



GLOBAL
PARTNERSHIP
for EDUCATION



Analysis of National Learning Assessment Systems [ANLAS]

ANLAS Ethiopia: Country Report

June 2019

ANLAS Ethiopia National Team

| | |
|--------------------------------------|-----------------------|
| Mesaye Demessie Zeleke (Focal point) | Aregawi Gidey Mesele |
| Yilikal Wondimeneh Demissie | Bekele Geleta Ayana |
| Arega Mamaru Yewore | Effa Gurmu Bati |
| Nega Gichile Bongasse | Getachew Abebe Tsigie |
| Abraham Mengistu Sertse | Belay Endeshaw Gizaw |
| Talefe Eshete Awoke | Robel Getachew Worku |

Table of Contents

| | |
|---|-----------|
| ANLAS Ethiopia National Team..... | 1 |
| Acknowledgements | 3 |
| List of Abbreviations | 4 |
| 1 Introduction..... | 5 |
| 1.1 ANLAS purpose | 6 |
| 1.2 The ANLAS model..... | 6 |
| 2 ANLAS in Ethiopia..... | 8 |
| 2.1 Implementation process..... | 8 |
| 2.2 National team training and planning..... | 9 |
| 2.3 Stakeholder briefings..... | 9 |
| 2.4 Completing the analytical tables | 9 |
| 2.5 Preparing and disseminating a country report to relevant stakeholders to inform education sector planning | 12 |
| 3 Findings of ANLAS..... | 13 |
| 3.1 Evaluation of key areas..... | 13 |
| 3.2 Aspects of improvement and recommendations | 15 |
| 4 Using the ANLAS findings | 41 |
| Appendix 1: Focal point and national team | 42 |
| Appendix 2: Stakeholder and document mapping tables | 43 |
| Mapping table II: Overview per key stakeholder group | 43 |
| Mapping table III: Overview per document..... | 44 |
| References | 46 |

Acknowledgements

The Global Partnership for Education (GPE), under the leadership of Ramya Vivekanandan and support of Medjy Pierre-Louis, has initiated the Analysis of National Learning Assessment Systems (ANLAS). It has assigned the development of ANLAS to the Australian Council for Educational Research (ACER).

The ANLAS development process includes three phases: 1) development of the ANLAS toolkit (July 2018 to January 2019); 2) pilot implementation of ANLAS in three GPE partner countries (January to June 2019); and 3) revision and refinement of ANLAS based on observations and experiences from the pilot implementations (June to August 2019).

Ethiopia is one of three countries participating in the piloting of ANLAS. The piloting countries are contributing significantly to the consultative development process of ANLAS, allowing to test both, the content of the ANLAS toolkit and the process of using these resources to analyze the national learning assessment system and to make recommendations for improvement to inform strategies within education sector plans.

The Ethiopia Ministry of Education endorsed the ANLAS study and nominated knowledgeable and responsible national team members. In this regard, special acknowledgement should go to H.E. Dr Tilaye Getie, Minister of MOE, and H.E. Ato Mohamed Ahmedin, state minister of MOE, for their concern and supporting the team in all respects. The team also extends its thanks to Ato Araya Gebre-Egziabher, Director General of NEAEA, for his facilitation and for creating a conducive environment for the team and the study as well.

Sally Robertson (ACER) supported the national team throughout the piloting process and Ursula Schwantner (ACER) provided support during the in-country visit.

Last but not least, the national team is grateful to acknowledge all the ANLAS participants drawn from teachers, senior experts in education, school leaders, training providers, development partners, and parents for providing the necessary information, active participation and their concern.

