



UNESCO
INSTITUTE
FOR
STATISTICS



GLOBAL
ALLIANCE
TO MONITOR
LEARNING



GLOBAL
PARTNERSHIP
for EDUCATION

Global Alliance for Monitoring Learning
Fourth meeting
28-29 November 2017
Madrid, Spain
GAML4/19

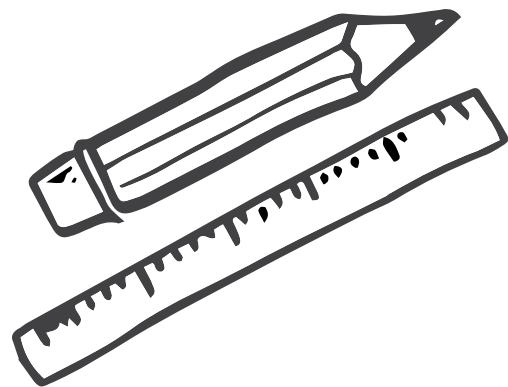
DRAFT

QUICK GUIDE
FOR IMPLEMENTING A
NATIONAL LEARNING ASSESSMENT



TABLE OF CONTENTS

1. Introduction.....	4
2. What decisions need to be made before launching the assessment?.....	6
3. How to measure learning?.....	9
4. How to select schools and students?.....	12
5. How to administer the assessment?	14
6. How to create a database?.....	16
7. How to compute results?	17
8. How to communicate results?	18
9. What personnel and facilities are required?	20
10. What are the main problems and challenges?	22
11. Where to find more information?	24



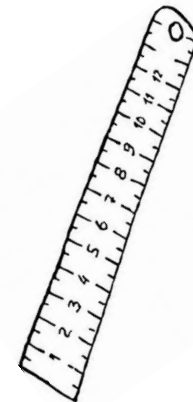
1. Introduction

Learning is key for the people to achieve their full potential and for the social and economic development of the countries. Therefore, it is important for countries to know how much students are learning.

To find out, strong learning assessments must be put in place. Countries implement national assessments for monitoring and supporting learning for all. This is done by administering assessments tests to measure what students know and can do with their knowledge. Assessments results are then used to inform educational policies and practices.

The purpose of this document is to support countries to implement their national learning assessments. It does so by answering to the following questions:

- What decisions need to be made before launching the assessment?
- How to measure learning?
- How to select schools and students?
- How to administer the assessment?
- How to create a database?
- How to compute results?
- How to communicate results?
- What personnel and facilities are required?
- What are the main problems and challenges?



Developing a national learning assessment is a long journey that requires political commitment, stable funding, and local technical capacity. Countries aiming to put in place their national learning assessments can greatly learn from the experience of other countries that are ahead on this journey. Mozambique offers a success story from a low income country that has been gradually implementing its national learning assessment (Box 1).

By putting in place strong learning assessments, countries are going to be better positioned to support learning for all.

Box 1. Mozambique's National Learning Assessment

Mozambique is a low income country in the south-east coast of Africa. After greatly expanding school coverage, the country's attention shifted to learning. This focus pushed for the development of a learning assessment with the purpose to monitor learning trends in key areas of the curriculum.

The first "National Assessment" was administered in 2013. A nationally representative sample of grade 3 students (6,764 students from 400 schools) was tested in reading in Portuguese (the language of instruction). The test measured students' ability to read letters, syllables, words, simple sentences, and short paragraphs. The test administration is standardized. External administrators read aloud the test instructions and most of the questions to whole classes of students. Teachers and school directors answered background questionnaires.

In 2016, the assessment was administered again, this time in both reading and mathematics. A new assessment is planned for 2019.

Assessment result shows the percentage of students by proficiency levels, with the highest level (Level III) matching the curricular expectations. They also show the 2013 and 2016 score differences. Results are published in a national report.

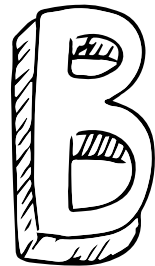
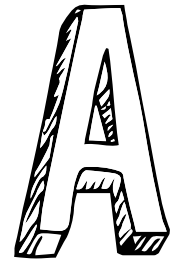
The National Assessment is in charge of [INDE](#), a semi-autonomous research institute within the Ministry of Education. INDE is in charge of test design and administration, data processing, analysis, and reporting.

