PDF ISBN 978-92-76-25180-4 ISSN 1977-6411 doi: 10.2841/984187 MN-BB-20-012-EN-N JANUARY 2020

PRACTICAL GUIDANCE NOTE 11

LEARNING ASSESSMENT

1. Topic Overview

In many countries, increasing attention is being paid to worryingly low levels of learning. The substantial gains in access to basic education have not been matched by gains in learning outcomes: too high a proportion of children leave primary school without even the minimum expected levels of proficiency in literacy or numeracy. With this recognition of a "learning crisis" has come greater acceptance and focus on learning assessment. "Countries that have been successful in improving learning usually have in place strong learning assessments".[i]

Sustainable Development Goal 4 (SDG4) emphasises the centrality of learning. Robust assessment is key to monitoring progress in SDG4. UNESCO's Global Alliance to Monitor Learning (GAML) is designed to improve learning outcomes by supporting national strategies for learning assessments and developing internationally-comparable indicators and methodological tools to measure progress towards key targets of Sustainable Development Goal 4.

Whilst many developing countries are now engaging in some form of large-scale learning assessment, as distinct from examinations which all countries do, the approach tends to be ad hoc and the connection between what the assessments find and the policy response remains weak.[ii]

Sustainable Development Goal 4 (SDG4) emphasises the centrality of *learning*. Robust assessment is key to monitoring progress in SDG4. UNESCO's Global Alliance to Monitor Learning (GAML) is designed to improve learning outcomes by supporting national strategies for learning assessments and developing internationally-comparable indicators and methodological tools to measure progress towards key targets of Sustainable Development Goal 4.

Summary

Assessment is concerned with the reliable measurement of learning and is a critical element in improving educational outcomes

There are 3 basic types of assessment:

1. Large-scale national and international surveys,
2. Examinations
3. Day-to-day, school-based assessment of pupils' learning

Major international interest is around large-scale surveys of learning. There are numerous types, ranging from wholly national assessments, to NGO-driven ones (e.g. ASER, UWEZO), and international assessments such as PISA and PIRLS (*details below*).

Effective assessment requires a sensitivity to context, strong government ownership and technical quality.

Each type of assessment has a range of pros and cons: choice should be determined by purpose and context.

What do we mean by assessment? Very simply put, assessment is finding out if what is taught (or is expected to be taught) is learned. There are many ways in which this is done in an effective education system: different methods serve different purposes and can tell us different things about learners and their learning.

> Learning assessment is part of evaluating and monitoring education more broadly. It refers to a wide range of methods and tools used to evaluate, measure and document learning outcomes and learning progress. It is about gathering information from multiple sources on what learners know and what they can do with what they have learnt. It also provides information on the process and context that enable learning, as well as those that may be hindering the learning process.
>
> **UNESCO, 2016**

i UNESCO Institute for Statistics, *Quick Guide No. 2: Making the Case for a Learning Assessment*, 2018, p. 8 http://uis.unesco.org/sites/default/files/documents/quick-guide2-making-case-learning-assessments-2018-en_2.pdf

ii UNESCO found that in Sub-Saharan Africa, only South Africa showed any policy response to the results. See UNESCO Institute for Statistics, *The Impact of Large-Scale Learning Assessments*, 2018 http://uis.unesco.org/sites/default/files/documents/impact-large-scale-assessments-2018-en.pdf

Why is assessment important? Assessing levels of learning is important for a number of reasons.

- It is a critical part of the teaching and learning process and, in measuring learning outcomes, is a high-level indicator of an education system's effectiveness.
- It provides critical outcome information to policy-makers, education managers, including curriculum developers, teacher trainers, textbook producers, school inspectors, principals and teachers.
- It raises the national debate around performance and accountability and, through open disclosure, puts pressure on providers by raising public expectations.
- It helps makes evidence-based investment choices around what to prioritise in order to improve learning.

Summative assessment measures learning at the end of defined period or course of instruction against established standards or benchmarks.