List of Abbreviations

| Abbreviations | Explanations |
|----------------------|---|
| ACER | Australian Council for Educational Research |
| A4L | Assessment for Learning |
| CTE | College of Teachers Education |
| DCPs | Developing Country Partners |
| DFID | Department For International Development |
| EGMA | Early Grade Mathematics Assessment |
| EGRA | Early Grade Reading Assessment |
| EMIS | Educational Management and Information System |
| ESA | Educational sector Analysis |
| ESDP | Educational Sector Development Program |
| ESPs | Education Sector plans |
| ETP | Education Training Policy |
| GAML | Global Alliance to Monitor Learning |
| GEQIP | General Education Quality Improvement Package |
| GEQIP-E | General Education Quality Improvement Package for Equity |
| GER | Gross Enrollment Rate |
| GPE | Global Partnership for Education |
| GP-LA | Good practice in Learning Assessment |
| MoE | Ministry of Education |
| n.d | No data |
| NEAEA | National Education Assessment and Examinations Agency |
| NLA | National Learning Assessment |
| NLSA | National Large-Scale Assessment |
| PISA | Programme for International Student Assessment |
| PR | Public Relation |
| REBs | Regional Education Bureaus |
| READ | Russia Education Aid for Development |
| SABER | Systems Approach for Better Education Results |
| SACMEQ | Southern and eastern Africa Consortium for Monitoring Educational Quality |
| SDG4 | Sustainable Development Goal 4 |
| SNNP | Southern Nations and Nationality People |
| TIMSS | Trends in International Mathematics and Science Study |
| UIS | UNESCO Institute for Statistics |
| UN | United nations |
| UNICEF | United Nations International Children's Emergency Fund |
| USID | Universal, Sustainable, Innovative Design for social change |

1 Introduction

Ethiopia is located in north-east Africa known as "the Horn of Africa". It has an area of 1.1 million square kilometers. The country has an estimated total population of over 90 million with highly diversified ethnic groups, languages, culture and topography. The country has adopted federal governance with nine regions and two city administrations. The Prime Minister is the head of government, and executive power is exercised by the government. The federal legislative power is vested in both the government and the two chambers of parliament.¹



The current Education and Training Policy (ETP) was enacted in 1994 to alleviate the problems of access, curriculum relevance, equity, quality and related challenges of the education sector.² As a result of the policy priority and financial input for the sector, Ethiopia made a significant progress to achieve the target set for primary education to be universal by 2015. For instance, the gross enrollment rate (GER) of primary education reached to 109.3% in 2017/18. During this time, the number of students and teachers were 26,905,580 and 583,461 respectively. The number of public and private schools had also reached 41,441 in 2017/18.³

The use of assessment in Ethiopia was started in 1938 with National Examinations. However, NLA has been implemented very lately in 2000. Although classroom assessment was introduced along with modern education in 1908, it has been given more attention in the new Education and Training policy of 1994. To monitor and evaluate the quality of education at a regular basis, National Educational Assessment and Examinations Agency (NEAEA) was re-established by Council of Ministers Regulation No. 260/2012 by integrating the two assessment systems, National Examinations and NLA.⁴ The National Examination is a census based in grade 10 and 12, and National Learning Assessment has been carried out on sample based in the exit cycle of primary and secondary education (grades 4, 8, 10, and 12). Early grade literacy and numeracy assessment on the other hand is being conducted at grades 2 and 3. Both national Examinations and national learning assessment have been administered on the basis of the national education and training policy and curricula of the country.

Though tremendous achievement observed access of education for all citizens, the quality of education as indicated in successive national learning assessments and national examinations is still a critical challenge for the nation. To this end, the assessment system should be strengthened in order to support evidence-based intervention and decision making.

¹ (The Federal Democratic Republic of Ethiopia National Commission Central Statistics Agency 2016)

² (Transition Government of Ethiopia 1994)

³ MoE 2018

⁴ (Council of Ministers Regulation No.260/2012 2012)

1.1 ANLAS purpose

Learning assessments are increasingly used by education systems around the world to inform evidence-based policymaking and reform processes, and to improve teaching and learning practice.⁵ To ensure the learning assessment systems sustainability and effectiveness, it is important to regularly review and evaluate their actual state and progress.

For the Global Partnership for Education (GPE), an important priority is the quality of learning assessment systems as a key component of strong education systems and a vital means to monitor and improve learning outcomes. GPE is a multi-stakeholder partnership and funding platform that focuses on supporting the efforts of more than 65 developing country partners (DCPs) in building effective education systems to improve equity and learning. In addition to its efforts to reinforce learning assessment systems through its grants to DCPs, GPE is working to strengthen learning assessment systems and promote a holistic measurement of learning through its Assessment for Learning (A4L) initiative, launched in 2017.