INDE has a small team that has been gradually strengthening its technical capacity in assessment. The team had previous experience implementing the international assessment [SACMEQ](#) and a formative classroom assessment project ([Quizz I Can Read](#)) that provided the basis for developing the "National Assessment". The team also benefited from participating on a Master's degree program, doing study visits and hands-on training (with the support of the [READ Trust Fund](#)).

The costs of the "National Assessment" have been covered by both the government and a pooled fund of donors.

2. What decisions need to be made before launching an assessment?

- **WHAT is the purpose of the assessment?** Countries develop learning assessments to monitor to what extent students are reaching key learning objectives as outlined in national curricula, and to support learning for all. Some countries also use assessment results to hold schools accountable and to provide students and their parents with information about learning progress. It is essential to identify the policy questions the assessment should answer (e.g., What percentage of students are reaching the math curricular objectives at the end of primary education? Are results improving over time?). The assessment should be designed to answer these questions.
- **WHAT will be tested?** Assessments usually measure curricular areas (learning domains) considered critical for succeeding in life, such as reading comprehension, math, or sciences; others also measure innovative areas like writing, health, or finances in order to obtain a picture of student learning. It is also important to decide in which language(s) the assessment will be administered.
- **WHAT other information will be collected?** Countries need to decide if they are going to use questionnaires to collect information on background variables that are important to understand learning (e.g., teaching practices, availability of textbooks). For instance, many countries administer questionnaires to the students, their teachers, principals, and parents.
- **WHO will be tested?** Countries usually select one or more target grades (e.g., last grade of primary education) for the assessment. Depending on how a country wishes to use the assessment results, it may need to test a sample of



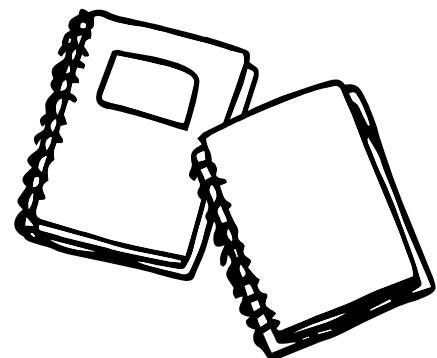
students that is representative at the country level; or it may need to test all schools and students (census). In both cases, countries need to have a comprehensive list of all schools so that they can either sample appropriately or assess all their students.

- **HOW will students be tested?** Currently the majority of learning assessments are administered in a paper and pencil format and use a combination of multiple choice and open-ended questions. Some countries administer similar types of tests online, as well as tests that are designed to assess skills like computer literacy. Most countries do group administration (e.g., to the whole class), while others do individual administration of the tests to the students. Test administrators must follow strict guidelines to ensure that all students are tested under uniform conditions (i.e., standardized administration).
- **WHO will administer the assessment?** External test administrators may be trained for administering the tests under the same conditions to all students. Teacher may also do so if they adhere to administration guidelines.
- **WHEN will the assessment be administered?** Some countries administer the assessment every year, others every two years or more. Some countries vary the grade tested (e.g., grade 3 is tested one year, grade 6 is tested the next year). The frequency of the assessment will vary depending on the purpose of the assessment, and on funding availability. Some countries administer the assessment at the beginning of the school year, others at the end. The timing of the assessment should take into consideration major school and national events. The same time will need to be adhered to for each assessment in order for results to be comparable over time.
- **HOW will results be published?** Results are ideally published as soon as possible. They should answer key policy questions (e.g., Do boys and girls have similar results in reading?). They can be published as percentage of students reaching different proficiency levels (e.g., Advanced, Basic, Below Basic), mean scores, percent of correct responses, or others. Frequent publication and discussion of results helps to educate different audiences and stakeholders

about how the country is addressing educational issues. This can promote confidence in the national assessment.

