

How to create a test specifications matrix (For the course manager)

To build a test specification matrix, one needs global knowledge of a course learning objectives. The **course's Task Team Leader** is typically in the best position to perform this step.

1. Examine the course outline or materials to **identify large pieces of content**. These may be modules, key content areas, or any rational general organization of the course.
2. **Create an initial matrix** (see example below) with the identified pieces or modules in the first column and two columns for cognitive levels. Label one column "knowledge and concept." Label the second column as "application." The matrix will be used to map the testable content of the course in terms of content areas and cognitive domains.

Cognitive domains Content areas	Knowledge and concept	Application
Content area 1 (describe learning objectives)		
Content area 2		
....		
Content area N....		

3. **Decide the number of test items—between 20 and 50—that should compose the post-test.**¹ The upper limit for the number of questions usually depends on the time needed to answer the questions. In multiple-choice testing, test developers generally allow about a minute per question when the test is given in the native language of the participants. However, non-native speakers generally take longer to answer a question.
4. Following the course's learning objectives, **determine the relative importance of participants gaining "knowledge and concepts" versus being able to "apply" the course teaching.** "Knowledge and concept" refers to items that test the content directly taught. Examples include definition of essential terms, recall of facts, and straight knowledge of basic principles. "Application" refers to items that test the participants' processing of the knowledge taught in the course. Examples include problem solving, analysis, and evaluation. After the percentages between "knowledge and concept" and "application" are set, multiply the total number of items by each percentage to determine the number of items needed for each cognitive domain. (The example below illustrates the case of a test with 25 items, of which roughly 40% would go to "knowledge and concept" and 60% would go to "application.")

Cognitive domains Content areas	Knowledge and concept 40%	Application 60%
Content area 1 (describe learning objectives)		
Content area 2		
....		
Content area N....		
Total (25 questions)	10	15

¹ . From an evaluative point of view, using more items is better. However, because no World Bank courses evaluated using this method ever needed more than 50 items, and because developing larger tabulation tools would slow down results processing, the Toolkit's tabulation tools can **only compute results of tests with up to 50 items**.

5. In the column “knowledge and concept,” following the course’s learning objectives, identify which content areas are the most important with respect to this cognitive domain. Then determine if there are content areas where there is little or no knowledge transmitted to **all participants in the same way**, (typically, these content areas correspond to sessions like field trips, where everyone is expected to draw different lessons from the experience). **Establish the relative importance of the various content areas** (see example below).

Cognitive domains Content areas	Knowledge and concept 40%	Application 60%
Content area 1 (describe learning objectives)	Most important	
Content area 2	Not important	
....	Second most important	
Content area N....	Marginally important	
Total (25 questions)	10	15

6. **Turn these degrees of importance into their equivalent number of items.** The sum should correspond to the total number of items needed (see example below).

Cognitive domains Content areas	Knowledge and concept 40%	Application 60%
Content area 1 (describe learning objectives)	5	
Content area 2	0	
....	4	
Content area N....	1	
Total (25 questions)	10	15

7. **Go to the column labeled “application” and repeat steps 5 and 6 for this column.** (Some rows may have no questions planned, if they reflect content areas that by design aim to provide knowledge that would not be known in advance or if different participants are expected to acquire a large variety of different knowledge. They are often related to field trips, case studies, presentation of individual or small group projects, etc.)
8. **Check if a test developed according to the matrix would truly reflect achievement in the course.** If not, make the final adjustments. If the course content changes after this step, check the matrix and adjust it accordingly.
9. The test is now mapped. **Next, content experts should write the items according to the matrix**, following the guidelines on [how to write test items](#). If a previous offering of the course (or part of it) was tested, some items may be re-used. World Bank users may use the guidelines on [how to check items used in past tests](#) and assess each item’s quality with the [guidelines on how to review items](#) before writing new items.