In this regard, the Analysis of National Learning Assessment Systems (ANLAS) is a key component of GPE's A4L initiative, aiming to support partner countries to build effective monitoring systems for evidence-based education sector planning.⁶

Therefore, ANLAS is designed to enable country partners to undertake a comprehensive analysis of national learning assessment systems, covering learning assessments in school education from primary or basic education to secondary education, all schools within the system (public, private and community), at central as well as decentralized levels. The aim of the qualitative analysis is to identify areas and recommendations for improvement of the national assessment system, to inform the development and implementation of improvement strategies as part of the wider education sector planning process.

As part of the intervention, to guide and assist developing country partners in undertaking ANLAS, a framework and toolkit have been developed. The toolkit consists of process tools to support the implementation of ANLAS, analytical tools to document and analyze the learning assessment system, and reporting and dissemination tools.

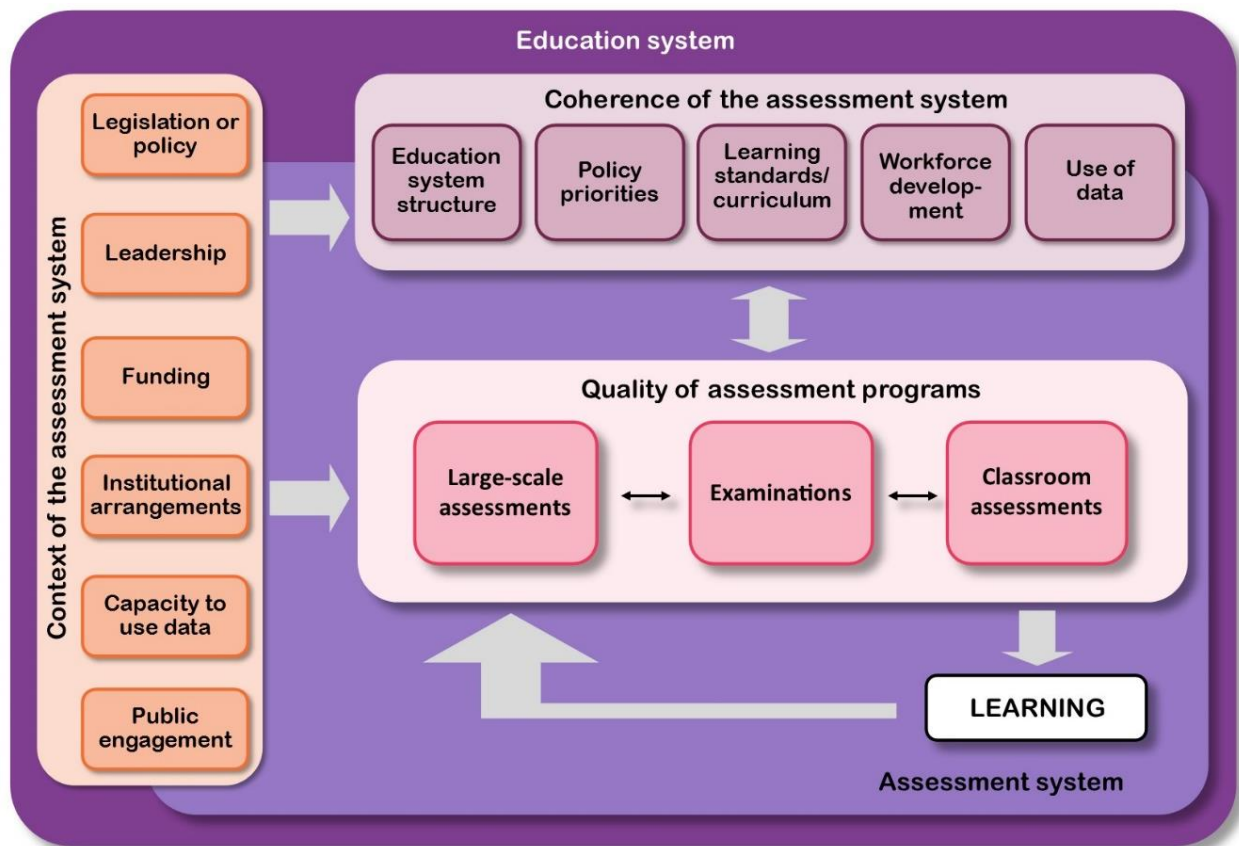
1.2 The ANLAS model

The ANLAS model illustrated in Exhibit 1 provides the analytical framework for the qualitative analysis of national learning assessment systems.