Formative assessment regularly monitors learning in order to give feedback to teachers and learners to improve learning.

Reliability is the capacity of a test to give the same results for a test-taker, irrespective of when or where the test is taken.

Validity is a test's capacity to give a correct measurement of learning.[iii]

Psychometrics is concerned with the theory and technique of psychological measurement. It refers to the field in psychology and education that is devoted to testing, measurement and assessment.

Item Response Theory [IRT] is a theory of testing based on the relationship between individuals' performances on a test item and the test takers' levels of performance on an overall measure of the ability that item was designed to measure. IRT allows for banks of different test items to be created that all reliably measure the same thing which can be helpful in minimising cheating.

National Assessments

These are variously referred to as National Assessment Surveys, National Achievement Surveys, or National Learning Assessments. These are used to provide national statements of learning at prescribed stages or age levels in specific areas, most commonly in literacy and mathematics. It may also include science and social studies. These large-scale, system-level sample surveys are usually done in one of two ways: most commonly sampling learners within schools (so they do not capture out-of-school children) or, less commonly, household sampling of particular school-age groups (so they do capture children not in school). There are many types of such assessments, each developed and used for different purpose, ranging from those run by national governments, for example in India[iv], to those run by NGOs, e.g. ASER[v], UWEZO[vi], as well as the now quite extensively used Early Grade Reading Assessments (EGRA)[vii] and Early Grade Mathematics Assessments (EGMA)[viii]. These are largely promoted by USAID and often used to measure the impact of specific early literacy and numeracy programmes. Assessments undertaken by NGOs (e.g. ASER in India and Pakistan, UWEZO in East Africa) are attractive because they are independent, get results relatively quickly and do not rely on government financing. However, they often have low levels of acceptance by governments who may contest findings. Independent assessments can have a shock effect and raise public awareness and debate, and can be an important part of pushing for greater accountability.

National Assessments can also capture other social and economic data of learners tested, as well as their school environment. This enables deeper analysis of factors that influence learning and provides important information for determining appropriate policy responses. Results can be reliably compared over time and location. However, because these assessments are sample based, they are unable to provide individual child or school-level information on learning. Results of national assessments are usually communicated through published reports. National assessments are of particular use to policy-makers, teacher-trainers, and curriculum developers, as well as development partners, as they provide system-level information that can shape both policy and programmes.

There is increasing interest and use of technology in assessment, largely spearheaded within OECD countries, e.g. Scotland[ix] but with some application in developing countries, e.g. Lesotho and Afghanistan[x]. Whilst a limited use of technology involving optical mark readers has been around for some time, the wider application of digital technology, through the use of online, tablet-based assessment tools, for example, is relatively new. These applications have a number of advantages in terms of accuracy, analysis and reporting, as well as reducing the load of marking. However, there are also significant challenges, primarily cost and accessibility, but also around institutional capacity and acceptance. The use of online surveys are likely to underrepresent data from schools that have limited access to electricity and technology such as those in conflict and disaster-affected areas, as well as in less accessible rural areas.

iii A simple way to understand the difference between reliability and validity is to think of weighing scales. They are reliable if they give the same result each time the same thing is weighed. They are valid if the weight they show is right.

iv National Achievement Survey(NAS) by the Ministry of Human Resource Development of India http://mhrd.gov.in/NAS/

v The Annual Status of Education Report in India http://www.asercentre.org/#8im15, The Annual Status of Education Report in Pakistan http://aserpakistan.org/index.php

vi Uwezo Learning Assessments https://www.twaweza.org/index.php?c=83

vii Early Grade Reading Assessment (EGRA) toolkit developed by RTI International https://shared.rti.org/content/early-grade-reading-assessment-egra-toolkit-second-edition

viii Early Grade Mathematics Assessment (EGMA) Toolkit developed by RTI International https://shared.rti.org/content/early-grade-mathematics-assessment-egma-toolkit

ix Scottish National Standardised Assessments (SNSA) https://standardisedassessment.gov.scot

x Adams, R., "Data to inform policy", International Developments, No. 5, Australian Council for Educational Research, February 2015, https://research.acer.edu.au/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1027&context=intdev

International large-scale assessments

There are a number of international assessments of learning, many of which are primarily designed for more developed countries, for example the OECD's Programme for International Students Assessment (PISA), Trends in International Mathematics and Science Study (TIMSS) and Progress in International Reading Literacy Study (PIRLS), though PISA has developed its PISA for Development (PISA-D) in 2015 which has to date conducted surveys in eight lower middle income countries.

