving students?

Look at the STASS Test Analysis Report and look for items in which the correct alternative has been selected by more students in the lower 27% than by students in the upper 27%. If this is not the case, then there may be a problem.

Example:

POS	RANGE	A	B	C	D	OMIT	DIFF	QUESTION
3.	TOTAL	3	2	7*	7	1	0.35	Z0044 01 02 02 003 3
	UPPER	1	0	0	4	0		
	LOWER	1	1	3	0	0		

In the above example, 3 persons in the LOWER range answered option C which is the correct answer. No one in the UPPER range answered this option.

What Might This Mean?	Possible Solution
<p>This can happen for multiple reasons:</p> <ul style="list-style-type: none"> • The item may have been miskeyed. • There may have been more than one correct answer. • Students may be “reading too much into the question.” • The wording on the item stem may have been confusing. • The item alternatives may have been confusing. 	<p>The responses to the other questions and reviewing the item with content experts will help you pinpoint the reason or reason(s) why more low-achieving students than high-achieving students answered this item correctly. The next section (3. Interpreting the Item Analysis) will guide you through this review process.</p>

Question 7: Which items, if any, were missed by none or almost none of the students?

Look at the STASS Test Analysis Report and look for items in which the difficulty index is very large (closer to 1.00) and/or the correct alternative for the item was almost always selected. Also, look at the students’ selection of alternatives and note any patterns of incorrect alternatives. There may or may not be a problem here.

Example:

POS	RANGE	A	B	C	D	OMIT	DIFF	QUESTION
3.	TOTAL	0	0	0	20*	0	1.00	Z0044 01 02 02 003 3
	UPPER	0	0	0	5	0		
	LOWER	0	0	0	5	0		

In the above example, the difficulty index (DIFF) of the item is 1.00 indicating no one missed this item.

What Might This Mean?	Possible Solution
<p>This can happen for multiple reasons:</p> <ul style="list-style-type: none"> • The item may have been compromised. • The incorrect alternatives may have been implausible. • The correct alternative may have contained a “give-away.” • Another question on the test may have answered this one. • The difficulty index for the learning objective may be closer to 1.00. 	<p>In some situations (such as assessing objectives with a high safety concern) it is ideal that all students are able to correctly respond to the question; however, you also want to make sure the reason they are able to respond to the question is because they have the knowledge and skills. Reviewing the question will help you assess whether one of the other reasons is occurring. The next section (3. Interpreting the Item Analysis) will guide you through this review process.</p>

Question 8: Which items, if any, have incorrect alternatives that were NOT selected?

Look at the STASS Test Analysis Report and look for items in which all of the incorrect alternatives have not been selected. Approximately the same number of students should have selected each incorrect alternative. If this is not the case, then there may be a problem, especially if the difficulty index is nearer to 0.00 than 1.00.

Example:

POS	RANGE	A	B	C	D	OMIT	DIFF	QUESTION
3.	TOTAL	2	15*	0	2	0	0.75	Z0044 01 02 02 003 3
	UPPER	0	2	0	3	0		
	LOWER	1	3	0	1	0		

In the example above, no one choose option C.

What Might This Mean?	Possible Solution
<p>This can happen for multiple reasons:</p> <ul style="list-style-type: none"> • This incorrect alternative may have been implausible. • This incorrect alternative may have a different format from the other alternatives. • The correct alternative may have contained a “give-away.” • Another question on the test may have answered this one. • This item alternative may have been 	<p>In some situations (such as assessing objectives with a high safety concern) it is ideal that all students are able to correctly respond to the question; however, you also want to make sure the reason they are able to respond to the question is because they have the knowledge and skills. Reviewing the question will help you assess whether one of the other reasons is occurring. The next section (3. Interpreting the Item Analysis) will guide you through this review process.</p>

confusing.	
------------	--

Question 9: Which items, if any, have a difficulty index larger (closer to 1.00) than the other items for this objective?

Look at the STASS Test Analysis Report and compare the difficulty indexes of items that measure the same objective. The difficulty indexes should be approximately equal. Look for items in which the difficulty index is closer to 1.00 than any other item for the objective.

Example:

POS	RANGE	A	B	C	D	OMIT	DIFF	QUESTION
3.	TOTAL	2	2	3	13*	0	0.65	Z0044 01 02 02 003 3
	UPPER	1	0	0	4	0		
	LOWER	1	1	1	2	0		
4.	TOTAL	0	19*	1	0	0	0.95	Z0044 01 02 02 003 4
	UPPER	0	5	0	0	0		
	LOWER	0	5	0	0	0		
5.	TOTAL	12*	2	4	1	1	0.60	Z0044 01 02 02 003 5
	UPPER	2	0	2	1	0		
	LOWER	1	2	1	0	1		
6.	TOTAL	14*	2	2	1	1	0.70	Z0044 01 02 02 003 6
	UPPER	3	0	2	0	0		
	LOWER	1	2	1	1	0		

The above example shows four items on a test that come from the same objective. Item 4 has a difficulty index (DIFF = 0.95) that is larger than the other items for the objective.