⁵ John C. Cresswell, "System-Level Assessment and Educational Policy," (Melbourne: ACER, 2017).

⁶ Global Partnership for Education, "Concept Note: Analysis of National Learning Assessment Systems (ANLAS)." (Washington, DC: Global Partnership for Education, 2018).

Exhibit 1: The ANLAS model



The qualitative analysis includes the following *three focus areas*:

1. **Context** refers to the broader context of the assessment system and the extent to which it is supportive of the assessment system⁷. Context covers six key areas as shown in the ANLAS model.
1. **Coherence** refers to the extent to which the assessment system is aligned with, or coherent with important aspects of the broader education system⁸. Coherence covers five key areas as shown in the ANLAS model.
2. **Assessment Quality** refers to the overarching key quality concepts for learning assessment programs: technical rigor, fitness for purpose, clarity and consistency of purpose objectivity and independence, transparency and accountability and ethicality and fairness. The assessment programs covered in ANLAS include large-scale assessments (national, international/regional), examinations and classroom assessments. Eight key areas are operationalized for quality of large-scale assessments and examinations, and seven key areas for quality of classroom assessments.

⁷ Marguerite Clarke, "What Matters Most for Student Assessment Systems: A Framework Paper: Systems Approach for Better Education Results (SABER) Student Assessment Working Paper No. 1," (Washington, DC: World Bank, 2012); Network for African Learning Assessments (NALA), "Learning Assessments Systems Evaluation Framework (Draft)," (n.d.), accessed January 4, 2019, http://www.adeanet.org/adeapmp/sites/default/files/activities/learning_assessments_systems_framework_v18.pdf

⁸ Clarke, "What Matters Most".

An important **cross-sectional element** of ANLAS is the domain of **21st century skills**. These skills, or ‘general capabilities’ and ‘transversal competencies’ as they are often referred to, are an increasing policy priority for education systems around the world⁹. The diagnostic nature, the cross-sectional consideration of 21st century skills and the integration of the analysis into the education sector planning process are essential, distinct features of the initiative.

2 ANLAS in Ethiopia

Education is used as a means for making the nation competent in the current knowledge-based economy. For this to happen, Ethiopia has been making every effort for Quality Education for All and ensuring that learning is taking place. In this regard, Ethiopia has been conducting national large-scale assessment to check students’ achievement levels in a timely and technically defensible manner in order to produce relevant data that help monitor and improve the quality of education by its sample based large scale assessment at particular grades. National examinations, which are census based, have also been practiced for many decades to select or certify students as they move from one level of the education system to the next. Classroom assessment is also used to provide real-time information that support ongoing teaching and learning.

To compare the country’s education system performance with other countries, participating in an international and/or regional assessment like PISA/TIMSS or SACMEQ seemed highly valuable. However, due to various reasons, Ethiopia has not yet participated in any international as well as regional assessments.

ANLAS provided a useful opportunity to undertake a comprehensive review of these different programs that comprised the national learning assessment system.

2.1 Implementation process

The implementation of ANLAS in Ethiopia was undertaken in six steps, commencing in July 2018 and ending in June 2019. Exhibit 2 outlines the dates and duration of the six ANLAS steps.

Exhibit 2: ANLAS implementation process in Ethiopia

| ANLAS implementation step | Date and duration |
|---|--|
| 1. Nominating a focal point and establishing a national team | 16– 23/08/2018 One week (the number of National team members was increased on 23/01/2019) |
| 2. Familiarization with the ANLAS toolkit | 14- 23/01/2019 3 days |
| 3. Team training and planning | 23- 25/01/2019 3 days |
| 4. Stakeholder briefings | 29/01/2019 |

⁹ Global Partnership for Education, "Concept note".

| | |
|---|--|
| | 1 day |
| 5. Completing the analytical tables: description, evaluation, recommendations for improvement, and synthesis of findings | February 11 – March 08/2019 4 weeks |
| 6. Preparing and disseminating a country report to relevant stakeholders to inform education sector planning | March 13 – June 27/2019 15 weeks |

2.2 National team training and planning

The national team consisted of 12 members, including the focal point, representing MoE & NEAEA. The members are listed in the Appendix along with their roles and organizations.

The national team training was undertaken to develop a shared understanding of the ANLAS framework, tools and processes. The national team planning was aimed to develop a detailed ANLAS implementation plan and identify the key stakeholder groups to be involved.