- **HOW MUCH will it cost?** Costs will vary greatly depending on the size of the assessment (i.e., number of students tested), administration mode (e.g., group administration vs individual administration), local costs of services (e.g., printing), and personnel (e.g., test administrators). It is important to estimate total costs and secure sufficient and stable funding (e.g., from government or donors).
- **WHO will be in charge of the assessment?** This could be the Ministry or Department of Education, a semi-autonomous government agency, or a university with a specialist assessment group. It is important to specify what personnel and facilities from the unit will be available for the assessment. The unit should be accountable to a clearly recognizable body (e.g., government department).
- **WHO will be in charge of making these key decisions?** It is recommended that a steering committee make key decisions about what should be assessed, by whom, of whom, and how frequently. The committee can be hosted by the national education system, and made up of representatives of key stakeholders (e.g., policy makers, educators, parents, assessment specialists). Involvement of stakeholders outside of the national education system can help secure national commitment, through their broader understanding of education and assessment issues. The Steering Committee should be advised by technical experts.

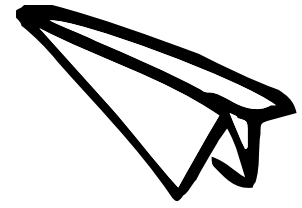
Each country will answer these questions according to its broader education goals, the local context, and the resources available. The decisions made as a result of these questions should be documented in an assessment framework that should be publicly available. An assessment framework specifies the key features of the national learning assessment, guides the assessment design and implementation, serves to inform stakeholders (e.g., policy makers,



educators, parents), and outlines the assessment approach to be taken by the country.

DRAFT

3. How to measure learning?



Countries usually implement learning assessments with the purpose to monitor to what extent students are reaching key learning objectives and to support learning for all. Key learning objectives are usually stated in the national curriculum.

Assessment tests must be aligned with the curriculum so that their results can be interpreted as students reaching (or not reaching) curricular objectives. In other words, the competencies, contents, and skills measured by the tests must be consistent with the learning expectations stated in the curriculum.

Assessment tests must also be attuned with the actual learning levels of the students. This is essential for reporting what students actually know and can do, and for knowing how far or close they are from reaching the curricular objectives.

It is essential that test development is a collaborative activity between curriculum experts, teachers, and assessment experts. Important steps to measure learning are:

- **Identify key strands, competencies, objectives, contents or skills.** For instance, reading literacy can be structured into the strands of decoding and reading comprehension. Math can be structured in content areas of numbers, geometry, measurement, and algebra. This step has implications for the format of assessments. For example, a paper and pencil test allows to measure reading comprehension, but not oral comprehension. As much as is possible, the format of the assessment needs to resemble the style of the actual knowledge and skills as performed by students in the classroom.
- **Identify key proficiency levels.** Each level (e.g., Advanced, Basic, Below Basic) describes what students know and can do at different stages of their learning trajectory. For instance, the highest proficiency level can describe the curricular objectives, whereas the lowest one can describe what lower performing students can actually do. Assessment results can be reported showing the percentage of students reaching each level.

- **Write test specifications for each learning domain.** Identify key categories of analysis to guide test design: strands, competencies, objectives, contents or skills; item formats (e.g., multiple choice or open response questions); type of texts (e.g., narrative, informative); complexity or difficulty of the items (e.g., easy, intermediate, hard). Specify the number of items or questions that will be needed overall and in each category. The distribution of items must be consistent with the curriculum.
- **Recruit and train item writers.** Item writers must have a good understanding of the assessment framework and test specification before starting to write items. Item writers can be from the national assessment or can be external to it (e.g., teachers, curriculum specialists from another unit of the Ministry or Department of Education).
- **Write test items.** Items (e.g., questions, tasks) must cover a wide range of difficulty levels in order to measure different learning levels. Good items (a) Present tasks that can be mapped back to the test specifications, (b) Are conceptually correct, (c) Focus on meaningful issues, not trivial details, and (c) Tell students what they are required to do. Good multiple-choice items also (a) Include a key that is indisputably correct, (b) Include distractors that are indisputably incorrect, while being reasonable and plausible, (c) Present response options that are equivalent in length and style. Open response items must be written together with their scoring guides, and they must be scored after the test administration.
- **Write test instructions.** Instructions should welcome the students, explain how to complete the test and in how much time, and provide some practice items. Students must understand how to answer the test, including multiple-choice questions. If a separate answer sheet is provided, students should practice marking their responses on the sheet. Instructions must be read by the test administrator before the students start answering the test.
- **Revised and pre-pilot the items.** Items should be reviewed by item writers, and by external panels of persons not involved in item writing. Revised items should be pre-piloted by administering them to a small number of students (e.g.,

50 students). Subsequent rounds of reviews and revisions should be carried out before assembling pilot test booklets.