What Might This Mean?	Possible Solution
<p>This can happen for multiple reasons:</p> <ul style="list-style-type: none"> • The item may have been compromised. • The incorrect alternatives may have been implausible. • The correct alternative may have contained a “give-away.” • Another question on the test may have answered this one. • The item may be written at a different K-level than the other items for the objective. 	<p>The responses to the other questions and reviewing the item with content experts will help you pinpoint the reason or reason(s) why the difficulty index of this item is larger than the other items for the objective. The next section (3. Interpreting the Item Analysis) will guide you through this review process.</p>

3. Interpret the item analysis.

Once the item analysis has been conducted, the results from the nine questions must be reorganized and interpreted. The results are organized into four sections: 1) [Test Conditions](#), 2) [Step 1: Length of time allotted to take test](#), 3) [Step 2: Items missed by most or almost all the students](#), 4) [Step 3: Items missed by none or almost none of the students](#)

Test Conditions

The reported test conditions are displayed to provide a context for interpreting the results of the item analysis. The test conditions may affect the results of the item analysis and it is important to consider them.

Step 1: Length of time allotted to take test

The results from questions 2 and 3 from the item analysis are used to help assess whether the students had an enough time to complete the test. By looking at the results from these questions determine if there are questions at the end of the test or scattered throughout the test that the students are skipping.

What Might This Mean?	Possible Solutions
Students may not have had an enough time to finish the test. It could be that they are either spending too much time on earlier questions or there are too many questions on the test.	<ul style="list-style-type: none"> • Provide more time to students taking the test • Evaluate the order of the items to determine if there are items positioned earlier in the test that students may be spending excessive amount of time on. • Have students time how long it takes them to complete each question or each section of the test.

Step 2: Items missed by most or almost all the students

The results from questions 2, 4, 5, and 6 from the item analysis are used to help assess the reasons why most or almost all the students are missing certain items. As indicated in [Conducting Item Analysis](#), there can be multiple reasons why a particular item is flagged based on these questions. In this section, items are reviewed by content experts to pinpoint the probably reasons and determine whether the item needs to be revised.

For the items in this section, you will need to 1) [check the items for any miskeys](#), 2) [review the items with a panel of content experts](#), and [indicate whether any items need revisions](#).

First, check the items for any miskeys. Prior to reviewing the items, you will want to check for any potential miskey. Make sure that the correct alternative is correctly marked on the answer key. If an item has been miskeyed, you will need to: 1) correct the miskey and 2) review the revised item statistics to see if it is flagged based on any of the other item analysis questions. If after fixing the miskey the item is no longer flagged, then the item does not have to be reviewed.

Second, review the items with a panel of content experts. Conduct an item review meeting for the flagged items; evaluate each item based on the Quality Checklist for Multiple-Choice (MC) Item Construction. Answering the following questions will help guide you through the review.

Review Question	Review Focus
Is there more than one correct answer or Is there not a correct answer?	<ul style="list-style-type: none"> • Look at the student selection of alternatives and note if there are a pattern within the incorrect alternatives. • Review the alternatives to ensure that there is clearly one correct alternative and the other alternatives are definitely incorrect.
Is the wording of the item	<ul style="list-style-type: none"> • Review the stem, alternatives, and item as a whole to

Review Question	Review Focus
confusing?	<p>make sure the item does not use any confusing vocabulary and/or statements.</p> <ul style="list-style-type: none"> Review the stem and alternatives to ensure that they are clearly stated and written. Review each alternative to ensure that they are similar in grammar, format, and content. Review each alternative to make sure it fits well with the stem. Verify that the terminology used in the item is the same terminology used in the course.
Has the item been written at too high of a reading level?	<ul style="list-style-type: none"> If the learning objective does NOT require a specified reading skill or level, verify the reading level of the item is below the reading level of the students in the course. If the learning objective requires a specific reading skill or level, verify that the item is assessing at and not above the specified level.
Does the item assess the intended objective? Was the item written at a k-level higher than the objective being tested? (If item was flagged for question 2 and/or question 5)	<ul style="list-style-type: none"> Compare the item with the learning objective and make sure the item assesses the appropriate information covered by the learning objective. Review the item to verify that it is assessing the learning objective at the correct k-level. Conduct a k-level alignment in the MC Item Construction Tab, if necessary.
Are some of the students “reading too much into the question?” (if item was flagged for question 4 and/or question 6)	<ul style="list-style-type: none"> Look at the student selection of alternatives and not if there are patterns of incorrect alternative. Review the alternatives to ensure that there is clearly one correct alternative and the other alternatives are definitely incorrect.