2.3 Stakeholder briefings

The stakeholder briefings aimed to create a common understanding of the objectives and implementation process of ANLAS among the national team and key stakeholders. The stakeholder groups involved were federal and regional government education senior official representatives, the NEAEA directors and senior experts, development partners, public relation (PR) officers. Thirty-two stakeholders participated in the stakeholders briefing session which was held in the NEAEA hall in Addis Ababa, Ethiopia.

2.4 Completing the analytical tables

Analytical tables and instructions for completing them were provided to document and guide the analysis of the three ANLAS focus areas. Within each of the focus areas, several key areas were analyzed by using sets of guiding questions. Besides, for each key area, a quality objective was defined, against which the key area was described and evaluated. These quality objectives are indicated as part of the ANLAS findings. The analysis of the national learning assessment system consisted of three major steps:

1. describing the key areas for each ANLAS focus area,
2. evaluating the key areas for each ANLAS focus area, identifying aspects for improvement and making recommendations, and
3. synthesis of the findings.

This section describes these three steps. The findings from the analysis are presented in section 3.

2.4.1 Describing the key areas of the national learning assessment system

The following assessment programs were included in the analysis:

1. National examinations (for grade 10, certification of secondary school and grade 12

- admission for higher educational institutes administered in each year),
2. Sub-national examinations responsible for each region (grade 8, certification of primary school administered in each year),
 3. National learning assessment (given for lower grades 2 and 3 like EGRA and EGMA and exit levels of primary and secondary at grades 4, 8, 10 and 12 for every four-year intervals), and
 4. Classroom assessments.

Currently there are no regional/international assessments implemented in Ethiopia. However, the analysis also provided some useful insights into future options for these assessments.

The description of the national learning assessment took place alongside the evaluation of the key areas. The processes and sources of data used to undertake this analysis are described below.

2.4.2 Evaluating the key areas against quality objectives and making recommendations for improvement

The sources of data used in ANLAS were public documents, and data collected from stakeholder consultations through interviews and focus group discussions (FGD). Semi-structured interviews and FGD were used with a total of 80 stakeholders from education leaders, experts, teachers of primary and secondary schools, cluster supervisors, parents, and Development Partners like USAID, UNIECF, DFID, and World Bank as shown in Exhibit 3. Primary school teachers and leaders were consulted from selected cluster schools in the three regions and one city administration: Oromia regional state (Qobo Luto Primary school), Amhara regional state (Chacha Primary school), SNNP regional state (Tiya Primary School) and from Addis Ababa City Administration (Minelik II Primary School). Training providers were also coming from different regions' College of Teacher Education (Asela CTE, Jigjiga CTE, Arbaminch CTE, Harari CTE, Hosaena CTE, Dilla CTE, Bonga CTE, Adwa CTE and Abiy Adi CTE). Senior experts from regional education bureaus (REBs) were seven in number: SNNP 3, Tigray 1 and Oromia 2, and Somali 1, were participated in ANLAS focus group discussions.

The stakeholders were consulted by briefly explaining ANLAS purpose and its toolkits by the national team and then participants were responded their ideas accordingly. The national team has facilitated and recorded their responses based on the ANLAS stakeholder mapping tables. Efforts were also made to triangulate the information provided by the interviewees using subsequent interviews or public documents. Furthermore, different documents like the current Education and Training policy, Policy frameworks for assessment and examinations, manuals and guidelines, assessment reports were used as summarized in Appendix 2.

Exhibit 3: ANLAS consulted stakeholders

| Number | Stakeholder group | Number of participants |
|--------|--|------------------------|
| 1 | National or sub-national level officials | |
| | Federal MoE senior experts and directors | 7 |
| | NEAEA senior experts and directors | 13 |
| | Regional/subnational Education Bureau senior experts and directors (SNNP 3, Tigray 1 and Oromia 2, and Somali 1) | 7 |
| 2 | Training Providers | |
| | College of Teacher Education (Asela CTE, Jigjiga CTE, Arbaminch CTE, Harari CTE, Hosaena CTE, Dilla CTE, Bonga CTE, Adwa CTE and Abiy Adi CTE) | 9 |
| 3 | School leaders and teachers | |
| | Primary school leaders and supervisors | 5 |
| | Primary school teachers | 19 |
| | Secondary school teachers | 9 |
| 4 | Development partners | |
| | USAID | 2 |
| | UNICEF | 2 |
| | DFID | 1 |
| | World Bank | 2 |
| | Total | 76 |

Before communicating with informants, prior communication through letter and/or phone was made. Face to face contact and briefings, summarizing the analytic tables and sending via email and printed copies were some of the modalities used for stakeholder consultations.