- **Do a pilot administration.** This is important for knowing the psychometric characteristics (e.g., difficulty, discrimination) of the test overall and of each item. It is good practice for the pilot to include two or three times the number of items required for the final test, and to administer each item to a minimum of 200 students with similar characteristics to those that will be taking the final test. The pilot is also important for testing the administration manual, logistics (e.g., transportation), and other procedures.
- **Analyze pilot results.** Compute item statistics such as difficulty (e.g., percentage of correct responses, scaled scores) and discrimination (e.g., point-biserial correlation), percentage of students by response option (for multiple choice items) or percentage obtaining full or partial credit (for open-ended questions), and percentage omitted. Key test statistics include mean scores, reliability or precision (e.g., Cronbach's Alpha). It is good practice that items cover a wide range of difficulty and that Cronbach's Alpha is greater than 0.70.
- **Assemble test booklets for main assessment administration.** Test booklets should be engaging without being distractive. They should include unique identifier numbers (ID) to track students and schools. The number of items in the booklets should be adequate so that students can answer the test within a reasonable amount of time. It is good to put easier items at the beginning of the booklets. Countries often use different booklets with different items, in order to increase the overall number of items tested. In this case, different booklets should be equivalent in terms of content and psychometric characteristics (e.g., difficulty). They may also share some common items.
- **Prepare for printing.** Printing is expensive and may take several weeks. Quotations should be requested as soon as key parameters are known (e.g., number of booklets to be printed, number of pages by booklet). All test booklets and other instruments (e.g., questionnaires, administration manuals) should be carefully checked before printing.

4. How to select schools and students?



Schools and students should be selected so that overall assessment results can be interpreted as pertaining to the country as a whole. In some cases, the selection of schools and students may need to be restricted due to financial or logistical constraints. In these cases, assessment results must be interpreted as pertaining to a sub-national level within the country (e.g., percentage of students reaching minimum proficiency level in urban schools within region X). All the decisions regarding the selection of schools and students should be documented.

Key steps for selecting schools and students for a learning assessment are:

- **Specify the target population of students.** Indicate to which students and schools within the country the results would apply (e.g., All grade 6 students in both public and private schools).
- **Construct or obtain a comprehensive list of schools.** This list (sampling frame) should ideally include identification and contact information for all target schools, school classification information (e.g., urban/rural, public/private, region/district), number of classes and number of students in the target grade.
- **Decide on the best approach for selecting schools and students:**
 - a. Select a sample of students that represents the country (or sub-national) level. This is the best approach if the purpose of the assessment is to estimate overall learning levels for the country. Sampling methods usually involve (i) sampling stratification according to school classification information (e.g., by region), (ii) systematic random sampling of schools from the sampling frame, with bigger schools having a greater probability of being selected, and (iii) sampling of an intact class or students within schools. It is best

practice to select replacement schools too, in case some of the selected schools cannot participate.

- b. **Select all schools and students** (i.e., census) in the target grade. This is usually the best approach if the assessment aims to report school-level results, or if the assessment is implemented in a relatively small country.

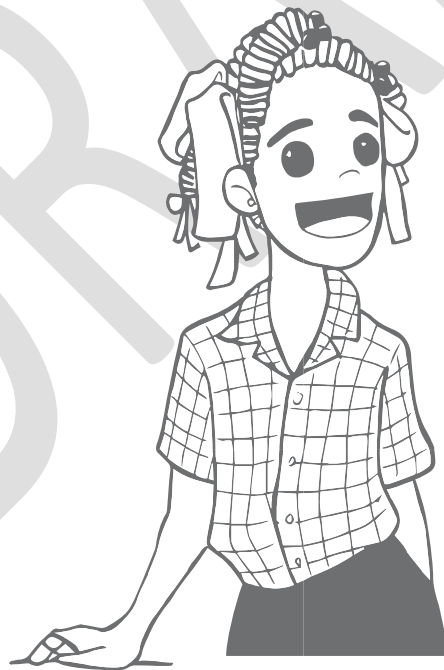


5. How to administer the assessment?

The learning assessment administration is a major undertaking. It usually requires contacting hundreds of schools, recruiting and training hundreds of test administrators and supervisors, distributing thousands of test booklets in a secure way all over the country, and sub-contracting services (e.g., transportation). Field operation refers to all the activities needed for a successful test administration. These include:

- **Design a field operation plan.** This should identify activities, responsibilities, and timelines. It state how schools will be allocated to field coordinators and test administrators (e.g., 5 schools per coordinator). A field operation manual and documentation (e.g., forms, control checklists) should provide guidelines and should serve for quality control purposes.
- **Train key personnel.** Training should emphasize assessment purposes, the importance of uniform testing conditions, and the roles and responsibilities of the trainees during the field operation. It should be conducted before the data collection, and should replicate as much as possible the real test administration conditions. For instance, by reading aloud the test administration manual during the training session and by filling administration forms.
- **Contact schools.** Invite selected schools to participate in the assessments. Ensure there is a contact person in the school who can help organize the test administration (e.g., informing the school teachers about the assessment, storing test instruments, etc.).
- **Print, check and distribute assessment instruments.** Check samples of the printed instruments (e.g., tests, questionnaires) to ensure they meet quality standards. Packaging and distribution may take several weeks. The security and confidentiality of the instruments should be a top priority.

- **Administer the assessment according to standardized procedures.** All students should be tested under the same uniform conditions. The test administrator manual should state those conditions and serve as a basis for quality control.
- **Send quality control observers to schools.** This is important to verify that the assessment was effectively administered under the same uniform conditions in all schools. It is also important to build trust around the assessment. Observers may use the administration manual or a checklist as a reference. Observers may include personnel from the Ministry or Department of Education, donors, parents, or others.
- **Verify the return of assessment instruments.** Usually countries need to collect all test booklets and other instruments (e.g., questionnaires, forms) from schools. All booklets and instruments should be registered (e.g., using ID numbers) and accounted for before and after the test.



6. How to create a database?

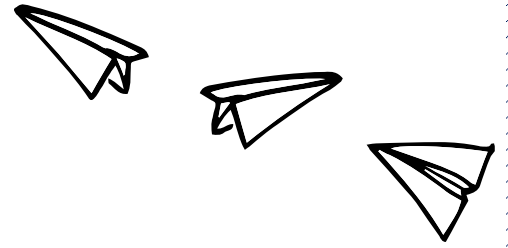
- **Prepare database structure for capturing data.** The database should include all the variables in the test booklets and other assessment instruments (e.g., IDs; responses to item 1, item 2, etc.). It must specify the type of information to be recorded in each variable (e.g., text/ numeric, valid ranges). This information should be documented in a database manual (i.e., codebook).
- **Prepare protocols for data processing.** Protocols establish rules for data capture, data cleaning, and verification. For instance, they provide instructions for data key punchers on how to enter the data, state rules for doing internal consistency checks (e.g., if a student is absent then all item responses should be missing) and automated corrections.
- **Capture raw data in database.** Methods include online data collection, scanning of test booklets or mark sheets, and manual key punching. Special hardware (e.g., high speed scanners) and software (e.g., Access, Excel) may be needed depending on the method. Provisions need to be made to secure the required hardware and software for this task.
- **Data cleaning.** This include procedures for checking and correcting unique identifiers (IDs), duplicated or dropped records, checks that the values entered are valid and within range. Data cleaning usually requires several rounds of iterations before producing the final or official database.
- **Compute or add new variables.** These may include:
 - a. IRT (Item Response Theory) scaled scores. These are useful for putting in the same scale and comparing scores from different booklets.
 - b. Proficiency levels. Students are classified in a level (e.g., Advanced, Basic, Below Basic) based on their final score in the test.
 - c. Weights to adjust for sampling design. These are important to produce valid results at the country level.