In addition to these global assessments, there are a number of regional organisations that carry out periodic learning assessments. The Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) has, since 1995 carried out four rounds of learning assessment in 16 countries in the region. Similarly, the Programme d'Analyse des Systemes Educatif de la CONFEMEN[xi] (PASEC) has, since 1993, conducted assessments of learners in French and Mathematics in around 30 Francophone countries worldwide. In South America, the Latin American Laboratory for Assessment of the Quality of Education (LLECE) has undertaken three rounds of literacy and numeracy assessments in Grade 3 and 6 since 1997, with a further round planned in 2019.[xii] International assessments enable cross-country comparisons of performance over time. As with National Assessments, they are of particular use to policy-makers and education practitioners.

National and international examinations

The overall purpose of examinations is fundamentally different to that of national assessments. Examinations are taken by all learners across a range of subjects at certain points in their education Most typically in developing countries, the major examination points are at the end of primary school (often the Primary Leaving Examination), at the end of lower (or junior) secondary and then at the end of full secondary (often referred to as Secondary Certificate or Matriculation).

In some countries, advanced-level examinations are taken at the end of upper secondary school. All these are often referred to as "high stakes" or "gateway" examinations as performance often determines entry to the next level of education, or are necessary qualifications for employment. Examination results attract high levels of parental and political interest, and are often the main way school performance is judged. Some countries use international examinations set by international boards such as the West Africa Senior Secondary Certificate Examinations (WASSCE) or the Caribbean Certificate of Secondary Level Competence (CCSLC). These enable countries with limited expertise to benefit from scaled operations, offer some degree of independent legitimacy and enable cross-country comparison of performance. There is less coherence at primary level. In many cases the primary leaving examination is designed and administered at provincial or district-level examination, so may lack national coherence. The Caribbean Examination Council has attempted a unified regional Primary Exit Examination (CPEA), but take up has been limited.

School and classroom-based, on-going assessment of pupils' learning

This is usually a less-systematised approach, that varies across schools and is largely determined by the commitment and skills of individual teachers. It is often referred to as continuous assessment. This includes the week-by-week, day-to-day, lesson-by-lesson checking by teachers that their pupils are learning. It involves a range of techniques embedded within their pedagogy, such as questioning routines, setting and marking of exercises, including homework, and periodic testing. Information from such assessment generally remains at the class or school level, being shared with pupils and parents through pupil's school reports, feedback at teacher-parent events and displayed test results.

The extent to which this is effectively carried out in schools varies widely, and whilst in more developed education systems continuous assessment often forms a significant part of overall learner assessment, its use in developing countries is limited. The validity of what such assessment may say about learner achievement is not standardised and often unreliable. Development in this important area is frequently constrained by teacher capability and motivation, as well as the feasibility of being able to reliably check the progress of all learners in heavily over-crowded classrooms. Yet, this type of assessment, potentially of every child's progress, is the most useful at the school level and one that can provide immediate opportunity to influence teaching and learning.