The participants' responses were crosschecked and triangulated with pertinent official documents. Moreover, the data collectors (the national team) were reading the points or summary notes of the focus group discussion to the participants for confirmation of consensuses reached.

For each of the three ANLAS focus areas, the key areas were evaluated against quality objectives (see Section 3). Three evaluation categories are differentiated as shown in Exhibit 4.

Exhibit 4: Evaluation categories

| Category | The quality objective is ... |
|----------|---|
| 1 | Achieved (currently no improvements needed) |
| 2 | Partly achieved (improvements can be made) |
| 3 | Not achieved (improvements are required) |

For focus areas with an evaluation category 2 or 3, recommendations for improvement were identified. These recommendations are aimed to inform the development and implementation of improvement strategies as part of the wider education sector planning process.

2.4.3 Limitations

Though the national team composition was from relevant departments of the ministry and as a result there was no significant challenge in accessing pertinent documents for the assessment and coverage of the focus and key areas, the availability of the key stakeholders for consultations at various levels was a serious challenge. It was very demanding for the national team to get key informants for face to face discussion for more than half a day. To tackle such problem, various data collection strategies/modalities like strategies such as identifying the relevant guiding questions from the analytical tables for the specific stakeholder group were used. Some key stakeholders couldn't attend the stakeholders' briefing and consultation session, therefore the national team was engaged for more additional time for briefings about the intention and implementation of ANLAS.

2.4.4 Synthesis

The information from the analysis of the three focus areas was consolidated. This involved developing an overview table which described the evaluation categories, the aspects and recommendations for improvement, for all three ANLAS focus areas. This information is presented in Section 3.2.

2.5 Preparing and disseminating a country report to relevant stakeholders to inform education sector planning

After completing the data collection, evaluation and synthesis, the report was organized by the national team that grouped into three focus areas (FAs). The national team also organized the necessary evidences that support issues related to this diagnostic study from different official documents.

This report was developed to present and disseminate the findings from the analysis to all stakeholders of government officials, developmental partners and senior education experts and teachers and the public at large through panel discussion, mass media, and uploaded on web site.

Then, the study findings will have been used for developing educational sector development plan (ESDP VI) to fill the gaps of assessment system using documents and discussion with policy makers and MoE officials since September 2019. Further information about how the ANLAS findings will be used is provided in Section 4.

3 Findings of ANLAS

This section provides a summary of the outcomes of the qualitative analysis of the national learning assessment system in Ethiopia.

3.1 Evaluation of key areas

The focus areas included in this part are context of assessment, coherence of assessment and quality of large-scale assessment, examination and classroom assessment. The evaluation of each key area in the three focus areas is summarized as follows.

In the first focus area 'context of assessment system' has six key areas. One key area, institutional arrangements, was rated as 'achieved'. The rest five key areas: legislation/policy, leadership, funding, capacity to use assessment data and public engagement were rated as 'partly achieved' indicating improvements may be needed in these areas.

The second focus area; coherence of the assessment system has five key areas. Among the five key areas the first key area 'structure of the education system' was rated as achieved. On the other hand, the remaining four key areas (education policy priorities, learning standards and curriculum, school education workforce development and use of data) were rated as partly achieved, which may need some improvements.

The third focus area is about the quality of large-scale assessment, public examinations and classroom assessment. The focus area has two parts (part A & B). Part A refers to large scale assessment and examinations. It consists of three sub-components: national large-scale assessment (NLSA), national examinations and sub-national examinations. The quality of large-scale assessment has eight key areas: organization and resources, assessment framework, assessment instruments, sampling, field operation, data management, data analysis, reporting and dissemination.