7. How to compute results?

- **Identify key policy questions that will guide the analyses.** For instance: What is the country mean score in mathematics? What percentage of students reached the minimum proficiency level in reading? Is there any difference in scores between boys and girls? What percentage of students answered correctly specific items in the test? Did results improved compared to the previous assessment?
- **Analyze data.** Compute descriptive statistics (e.g., Mean scores, percentage of students) and basic inferential statistics (e.g., mean comparisons) to answer the most basic research questions. Acknowledge that scores are not perfectly precise (reliable) by reporting standard errors or by indicating if differences are statistically significant.
- **Apply weights** according to the sampling design.
- **Document all your analyses** so that others can understand and replicate them.



8. How to communicate results?



Countries put in place learning assessments to monitor and support learning for all. To meet this purpose, assessment results must be communicated effectively to stakeholders. Key steps to ensure an effective communication of assessment results are:

- **Prepare a communication plan.** Identify the key policy questions and messages to communicate, the communication products (e.g., national report, flyer), the stakeholders to whom each product will be directed (e.g., teachers, parents), and the timing when these products will be published (e.g., 3 months after the assessment administration). Make sure to secure enough funding to implement this plan.
- **Identify key policy questions and messages.** For example, how many students are meeting the curricular objectives? How do the regional results compare to the national results? Are results improving over time?
- **Select the best communication products.** Different products serve different objectives and are useful to reach different audiences or stakeholders:
 - a. National reports for the general public. Provide an overview of the assessment and its main results.
 - b. Workshops for teachers and school leaders. Provide an opportunity to explain the importance of the assessment, and to discuss how to use the results to improve student learning.
 - c. Media reports. Contain all the information needed to communicate the results to a wider audience, including key messages, graphs and tables.
 - d. Websites, videos, blogs and social media for engaging the general public on a regular basis.

- e. School reports and parent reports are effective tools for accountability purposes.
 - f. Conferences and presentations (e.g., in Power Point) offer an opportunity to engage the academia and other institutions (e.g., funding agencies)
 - g. Flyers, brochures, posters may be used to communicate brief messages to a wider audience and to those that do not have time to read more elaborated results (e.g., policy makers).
- **Identify key audiences or stakeholders:** Parents, teachers, the Ministry or Department of Education, funding agencies, media, and the general public. Make sure to address their information needs in the communication products. Ensure that they have access to the communication products (e.g., by distributing a school report to all schools). Build an assessment culture where stakeholders understand, value and use assessment information to improve learning.
 - **Prepare communication products.** Make sure that the products respond to key policy questions (e.g., What percentage of students are reaching the minimum proficiency level in reading?), that the information reported is correct, clear and simple; and that the reports look nice and professional. Printing some reports may require additional databases (e.g., school report database allowing for printing specific scores for each school).
 - **Present actions to improve assessment results.** This is especially important when reporting poor results. For instance, the Minister of Education may say that these results are a diagnosis or starting point, and launch a new program to support learning. Schools may set learning objectives to be reached the next time assessment results are published.
 - **Monitor how assessment results are being used.** Check if reports are reaching the schools, parents, or other stakeholders as intended. Meet with stakeholders and ask them about the different assessment products: Did they get them? Did they find the results informative for their particular schools? Could they understand them? How did they use the results to support learning?

DRAFT

9. What personnel and facilities are required?

Key staff for implementing a national learning assessment include:

- **National coordinator.** Should give general directions and leadership, ensuring that the assessment address key policy questions, is technically adequate, and is carried out on time and within budget. Sub-national coordinators may also be needed depending on the country.
- **Test development coordinators.** These persons are in charge of producing the tests to measure learning in each learning domain (e.g., math, reading). They should be knowledgeable about the curricular objectives and the actual learning levels of the students in the classrooms. Main responsibilities include recruiting, training, and supporting teams of item writers, and scoring open-ended questions.
- **Field operation coordinator.** Is responsible for the standardized administration of all assessment instruments to the students. Is in charge of recruiting and training test administrators, supervisors, regional or district level coordinators, and quality control observers. This personnel may be full time staff from the national assessment, staff from other branches of the Ministry or Department of Education, or may be employed on a temporary basis. The field operation coordinator is also responsible for producing the administration manual and of all the logistics surrounding the assessment (e.g., distribution and recollection of instruments).
- **Data manager coordinator.** Is responsible for preparing the databases, i.e., capturing and cleaning the data, computing weights, scores, cut-off points associated to the proficiency levels, and results. This person has to make sure that the databases meet the data requirement of the communication products. Should be competent in running statistical and psychometric analyses using different software (e.g., SAS, SPSS, STATA, IATA).

