

How to interpret the course Level-2 evaluation results (For course manager/content experts)

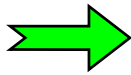
A. Overall, did participants learn?

1. The first page of your Course Evaluation Results (from the tabulation file) indicates whether the learning gain was statistically significant or not.
 - a. If the Learning Gain is statistically significant (appearing in green in the tabulation file), you can be over 95% confident that your participant learned some of the course contents.
 - b. If the Learning Gain is not statistically significant (appearing in orange in the tabulation file), the test could not detect learning with sufficient confidence. The reason(s) may be:
 - Participants overall did not learn enough for the gain to be significant; and/or
 - The evaluation could not detect learning, possibly because:
 - ❑ The test items did not represent well what was taught in the course;
 - ❑ The test items were not written sufficiently clearly for the audience;
 - ❑ The test forms were unclear (e.g., poor formatting/production); and/or
 - ❑ The setting prevented participants from answering thoughtfully (e.g., too short).

If all Toolkit guidelines were implemented, the evaluation should not be the issue.

A measure of test quality is the Internal Consistency Reliability Coefficient or Cronbach's Alpha shown in the Item Analysis Summary results for the "Additional Statistics" sheet. Roughly speaking, if it is below 0.50, the test may not have been reliable enough to detect learning.

| | A | B | C | D | E |
|----|---------------------------------------|---------------------------|-----------|-----------------|-----------|
| 1 | | Matched Respondents | | All Respondents | |
| 2 | Characteristics | Pre-test | Post-test | Pre-test | Post-test |
| 3 | Mean Test Score (% correct) | 41.8 | 50.6 | 41.1 | 50.3 |
| 4 | Median Test Score (% correct) | 40.0 | 50.0 | 40.0 | 48.0 |
| 5 | Standard Deviation | 15.5 | 18.9 | 15.5 | 18.1 |
| 6 | Standard Error of Mean | 2.4 | 2.9 | 2.3 | 2.6 |
| 7 | Skewness Ratio | 0.07 | 0.35 | 0.16 | 0.39 |
| 8 | Cronbach's Alpha | 0.67 | 0.78 | 0.67 | 0.75 |
| 9 | Measurement Error | 8.9 | 8.8 | 8.9 | 9.0 |
| 10 | Number of Respondents | 42 | 42 | 44 | 47 |
| 11 | Response Rate | 86% | 86% | 90% | 96% |
| 12 | Number of Items | 25 | 25 | 25 | 25 |
| 13 | Note: For definition of terms, click: | Definitions | | | |
| 16 | Total Number of Participants | 49 | | | |
| 17 | Learning Gain | 8.8 (in percentage point) | | | |



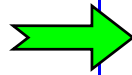
B. Overall, how much did participants learn?

2. Page 1 of your Course Evaluation Results (tabulation file) indicates the learning gain.
3. To compare your learning gain with the gain of other WB courses evaluated in the same way, click the [test items and results database](#) (only accessible from the World Bank's intranet). Then, click "compare a course results."

A. Why use it?

The Level-2 Evaluation Analysis Tool and Test Item Repository can be used to:

- a. [review items used in the past](#) (before writing test item). See [guidelines](#).
- b. [compare a course results](#) with other courses. See guidelines on [how to interpret results](#).
- c. [create ad hoc queries](#) using over 50 criteria to return a wide variety of Level-2 evaluations results presented by item, by offering, by course, by thematic programs, by division in different tables and charts and computing statistics.



- The average learning gain for over a hundred World Bank courses is about 13 percentage points. The standard deviation of the gain scores is about 9 percentage points. Based on these data, the table below gives you a rough guide to how your course Learning Gain compares to other courses.

| If your course Learning Gain is: | It is higher than: |
|----------------------------------|-----------------------------------|
| 28 percentage points | 95% of all other recorded courses |
| 23 percentage points | 87% of all other recorded courses |
| 18 percentage points | 71% of all other recorded courses |
| 13 percentage points | 50% of all other recorded courses |
| 8 percentage points | 29% of all other recorded courses |
| 3 percentage points | 13% of all other recorded courses |

- If you have access to the World Bank's intranet, and if you want to compare your courses with courses that share some of its characteristics (e.g., topics, duration, class size, delivery mode, language, etc.), query the database according to any combination of these criteria at: <http://www.wbitembank.com/NewL2/definequery.cfm>
- For example, if you want to return the results of DL courses only. Click the Triangle in front of "Course Offering Information."

Define Query
Current Query

Click the triangles to expand or collapse search categories and Define your search using the fields and search criteria within each category

- ▶ Course Offering Information
- ▶ Course Offering Evaluation Information
- ▶ Course Offering Results Data
- ▶ Course Offering Test Statistics
- ▶ Item Information
- ▶ Item Statistics

Choose Output:

Define Item Display LG
 LG R Pre R Post

- This expands the course offering information category:

Define Query
Current Query

Click the triangles to expand or collapse search categories and Define your search using the fields and search criteria within each category

▼ Course Offering Information

| Display | Search Criteria | Statistics | Sort |
|---|-----------------|------------|--|
| <input type="checkbox"/> Client or Staff Learning | CL | | <input checked="" type="radio"/> Ascending <input type="radio"/> Descending |

- Placing your cursor over a specific area provides details about the function it performs and the criteria selected.

9. Scrolling down brings you to the criterion “Distance Learning.”
10. Pull down the arrow, by distance learning and click “Yes”.
11. Select **at least one** field to display by clicking the box in front of a criterion, e.g., number of participants.
12. In the “Choose output” area, select the LG chart (it returns the distribution of learning gains for the queried courses).
13. Click “Submit Query”

The screenshot shows a query builder interface with the following elements:

- Step 1:** A green arrow points to the 'Distance learning' dropdown menu, which is currently set to 'No'.
- Step 2:** A green arrow points to the checkbox for 'Number of participants', which is checked.
- Step 3:** A green arrow points to the 'LG' (Learning Gain) chart icon in the 'Choose Output' section.
- Step 4:** A green arrow points to the 'Submit query' button.

Other visible elements include:

- Criteria: 'Delivered largely by partner' (No), 'Fee-based' (No).
- Sorting options: Ascending/Descending radio buttons for each criterion.
- Navigation: 'Reset all', 'Submit query', and 'Bookmark' buttons.
- Other output options: 'Define Item Display', 'R Pre', and 'R Post' charts.

14. The results will appear on your screen. The database has over 50 criteria for queries and 6 main output formats. Feel free to explore it to refine your comparisons.

C. What did participants learn?

15. The third and subsequent pages of your Course Evaluation Results (from the tabulation file) and the Item Analysis Summary (from the item analysis file) give the test results per items.
16. Review these results systematically, notably:
 - a. Was the matched post-test score higher than the matched pre-test score?¹
 - b. Did enough participants answer the item correctly at the end of the course?
 - c. Within each item, see where most responses are. Did participants move from a common misconception to the correct answer?
 - d. Were some items confusing (e.g., see discrimination index)?
17. The color-coded footnotes in the item analysis summary give you more details about how to interpret the statistical item analysis. However, statistics only provide one insight into an item's ability to detect learning. The review by a content expert from the content's perspective should be added for a complete analysis. Revise the course/test based on both reviews.

¹ A small decrease is no reason for alarm. From one time to another, participants typically vary slightly in their responses even without any intervention.