When working with tests that have higher stakes (high consequences for performing poorly on the test), try to have a panel of at least five experts reviewing the item. It is recommended to have the experts review the item as a group so that the item can be discussed from different perspectives and all can agree on any revisions that may be needed.

Finally, indicate whether any items need revisions. When completing the item review meeting, indicate the items that need revision in the appropriate space in step two of the Interpreting Item Analysis screen.

Step 3: Items missed by no or almost none of the students

The results from questions 7, 8 and 9 from the item analysis are used to help assess the reasons why no or almost none of the students missed certain items. Ideally, this may have occurred because each student clearly understood the objective; however prior to making that interpretation, verify that either the test has not been compromised or there is nothing within the item or the test that is giving away the answer to the question.

Like in Step 2, this section is used to determine which items, if any, need to be reviewed and/or revised by content experts. For the items in this section, you will need to 1) [determine if any items](#)

[have been compromised](#), 2) [determine if items on the test give away the answer to any of the flagged question](#), 3) [review the items with a panel of content experts](#), 4) [indicate whether any items need revisions](#).

First, determine if any items have been compromised.

Test compromise may be suspected when the difficulty index of an item changes drastically from the previous offerings of the course. For example, if the difficulty index of an item is usually 0.5 (1/2 of the trainees answer the item correctly) and on one test, all the trainees answer the item correctly (difficulty index (DIFF) = 1.00), compromise of the item may be suspected.

If compromise is suspected, the MC item should be marked as “Suspected” for “Status on Compromise” in the MC Item Construction tab. Items that are suspected of being compromised should be taken out of the item pool when marked as suspected. These items can be replaced with different items. Retaining the suspected item ensures that a similar item will not be added to the active test bank.

Second, determine if items on the test give away the answer to the flagged question.

Review all the items on the test and see if any of them provide a clue that gives a-way the answer to the item under review. If so, you should 1) remove item in question, 2) replace the item with another item in the item bank covering the same Learning Objective, 3) review the new item to verify that it does not provide clues to other items on the test and visa versa.

Third, review the items with a panel of content experts.

Conduct an item review meeting for the flagged items; evaluate each item based on the Quality Checklist for Multiple-Choice (MC) Item Construction. Answering the following questions will help guide you through the review.

Review Question	Review Focus
Are the incorrect alternatives plausible?	<ul style="list-style-type: none"> Review the alternatives to ensure that the incorrect options are plausible but incorrect.
Is there a “give-away” in the alternatives or other part of the item indicating the correct answer?	<ul style="list-style-type: none"> Review the stem and alternatives to ensure the same word does not appear both in the stem and the alternatives. Review the alternatives to see if they are the same in grammar, format, and content.
Does the incorrect alternative have a different format from the other alternatives or one of the other alternatives provides a clue that eliminates the alternative? (If item was flagged for question 8)	<ul style="list-style-type: none"> Review the stem and the alternatives to ensure the same word does not appear both in the stem and the alternatives. Review the alternatives to see if they are the same in grammar, format, and content. Make sure that each alternative fits in well with the stem.
Is the item’s alternative confusing? (If item was flagged for question 8)	<ul style="list-style-type: none"> Review the alternative to make sure the item does not use any confusing vocabulary and/or statements. Review alternatives to ensure that they are clearly stated and written. Make sure that each alternative fits in well with the stem. Verify that the terminology used in the item is the same terminology used in the course.

Review Question	Review Focus
Is the item written at a k-level different than the other items in the objective	Compare the item being reviewed with the learning objective and make sure the item assesses the appropriate information covered by the learning objective. Conduct a k-level alignment, if necessary.

When working with tests that have higher stakes (high consequences for performing poorly on the test), try to have a panel of at least five experts reviewing the item. It is recommended to have the experts review the item as a group so that the item can be discussed from different perspectives and all can agree on any revisions that may be needed.

Finally, indicate whether any items need revisions.

When completing the item review meeting, indicate the items that need revision in the appropriate space in step three of the Interpreting Item Analysis screen.

4. Identify areas for course review and student performance.

Now that the item analysis has been conducted and interpreted, enter notes that may be helpful in evaluating student performance or to use when the course is under review. The focus of this section is looking at how the students performed on the objectives as a whole. When reviewing this section look at the objectives that contain many items that were missed by a high percentage of students. Use the following questions to guide you as you are reviewing the items.

- How did the students perform on this item compared to the learning objective as a whole?
- Were most or all the students missing all or most of the items in the learning objective?
- Is the learning objective particularly difficult to learn?
- Were the students able to grasp the objective when it was presented in the course?
- Were students lacking in necessary prerequisites to accomplish this objective?
- Was there ample time in class to present the learning objective?
- Was there ample time for the students to practice?
- What was the time elapse from the time of teaching the learning objective and taking the test?
- Are the students able to make linkages between lessons or topics?