- **Communication coordinator.** Is responsible for putting in place and implementing an effective communication plan. Working with a design and communication person is highly recommended.

The national assessment need to have access to basic facilities to operate, such as:

- **Meeting rooms and offices** for key staff with computers connected to the internet, printers, and other normal offices supplies.
- **Hardware** (e.g., scanners) **and software** (e.g., SPSS, IATA) with licenses to operate.
- **Vehicles for transportation** of boxes with assessment instruments.
- **Secure storage facilities** to accommodate large amount of boxes with assessment instruments.



10. What are the main problems and challenges?

- **Failure to secure political support and stable funding.** As a consequence, the assessment stability is at risk. The assessment may be administered only once, or on an irregular basis. Involving stakeholders and transparency are essential to minimize this risk.
- **Need to secure sufficient staff.** It is common to see assessment units whose members can only devote a fraction of their time to the assessment. It is important to agree on the number of staff and the amount of time they would put in the assessment, and to plan the assessment accordingly.
- **Need to develop local capacity.** The best way to do so is by providing hands-on training while implementing an assessment. Other capacity building activities may involve courses or workshops (face-to-face and distance), study visits, and attending conferences. Efforts should be made to retain the trained staff.
- **Tests are too difficult.** Assessment tests are usually good at measuring the curricular objectives or expectations, but are not that good at measuring what students at the low end of the learning distribution can do. Consequently, results may end up showing that most of the students are clustered in the lowest proficiency levels. Therefore, make sure that test design is informed by both the curriculum and the actual learning levels of the students.
- **Poor sampling.** Assessments that are supposed to be representative at the country level may end up using samples that are representative of a few regions, or a certain type of schools (e.g., urban schools) only. Or, during the field work, only the schools that are “easy to reach” may end up being tested. This may happen because of poor planning, logistical difficulties, or being out of budget. Sampling and field work should be planned in detail and well in advanced to avoid these problems.

- **Lack of standardized procedures.** When there are not clear guidelines and manuals to administer the assessment, when tests administrators do not have appropriate training, when scoring rules are not clear, when there not clear rules or protocols for creating the dataset... the whole assessment effort may end up being wasted. A minimum level of standardization is needed to ensure that test scores can be interpreted as intended. Manuals, training, and quality control procedures are important tools to ensure standardization.
- **Assessment results are not comparable.** This is a major issue when the aim is to report changes in learning across years. Comparability is at risk when (a) the sampled schools and students are not equivalent from one test administration to the other, (b) the tests are not parallels or have not been put on the same score scale. Another related problem is that the uncertainty (error) associated to the tests scores is unknown. Then it is not possible to know if differences in scores are meaningful or statistically significant. When designing the assessment, make sure it has the technical features needed for comparing results.
- **Assessment results are not published.** This is most likely to happen in the context of poor results. It is important to anticipate this scenario with the Minister of Education or equivalent. For instance, planning for a communication strategy where poor results are used as a baseline to start improving.
- **Lack of an assessment culture.** Stakeholders (e.g., teachers, parents, politicians) may not understand, value and use assessment results as intended. This may happen because results reports are too long and complicated. Produce simple reports, flyers, websites that address a few key research; and offer workshops to explain results and the assessment in general.
- **Lack of documentation.** Technical documentation (e.g., assessment framework, test specifications, sampling plans, data processing protocols) is needed to provide evidence that the assessment can be interpreted and used as intended. This is key to ensure the credibility of the assessment.

11. Where to find more information?

This Quick Guide is based on the following references, which cover the topics presented here in much more detail:

UNESCO-UIS and ACER. 2017. Principles of Good Practice in Learning Assessment.

[Link](#)

World Bank. 2008. National Assessments of Educational Achievement Series (available in English, French, Portuguese, Spanish, and Russian): [Link](#)

- Assessing national achievement levels in education (Vol. 1)
- Developing tests and questionnaires for a national assessment of educational achievement (Vol. 2)
- Implementing a national assessment of educational achievement (Vol. 3)
- Analyzing data from a national assessment of educational achievement (Vol. 4).
- Using the results of a national assessment of educational achievement (Vol. 5)