In its most organised form, groups of schools (i.e. those in a district, or belonging to a particular board) may conduct termly or yearly tests or examinations, but the setting, administration, marking and reporting of these is often highly questionable and results cannot form part of a serious discourse on learner achievement. Furthermore, these tests or examinations may be used at district or board level to decide student progression into the next grade or level (high stakes testing) but this is inadvisable. In crisis-affected contexts, these tests often gain prominence in areas with limited government services, and are then used to decide student progression or entry to school for displaced children.

xi La Conference des Ministres de l'Education des Etats et gouvernements de la Francophonie

xii UNESCO Regional Bureau for Education in Latin America and the Caribbean, "La UNESCO prepara aplicación de evaluación regional de aprendizajes con encuentro de coordinadores nacionales y seminario sobre currículos", 24 March 2019 http://www.unesco.org/new/en/santiago/education/education-assessment-llece/events/41-meeting-coordinators-llece/reunion-coordinadores-havana-2019/

2. Key Issues

Investing in assessment in itself will not improve learning. It is the acceptance of the results, and using these to address weaknesses in the system, that brings about change. It is important to see assessment not simply about measurement but as an integral part of improvement.

Policy dialogue and decisions around assessment need to be framed within the broader system. This section focusses largely on the main policy dialogue concerns that are likely to arise around national and international assessments. However, these need to be seen as part of wider policy and practice on all types of assessment. Firstly, then, it is helpful to summarise some of the key issues associated with examinations and classroom assessment.

National Examination Systems: key issues

- At what stages of the education cycle do students sit national examinations?
- Who is responsible for the setting, management, marking and reporting on national examinations?
- Is this an entirely centralised function for all levels, or is there some state, provincial, regional or district autonomy?
- What is the general view of the relevance, reliability and validity of these examinations?
- To what extent are these examinations used to control the flow of students to the next stage? It is important to understand the effect that examinations have on students and the system as a whole.

> An interesting and infamous example comes from Liberia where, in 2013, following the WASCCE examination results where the government had lowered the pass mark, the University of Liberia carried out its own entrance examinations and failed all 25,000 applicants, finding them semi-literate![xiii]

- Potential negative effects may include:
  - High repetition rates as weaker students are held back to improve chances of success
  - Drop out, as students with little chance of success leave school
  - Cash for extra-tuition. Teachers see exams as an opportunity to earn money through additional classes and may even scale back on normal classes to increase the demand
  - Rent-seeking by officials who control school placements
  - Cheating by teachers and schools to inflate their pass rates, and by examiners offering grades for cash
- Engaging in examination reform is difficult. Governments are often reluctant to accept external interference and there is much vested interest in the status quo.
- Linking national examination boards with those from more developed systems can be a useful approach in strengthening the validity and reliability of national examination systems. This also builds wider, international legitimacy, as awards in one country are accepted in others.
- In theory, encouraging membership of regional examination systems (e.g. the West African Examination Council) can help to strengthen a national system, but there is no guarantee of the fidelity of a national system which remains susceptible to malpractice (e.g. Nigeria).[xiv]

Classroom assessment: key issues

- Regular and reliable assessment of every pupil's learning is a critical first step to improving learning outcomes. This is both assessment of learning (what a child has actually learned against what is expected) and assessment for learning [enabling a teacher to understand progress and adapt their teaching accordingly). It provides the bridge between teaching and learning.
- What methods are actively promoted and being used in schools to routinely and regularly assess pupil's learning? These may include: regular marking of pupil's work, including short tests; oral questioning; more formal school tests or examinations (e.g. end of term or year).
- Are there any programmes focussing on this, or are there any tools and guidance materials available to teachers to support them?
- How is pupils' performance communicated to parents? This is typically done through formal, written end of term or year reports, which may be based on test performances as well as other methods. These may often present an inflated picture of a child's achievement, as schools and teachers may feel exposed to criticism if parents think their children are doing badly.
- What methods are being used to assess the root causes inhibiting pupil learning (e.g. malnutrition, poor school infrastructure and deterioration of natural resources)?
- Are there reosurces in place for schools and teachers to support children who fail assessments?
- Features of many public school systems make classroom assessment difficult. The most significant features are poor infrastructure, large class sizes, overcrowding, weak school management and generally poor conditions of service and resources for teachers. All of these impact on teachers and teaching, and their ability to assess learning.
- Negative effects may include:
  - Confusion of purpose where tests are used for grade progression
  - Rent-seeking by teachers who push extra coaching for weaker pupils
  - Exploitation of pupils – cash or sex for grades
  - Incentives for schools and teachers to inflate results and conceal failure