In NLSA, five key areas (Assessment instruments, sampling, field operations, data management & data analysis) were achieved while other three key areas (organization and resources, assessment framework, reporting and dissemination) are partly achieved. For the national examination, five key areas (assessment instruments, field operations, data management, Data analysis, reporting and dissemination) were achieved whereas two key areas (Organization and resources, assessment framework) were partly achieved. Moreover, in the sub-national examination two key areas (field operations and reporting and dissemination) were achieved and the other five key areas (organization and resources, assessment framework, and data analysis) were partly achieved. Since the national and sub-national examinations are census based, sampling was not applicable.

Part B refers to classroom assessment and has seven key areas: availability of guidelines, capacity building for teachers, resources and tools, assessment methods, assessment content, mechanism of quality control, and use of assessment data. All the key areas were partly achieved. Exhibit 5 shows the results of the evaluation categories for each ANLAS focus area and key area.

Exhibit 5: Evaluation results

| 1. Achieved (currently no improvements needed) | 2. Partly achieved (improvements can be made) | 3. Not achieved (improvements are required) |
|--|---|---|
| Focus area 1: Context of the assessment system | | |
| 4. Institutional arrangements | 1. Legislation or policy | |
| | 2. Leadership | |
| | 3. Funding | |
| | 5. Capacity to use public engagement | |
| | 6. Public engagement | |
| Focus area 2: Coherence of the assessment system | | |
| 1. Structure of the education system | 2. Education policy priorities | |
| | 3. Learning standards and curriculum | |
| | 4. School education workforce development | |
| | 5. Use of data | |
| Focus area 3.A: Quality of large-scale assessments and examinations | | |
| <i>[Program 1 – National Large-scale Assessment (NLSA)]</i> | | |
| 3. Assessment instruments | 1. Organization and resources | |
| 4. Sampling | 2. Assessment framework | |
| 5. Field operations | 8. Reporting and dissemination | |
| 6. Data management | | |
| 7. Data analysis | | |
| <i>[Program 1 – National Large-scale Assessment (NLSA)]</i> | | |
| <i>[Program 2 – National Examination]</i> | | |
| 3. Assessment instruments | 1. Organization and resources | |
| 5. Field operations | 2. Assessment framework | |
| 6. Data management | | |
| 7. Data analysis | | |
| 8. Reporting and dissemination | | |

| [Program 3 – Sub-national Examination] | | |
|---|-------------------------------|--|
| 5. Field operations | 1. Organization and resources | |
| 8. Reporting and dissemination | 2. Assessment framework | |
| | 3. Assessment instruments | |
| | 6. Data management | |
| | 7. Data analysis | |
| Focus area 3.B: Quality of classroom assessments | | |
| | 1. Guidelines | |
| | 2. Training | |
| | 3. Resources and tools | |
| | 4. Assessment methods | |
| | 5. Assessment content | |
| | 6. Quality assurance | |
| | 7. Use of assessment data | |

3.2 Aspects of improvement and recommendations

Based on the analysis of the data obtained from different stakeholder groups and documents, some major gaps that need improvements were identified and possible solutions were recommended by the stakeholders in group discussions to enhance the effective implementation of the assessment systems. The findings, improvement and recommendations in each FAs are briefly discussed below.

Focus area 1: Context of the assessment system

The context of the assessment system comprised of six focus areas. Five of them were evaluated as partly achieved. One key area (key area four – institutional arrangement) was evaluated as achieved. The details for each key area are described below and summarized in Exhibit 6.

With regard to Legislation or Policy, the data obtained from the analysis of stakeholders and document consultations depicted that the country has Education and Training policy. The policy guides the overall education system and policy frameworks, manuals, guidelines and internal directives are also developed for NLA and public examinations under the education policy. However, the analysis identified that there is a gap between policy provision and implementation of policy. Particularly, it was found that there is inconsistency of the evaluation criteria in the policy document and the grading system in national examinations. While fixed criteria for evaluation in the education policy, norm referenced grading system is practiced in grade 10.

According to the information obtained from FGD, it is clearly described that there is encouraging political commitment by leadership on promoting and supporting the assessment system by developing policy documents, allocating budget, establishing

institutional arrangements that can help to carry out the assessment, communicating the public about the purpose of the assessment and promoting different development partners to engage in the assessment system. In spite of this, the leadership has a limitation in the capacity to use the assessment results for making evidence-based decision and intervention.

