Planning a Test

It is important to carefully prepare the instruments used to collect the information necessary for conducting a learning evaluation. Several activities must be performed to ensure the availability of appropriate instruments linked to the mechanism that we plan to use. When planning a test, we consider that some aspects have to ensure data collection tools for the evaluation mechanism to prepare. The EduStat software can be an useful tool when planning a test. Defining the parameters of the EduStat format database will link the items (that is to say the variables) to the specification table.

A. Steps in the development of a test

At the time of planning, designers must perform the following activities:

- *The definition of the domain or the preparation of the specification table* - Step during which we identify, describe and weights the specific objects of evaluation.
- *Validation of the domain of the test* – We have to check with a group of specialists if the selected area adequately represents what we want to evaluate; also we examine if the specification of the test is relevant.
- *The specification of the test* – When the type of test has been chosen, we define its parameters such as length, authorized material, marking mode, and type of items, exercises or tasks.
- *The specification of items, exercises or tasks* – This is the stage at which we set up the characteristics of items or exercises to produce (instructions, stimulus, and response).

B. Tasks to be done

After planning the production of a test, we have to develop the items that will be part of the test required for the completion of the proposed assessment; we have to perform the following tasks:

- *The drafting of items, exercises or tasks* – Items, exercises and tasks must be written according to the criteria adopted at the other steps.

- *The validation of content and docimological revision* – We should have here two types of specialists to be associated: first, the congruence between the items, exercises or tasks selected and the dimensions to be measured; secondly, respect for docimologycal requirements.
- *Writing the test* – During this step, we assemble and structure items, exercises or tasks; also produces the administration guide, student workbook, instructions to students and answer key.
- *Experimentation* – We have to verify the relevance of the evaluation material produced.
- *The revision of the instrumentation* – We have to modify assessment materials, if any, in the light of comments or suggestions obtained during the field test.

C. Documents to be prepared

At the same time that the officials prepare the test that will be used to gather information from students, we must write the various instructions related to each item or all of the test. We must produce the following documents:

- *The student booklet* – Students should know the questions or tasks that we intend to submit. Usually these questions are compiled in a booklet for students. Additional terms may however be provided; as examples, we can enumerate the oral examination and registration items on a board in the classroom. Whatever the means chosen, the student must have all the necessary information enabling him to demonstrate its performance.
- *The marking instructions* – Some questions require for students a short or long answer or an essay. This is mostly what is referred by the term "items corrected for the determination of partial credit." This kind of items is not corrected binary that is to say that the answer provided by the student is either correct or incorrect. The response should be assessed instead by the controller relative to a rubric that includes the award taking values corresponding to a numerical scale. We must then prepare a document outlining the marking instructions to ensure the consistency of scores for students. This applies even when the markers must codify some responses from students.
- *Administrative instructions* – We must ensure uniformity in the administration of a test. Several persons may be associated to the administration of a test. It is important to provide clear instructions for each person associated with the administration uses the same protocol.

Let us recall that the decisions taken in the context of learning assessment for certification purposes or selection of candidates have significant consequences. It is better to take the time to plan the operation to ensure the quality of decisions taken following the examination of results achieved by the students. As for the evaluation process for the purpose of piloting (formative evaluation and surveys on learning achievement mainly) requires that the mechanism allows the collection of information relevant to decision making.

D. Considerations

Several factors lead to pay particular attention to planning the development of instruments approach:

- the scope of the organization to develop;
- the time to invest to reach consensus, to develop the tests, to make the necessary improvements and to ensure the physical organization;
- the diversity of agents involved in the process.

E. Conditions to be respected

Several conditions must be met when planning the instrumentation necessary for the realization of an investigation.

1. Express clearly the purpose of the instrumentation

It may seem trivial or even pointless to report fault that clearly express the purpose of the test. However, it is important to ask the question: why do we administer the test? The answers to this question are various and they impact on writing, usage and marking of the test.

When one wants to achieve learning assessment to certify skills acquired by students after a year or at the end of a program of study, tests of quality requirements prove to be especially large if the measured customer is large. If instrumentation is taking place as part of the implementation of educational activities (formative evaluation), we need to quickly exploit the information gathered in order to adapt the educational activities at the current situation. Finally, if there is an investigation on learning outcomes, instrumentation should allow to have useful information for decision making.

2. Decide on the time of administration of the test

People who have responsibility for administering a test and additional questionnaires must determine when this instrument will be administered. If

possible, time must be chosen so that the conditions allow students to better demonstrate their skills.

When the test is prepared by an external organization, whether a ministry or a specialized firm, students and teachers must be informed early enough in order to be set.