xiii Smith, D.,"All 25,000 candidates fail Liberian university entrance exam", *Guardian*, 27 August 2013 https://www.theguardian.com/world/2013/aug/27/all-candidates-fail-liberia-university-test

xiv Lagos Police Nabs 4 Suspects In WASSCE Questions Leakage Scandal", *Sahara Reporters*, 27 April 2018 http://saharareporters.com/2018/04/27/lagos-police-nabs-4-suspects-wassce-questions-leakage-scandal

Large-scale assessments

Large-scale assessments are often politically contentious. They can reveal the inconvenient truth that, despite increased levels of investment (more schools, more teachers, more books), learning levels are low. Results are often at odds with the examination system. National assessments may show low levels of literacy and numeracy, yet primary leaving pass rates are high. International assessments can be even more politically sensitive, exposing the differences between countries. Results are frequently contested, disowned and ignored.

A checklist of the main considerations is organised under three areas: (1) understanding the context; (2) building ownership; and (3) ensuring quality. A final section looks at the main strengths and weaknesses of various types of large scale assessment.

Understanding the context

Understanding the overall context within which assessment is taking place is essential in determining appropriate support.

Dimension	Issues to consider
What has been done before?	• What assessments have previously been done? Who did them and what did they show? • How have the results been shared and used? Has there been any response to findings in the form of change in policy or approach? • How are results viewed by stakeholders, especially government?
What are the institutional arrangements for assessments?	• Which organisations have responsibility for undertaking assessment? What is their official mandate? • To what extent are they autonomous and able to operate independently, without undue influence from other parts of the system? Do they have adequate funds and capacities? • What support, technical and financial, might be needed? • Is there any discussion of how assessment findings should be used to improve learning?
What is the state of discourse on learning?	• Are reports/results of assessment clearly presented, readily available and accessible to the public? • Does the debate on education quality and learning in particular, both within the education sector as well as in the media, include evidence drawn from assessments?

Building national ownership

For assessment to ultimately have lasting and sustainable impact on learning, the process has to be truly owned and driven by national governments. Externally executed assessments run a higher risk of failing to establish the necessary credibility with policy and decision-makers compared to those that are implemented through national systems. However, independent assessments, such as UWEZO and ASER have high value in terms of accountability and can provide a vital challenge to government complacency. In addition, there are trade-offs in terms of timeliness and quality, but government-owned and managed national assessments are valid in terms of building sustainable systems and capacity.

Dimension	Issues to consider
Who is driving the process?	• Is assessment largely a donor-driven effort or is there a genuine commitment on the part of government? • Is there any overarching policy or legal framework for assessment? • What is the appetite for taking positive action based on results? • Who are the likely and obvious champions of assessment? Do they have the necessary influence to persuade others?
Is the purpose clear?	• Is the purpose behind assessments clear and understood by all major stakeholders so they are not used for unsuitable purposes? • Has there been a clear and open process of deciding what type(s) of assessments is/are best suited to national needs? • Is there a clear and agreed policy for sharing the results, even if they are disappointing?
How strong is national capacity?	• Does the institution (or institutions) responsible for national assessment have adequate organisational and technical capacity? Consider an institutional analysis or audit prior to any investment. • Has there been any institutional support for assessment in the recent past? What impact did it have? Has there been any institutional twinning between the national assessment unit and more established institutions elsewhere? Consider this as an option for possible support. • It may be that certain highly specialised functions (e.g. psychometrics, item response theory) may not be feasible to develop and sustain in country. Governments may need to consider having to contract in such expertise.

Ensuring assessments are of sufficient quality

Assessments need to be technically sound, effectively implemented and accurately marked with skilled analysis in order to produce findings that have high validity and reliability. A national assessment is a highly technical exercise, often requiring very specific expertise. Analysis of the context will go some way to determining the technical capacity for implementation. Donors can both help build this where weak and support the assessment through well-placed technical support.

