

Field Testing

The field testing of an evaluation mechanism serves several purposes. The treatment of the information collected provides valuable information to finalize the documents necessary for the implementation of the assessment and administration procedures. During this activity, we have to perform the activities usually used for administering an assessment of learning. The sample will receive a limited number of students. The EduStat software could be retained in order to carry out the planned technical work.

This document presents some issues related to the process of field testing.

A. Metric qualities of tests

The first objective concerns the verification of the metric quality of developed instruments in the context of an assessment. Ensure compliance tests have the qualities usually accepted in relation to the difficulty, discrimination and lack of bias items and the consistency of the instruments. That is why there is typically a greater number of items to try to be able to identify those that will be used in the final administration. Sometimes it is necessary to provide several "forms" of tests to evaluate the same discipline or a complementary set of discipline; this requirement may be motivated either by the extent of the areas for evaluation, or to have a sufficient number of questions related to each area evaluated by the production of the desired statistical indices.

B. Verification of administrative procedures

A field testing can also be considered a "repetition", a "practical simulation" of the operation taking place in its "official" administration. It is desirable to carry out this activity in relation to the conditions expected to be preferred when the final administration. The comments received as a result of the field testing will better suit the different administrative guides that will be drafted.

C. Choice of compilations techniques

Data collected at the time of field testing are used to check the quality of the developing instruments, not to generalize the results to the entire population for the assessment activity. However statistical compilations will be interested in

certain subgroups of students to check the operation of the instruments compared to some of their characteristics. It will then make sure to have a sufficient number of individuals per subgroup to consider if we want to get reliable statistics.

Following the conduct of a field testing, it is also possible to set up and test the mechanisms that will be used for data processing and statistical compilations. Although the number of individuals assessed may be insufficient to be able to perform reliable analyzes, managers will have real and credible data to identify and test the most appropriate techniques in the context of the evaluation program.

D. Treatment of data in a field testing

At the time of planning a field testing of a future assessment mechanism, we must pay special attention to a set of interrelated aspects. Sampling of pupils' instrumentation in preparation is important. It should however be concerned, from the planning phase of activities, to be in place for the treatment of collected information. The very tight schedule requires that officials of the evaluation mechanism soon have the data necessary for the revision of the instrumentation and the finalization of administrative procedures. We would like to emphasize certain points concerning the action to grant the administration of the instruments tested on a representative sample.

E. Information collected from administrator

Administration of an assessment should be standardized if we do not want to introduce bias and errors due to data collection. The field testing helps gather important information on the difficulties encountered in the various schools chosen. The observation by administrators of students subjected to the tests often can detect inaccuracies in the instructions or inaccuracies in experienced test.

Consequently, it may be useful to gather feedback from directors associated to the field testing. A questionnaire should be provided to obtain such information from people closely involved in the administration of the experimentation of the evaluation mechanism.

F. Coding and scoring of assessment booklets.

A set of data will be collected via the booklets given to selected students. There will probably be "marked" certain responses (essay questions, for example); if

this is the situation, we usually retain quality control mechanisms affecting, in particular, the consistency of scoring. Other information must be codified before their introduction into databases. It is prudent to prepare this phase of processing information. Several methods can be retained; it may be useful to provide coding forms that will allow faster and more rigorous data capture after either codified or marked. If technically possible, we might consider the use of forms suitable for scanning.

G. Data entry

Data entry must be executed with rigor and speed. The choice of techniques to use should be done quickly, even before the administration of field testing. It should be noted that before performing data entry, it will be important to prepare the database that will host the information. This is a technical operation that requires time and many checks. Again, we could study the possibility of using the optical reader for computer character recognition.

H. Statistical compilations

From the beginning of data entry, it is desirable to immediately undertake the planned statistical compilations. Such an approach based on real data quickly identifies some unanticipated problems. It is also possible to detect anomalies in files created at the time of data entry; then it is easier to make corrections before beginning statistical compilations.

Long before the administration of the field testing, we will selected the type of statistical analysis that is assumed helpful to run. It should be noted here that the objectives pursued by field testing do not apply to the production of generalizable statistics to the entire population. This is to check on the metric qualities of the instrumentation. The results will enable designers of tests to perform a revision of the instruments that will be used at the final administration. The selected statistical compilations usually focuses on the items analysis and through examination of bias. It should be noted here that it is not possible to carry out analyzes of items according to item response theory (IRT), since the number of students exposed to the tests is limited; this analysis technique is the generalization of the results.

We will also provide quickly the communication of results to persons in charge of development of tests that will be used during the administration of the evaluation mechanism. Often it is necessary to provide information on the statistical ratios products to promote optimal use of available statistics.

I. Conclusion

A field testing of an evaluation mechanism is important in order to assure the rigor required by the various stakeholders interested in the information that will be collected. The revision of instruments and administrative procedures will be mainly affected by this activity. It is also an opportunity to experience the tasks to be accomplished by specifying the responsibilities of each partner and also the methods and techniques to be selected.