3. Select the type of test

The test form is often reduced to the written questionnaire given to a group of students. However, we should also check whether another form would not be more appropriate. For example, the written questionnaire is not always appropriate when we want to measure the ability of a student to solve problems. In this case, a practical test is often the only truly appropriated. If, for economic reasons, officials decided to measure this type of performance from a written questionnaire, they must be considered when interpreting the results and be aware of the limitation of this measure.

There are three main families of tests:

- oral test;
- written test (pencil);
- practical test.

We will focus on the written test. This is the form usually used to gather information on students' performance.

4. Provide for the drafting and finalization items

The writing of items of a test requires time and specific skills. It is necessary to provide sufficient time to accomplish this step and determine who will participate, when and how. We will address these issues later.

5. Provide the physical conditions of administration

This is particularly important when the test is administered nationally. Managers must ensure that the conditions of administration are realistic according to the target clientele. They must also ensure that administrative requirements are the same for all students and provide mechanisms to identify all problems encountered during the execution of the test. We return to aspects related to quality control.

6. Provide the marking terms

In the case of tests with selected response, officials must plan the technical aspects related to the format of the documents provided to collect information from students and staff. The marking of tests with developed response is often time- consuming. Managers must predict the number of markers taking into account the average length of answers and include in the marking protocol, control steps to ensure the quality of the marking.

7. Provide the conditions for success

In order to judge whether a student has achieved the learning provided for in a program, we have to compare its results with the requirements that were established by consensus in defining what is essential according to the goals of education. The important thing is that realistic requirements with the program taking into account the context. This aspect is a key mechanism for the assessment of individual students or a survey of a population.

To assess the competence of a student, in the context of a test, must ordinarily determine a passing grade (e.g. 10/20) or a pass mark. The pass mark is used in a normative context and remains the same regardless of the used test. In a reference to test "criterion-referenced" to judge the competence of a student is determined a threshold of success. The pass rate is the level of quality at which a performance is considered be successful. It is established by consensus and differs from one test to another. The pass rate can be defined as a result that divides a distribution results in two distinct parts: the first containing the results from which we infer competence of individuals and the second, the results from which we infer the achievement of an insufficient degree of competence.

We must not consider the pass or the pass mark as a magic number. There is always a degree of uncertainty around the outcome chosen as the threshold of success or as a passing grade. The risk of error is present. A student can be judged competent when he is not, or one can judge a student so incompetent in reality this student mastery assessed learning. This is a statistical problem but also an educational problem. We must therefore establish mechanisms to judge borderline cases and exceptional situations. In doubtful or unforeseen cases, it convenes a committee whose members analyze data and compare their judgments justifying. In all these cases difficult to solve, concern for justice must guide people who judge.

When it is a survey of learning achievement, it may be useful to determine the needs of individuals through analysis of results (more on this subject when analyzing the data collected and preparation of evaluation reports).

8. Choose how to keep data

In assessment, planning instrumentation is also how to save data. We can proceed in different ways. The methods used reflect the technical resources that the school organizations have. We will deal with that at the time of publication of the results obtained following the administration of the evaluation mechanism.

9. Consider some technical aspects

Conducting a learning evaluation activity includes several technical aspects to control if we want to ensure rigor of the mechanism and availability of information in relation to the timetable set out. We present below some points to look at the development of tests, knowing that a much longer list should be drawn up if we wanted to be exhaustive.

9.1 Items arrangement

After developing a set of items or exercises, responsible for the design of a test must carry out the assembly. In completing this task, we must take into account the students' capacity. We should also consider the tasks related to data entry; available items must facilitate the work proposed to students and technical work related to the input information.

9.2 Student's responses

Designers must choose the most appropriate strategies for students to provide their responses. It should also state whether the students are entitled to certain instruments (reference books, scrap paper).

9.3 Marking and coding

Often it is necessary to mark the answers provided by students. As we noted above, then there should be the criteria used by the examiners to accomplish this task with rigor. Often we will have to code information in order to be able to treat it.

9.4 Data entry

It is important that, from the design of instrumentation phase, the input type of the information will be retained. It will be easier to make the test more suitable at this stage of an evaluation. If it turned out that some information cannot be entered with the available means, then it may be not useful to collect such information.

9.5 Statistical compilations

The information collected using the instrument provided will be compiled to provide a "result sheet" of the performance of assessed students. Previously, we

took treating responses. It is desirable to know, at the time of preparation of instrumentation, techniques to be used to perform certain tasks. Then we will have the information in relation to formats compatible with the technology used.

9.6 Links to be established between information elements

Leaders of a learning evaluation may wish to link multiple elements collected using the instrumentation (e.g., demographic information or expression of opinion). For this, we need that information can be merged to test and questionnaires.