Dimension	Issues to consider
Overall design	• Is there a clear rationale underpinning key aspects of the assessment? Most importantly: – **what** is to be assessed? Most commonly, this is reading and numeracy, however important areas such as social and arts education, environmental awareness, values and attitudes, oracy, citizenship and team-working maybe considered important but receive minimal attention. – **who** is to be assessed (e.g. Grade 2 and Grade 5 pupils)? – **what is the frequency** of assessments (eg. repeated every two or three years)? • Is the sample size and design appropriate to generate credible results at desired levels of the system (i.e. by province, region or district) as well as by other considerations (e.g. by ethnicity or language groups)?
Test construction	• Have test items been appropriately developed, pre-tested and standardised so that they are credible and reliable (psychometrics)? • Can results be reliably compared to previous assessments? • What additional information will be collected from learners and others (e.g. teachers, parents) to enable a deeper understanding of learning? • For literacy assessments, what languages will be used and what are the arrangements for minority language users?
Administration	• Are the arrangements for conducting the assessment robust enough to ensure integrity and confidentiality? • What quality assurance measures are in place? • Have all those involved been briefed and trained, and are appropriate safeguards in place to deter malpractice? • Are marking and recording systems robust, with appropriate validation measures?
Reporting	• How are results of assessment communicated? • Are reports produced and disseminated widely? • Are reports written in non-technical language that can be easily understood by stakeholders, including the general public? • Will the underlying causes that affect pupil performance and attendance be identified and reported? • How will the report results be used to influence policy and strengthen the national curriculum, including in subjects that receive mimimal attention (e.g. environment and climate change education)?

Major strengths and weaknesses of various large-scale assessment types

There are numerous pros and cons associated with the three main types of national assessments of learning. Choice will be largely dependent on purpose.

- If the intention is to quickly establish information about what children can do (perhaps a baseline for a major investment to improve education quality) then an externally managed assessment (Type 2) is probably the best option, particularly where any results from national assessments are unreliable.
- If the long term objective is to build national capacity for national assessment then clearly Type 2 is the best option, but will involve substantial national commitment and long term support.
- If the context is more developed and a national assessment system is well established then using an established international assessment system (Type 3) may be the best option.

Type	Strengths	Weaknesses
1. National Assessment implemented through government systems	• Involve local stakeholders • Greater sense of national ownership • Easier to align to the national curriculum • Better able to measure at the right grade or academic level • Usually less costly than external international assessments • Good for accountability purposes	• Open to political interference • May have lower degrees of transparency and credibility • Quality may be compromised by lack of technical expertise and poor administration (this could be mitigated with technical assistance).
2. National Assessment implemented through projects or NGO, e.g. EGRA/EGMA, ASER	• High level of independence from political interference • Technically good: high levels of reliability and validity • Often household based, so includes out-of-school children • Relatively quick • Uses internationally established and tested methods	• Can lack traction with government who may ignore findings • Can exclude local stakeholders • May clash or overlap with other national assessments
3. International and Regional Assessments (e.g. SACMEQ, PASEC)	• Independence from country politics and interference • High transparency and credibility • High technical standards • Can provide capacity-building opportunities	• Can lack alignment to national curriculum • Can underrepresent subjects receiving minimial or no attention (e.g. environmental education) • Do not fully involve local stakeholders • High cost • Less suited for accountability purposes

Conclusion

- There are a number of ways the EU can engage and support countries around learning assessments. Some options include:
- At the very basic level, being aware of the context and issues, including support from other partners and participating in the dialogue as part of the broader dialogue around the education sector.
- Ensure that learning assessment is included in any sector analysis being undertaken in preparation for a possible education support programme, given its importance as part of efforts to improve education quality.
- Commission further analysis to identify specific options for support where learning assessment emerges as a potential component of a support programme. This should include a political economy analysis given the sensitivities around assessment.
- Where support for large scale national assessments is being considered, ensure that support includes both technical support for effective assessment and that results are published and used as a means to improve learning.

3. Case Study

Source	Silvia Favret, EU Delegation in Morocco
Programme	EU support to educational assessment in Morocco has initially focussed on strengthening national capacity to undertake regular learning assessments at primary and lower secondary. Secondly, it aimed to support the response to the results that emerged. The national learning assessment system of Morocco (the PNEA - Programme National d'Evaluation des Acquis) has been supported by the EU through a budget support programme, the Education II programme (2014 - 2019). It has also been supported through an institutional twinning programme between the "Instance Nationale d'Evaluation" (INE), the entity in charge of national learning assessments and monitoring and evaluation of education public policies, and the French Ministry of Education and CIEP[xv]. Additional technical assistance also contributed to the implementation of the education reform in Morocco, with interventions supporting the improvement of education quality, and a focus on the pre-service training of teachers.
Context and challenges	The national programme of learning assessments (the PNEA) of Morocco is the responsibility of the national entity in charge of evaluation, which is part of the Higher Council for Education, Training and Scientific Research. The PNEA complements the exams and continuous assessments undertaken by the Ministry of Education though ad-hoc assessments of students competences. In this way it enables an external evaluation of the education sector performance, in relation to the expected results of the education sector strategy 2015 - 2030 (Vision Stratégique à l'horizon 2030). The first PNEA assessment took place in 2008 (for 2 grades of primary and 2 grades of lower secondary). The results of this first exercise showed average learning achievements below 50% and even lower in rural areas. This had a strong awareness-raising impact on the importance of assessing the performance of the education system in terms of (i) student's learning achievements and (ii) education service quality. This first PNEA, helped to deepen the reflection on the functioning of the education system in Morocco, which resulted in the new "Education Vision for 2030". The PNEA results highlighted two challenges: (i) the need for effective follow-up on PNEA findings at school level, and (ii) the necessity to take specific measures to address system weaknesses. PNEA results confirmed Morocco's results in international assessments – such as TIMSS (assessing and comparing competencies of 4th grade of primary and 2nd grade of lower secondary) and PIRLS (comparing students achievements of 4th grade of primary). Morocco's performance is below international learning achievements standards.
Action taken	EU support to the national assessment programme started in 2014 and has been two-fold: *(i) Institutionalisation of the national learning assessment system (PNEA)* EU supported the institutionalisation of the PNEA via an indicator with three successive annual targets that were included in Education II. The indicators targets focussed on: - the preparation of a second PNEA for 2016, which was realised and published in 2017 and that assessed upper secondary grades, with national coverage - the definition and adoption of the PNEA assessment for the long term, with a planning of PNEA learning assessments every four years, with national coverage, to gradually be expanded to more education levels. Budget support action, was strengthened by regular policy dialogue and technical expertise to improve teachers' competences on students learning assessment activities. In parallel to *Education II*, the programme financed institutional twining aimed at strengthening INEs monitoring and evaluation role, through better data collection, data analysis and the development of a shared information system. *(ii) Reinforcing the set-up of a structured school-level tutoring system for students with learning difficulties* Taking into account the low results of PNEA, TIMSS, PIRLS,PISA assessment and the high dropout rates for secondary levels, the EU also supported the structuring of a school-level tutoring systems for primary and lower secondary-level students experiencing learning difficulties. Morocco EMIS system, the MASSAR, is well-established and students' results are therefore regularly collected. Nevertheless, to anticipate tutoring needs and potential student difficulties, a preliminary assessment is now organised at the beginning of every school year. Its results are also registered in MASSAR. The development of such system was supported by the EU via budget support performance indicators engaging the Government on:

xv CIEP has recently been renamed "France éducation international" https://www.ciep.fr/en

Action taken *continued*	- the development in MASSAR of a specific pre-requisites technical infrastructure and database - at the latest one year later, the deployment of tutoring through remedial class support mechanisms in more than 500 schools (primary and lower secondary) based on the preliminary assessment. Budget support was accompanied by EU technical assistance to the national (and regionals) teacher training institutes (CNIPE – CRMEF) to support the reinforcement of teachers' capacities to detect students' difficulties and changes in their pedagogy to meet the needs of pupils with learning difficulties. Besides the upgrading of teachers training curricula, the EU support also accompanied the revision of secondary-level science and physics manuals, to integrate a remediation approach for learning (for teachers and students).
Impact	The capacity of the entity in charge of national evaluation (INE) is strengthened and the PNEA system for the national assessment of learning has been institutionalised and operationalised (next PNEA will be published in 2020). This is, however, still work in progress. PNEA should be gradually expanded to cover more grades and levels to allow comparability between assessments over time and to better evaluate the impact of the reforms in terms of learning achievements. A system of remedial support at school level has been set up for primary and lower secondary, which is responding to specific needs of the students. This system is yet to be progressively implemented nationwide as, for the moment, it is not yet in place in all the schools in Morocco. More than 2500 schools (primary and lower secondary) have operationalised the tutoring, out of nearly 9000 establishments. As for teacher training, the "remediation approach" has been integrated in some manuals as well as in teacher training curricula.
Lessons learned	The set up of the national learning assessment system is an important element, feeding into the implementation of the education sector reforms. The findings of the first two rounds prompted the set up of a remedial system at basic-education level. The full articulation of the national assessments and the measures to be taken at school level are still work in progress. Further work will be needed to ensure that the national assessments reports are shared at school level and that the findings are used to possibly adjust interventions at that level. In parallel, it is important the Ministry of Education also reinforce its capacities in terms of more routine learning assessments.
Further information	PNEA / INE: https://www.csefrs.ma/ine/presentation-de-linstance-nationale-devaluation/linstance-nationale-devaluation/?lang=fr Education II: http://educ2.men.gov.ma/

4. References and Further Reading

Clarke, M., *What Matters Most for Student Assessment Systems: A Framework Paper,* The World Bank: Washington D.C., 2012 http://siteresources.worldbank.org/INTREAD/Resources/7526469-1335214323234/WP1_READ_web_4-19-12.pdf

HEART, *Helpdesk Report: Learning Assessments,* 2015 http://www.heart-resources.org/wp-content/uploads/2015/10/HEART-Helpdesk-Learning-Assessments-FOR-WEB1.pdf

UNESCO, *Learning assessment at UNESCO: Ensuring effective and relevant learning for all,* 2016 https://en.unesco.org/system/files/private_documents/learning_assessments_brochure-eng-web.pdf

UNESCO Institute for Statistics, *The Impact of Large-Scale Learning Assessments,* 2018 http://uis.unesco.org/sites/default/files/documents/impact-large-scale-assessments-2018-en.pdf

UNESCO Institute for Statistics, *Quick Guide No. 2: Making the Case for a Learning Assessment,* 2018 http://uis.unesco.org/sites/default/files/documents/quick-guide2-making-case-learning-assessments-2018-en_2.pdf

International Assessments Of Learning

PISA – OECD's Programme for International Students Assessment http://www.oecd.org/pisa/

PISA-D – OECD's PISA for Development http://www.oecd.org/pisa/pisa-for-development/

PIRLS – Progress in International Reading Literacy Study https://www.iea.nl/studies/iea/pirls

TIMSS – Trends in International Mathematics and Science Study https://www.iea.nl/studies/iea/timss

SABER – World Bank's Systems Approach for Better Education Results http://saber.worldbank.org/index.cfm?indx=8

SACMEQ – The Southern and Eastern Africa Consortium for Monitoring Educational Quality http://www.sacmeq.org/

PASEC – Programme d'Analyse des Systemes Educatif de la CONFEMEN http://www.pasec.confemen.org/

LLECE – The Latin American Laboratory for Assessment of the Quality of Education http://www.unesco.org/new/en/santiago/education/education-assessment-llece/

Publications Office of the European Union