From the analysis in Exhibit 6 under key area 3, it is found that NLSA, the government covered only 24% budget, the rest (76%) of the budget is covered from donor sources. However, National examinations, the government allocated better budget amount (75%), while the rest 25% is collected from private students. Moreover, as per the response of key informants showed that there is no clearly defined budget for class room assessment. In conclusion, there is budget inadequacy for the large-scale assessment, national examinations and classroom assessment practices.

Regarding the instructional arrangement, the study found that, the government has a separate institution (NEAEA), responsible for designing, analyzing, implementing analyzing and use of data for evidence-based decision making. Besides, there are different departments in the ministry, which are expected to document and utilize the assessment results for better intervention.

The purpose of the learning assessment is to determine the level of students' academic achievement and identify factors influencing achievement and finally to generate recommendations for policy making to improve education quality. The government has the capacity to use data from learning assessments for evidence-based decision making in education policy and practice. However, there is a limitation in dissemination and utilization of the assessment results for intervention during formative evaluation.

Concerning the public engagement, there is a good experience in community participation to support the assessment system. However, there is a gap in having common understanding about the purpose of the assessment. Yet, there is a limitation in addressing updated guidelines and manuals for stakeholders like teachers, principals, supervisors.

Focus area 2: Coherence of the assessment system

Focus area two – Coherence of the assessment system – was comprised of five focus areas. Four of the five key areas were evaluated as partly achieved. One key area (key area one – structure of the education system) was evaluated as achieved. For each of the areas that were partly achieved, the aspects and recommendations for improvement were identified. These are described below and summarized in Exhibit 6.

Concerning the second key area- education policy priorities, although 21st century skills were well addressed in the Ethiopian curriculum framework. However, the assessment system as well as the implemented curriculum like students text books and other instructional materials were not articulated in line with the curriculum framework so that 21st century skills were not well addressed. More specifically, these 21st century skills were not adequately assessed within the large-scale assessment system since the instruments were only multiple response test items which are not suitable to assess most performance skills. Therefore, it is recommended that there is a greater focus to incorporate the constructed response test items that are helpful to assess the 21st century skills in the future. There is also a lack of fair alignment among the assessment programs to provide relevant information for policy priorities. EMIS include only the performance data from the national and regional examinations in its annual report. Therefore, it is recommended that

EMIS should organize and provide the aligned data, from major NLA findings and classroom assessment results.

With regard to learning standards and curriculum, aligning classroom assessment with the curriculum standards (competencies) and making continuous classroom assessment part of the instruction are key areas that require improvement. Attempts should be made to assess language practical skills and activities like listening, writing and speaking in language as well as laboratory and other hands on performances. The data from classroom assessment should be properly documented and utilized for further investigations of policies, standards, curriculum and instructional methods that the teacher used. In this case, provision of capacity building and development of appropriate instructional materials and improvement of the assessment system are suggested.

Concerning school education workforce development, alignment of the assessment programs with national official documents and integrating use of data with capacity building programs were issues that need improvement that is why the students result in classroom assessment and national assessments were significantly different according to the respondents. This was supported by findings in the NLA at each grade levels every cycle in which they achieve high in the classroom assessments. As recommendations, aligning training materials with the official documents and incorporating the findings of the assessment data in capacity building programs were forwarded.

Assessment data is used by various key stakeholders for evidence-based decision making in education policy and practice. However, producing quality classroom data and relevant data to particular regions according to their contexts, disseminating the findings to all stakeholders on time particularly to the lower levels of education systems such zones, districts and schools, limitation in using the data (recommendations) for example, curriculum revisions and instructional improvements, and establishing accountability to use the data from the assessment systems were areas identified as drawbacks.

Hence, producing quality and relevant data, improving its accessibility, establishing policies and regulations regarding accountability as well as providing appropriate trainings on using data are recommended.

Focus area 3.A: Quality of large-scale assessment and examination

Large-scale assessment in Ethiopia's context indicates National Large-Scale Assessments (NLSA) programs including EGRA and EGMA supported by government and donors. It is carried out to provide information on education system accomplishment and the factors contributing to educational outcomes, to inform education policy and practice. Large-scale assessments in Ethiopia use sample of the target population.

Similarly, Examination in Ethiopia's perspective describes both National and Sub-National Examinations under the auspices of federal government. National Examinations conducted at national level to make decisions for certification at the end of grade ten and for university entrance selection purposes at the end of grade twelve. Sub-National Examinations conducted at national states level to make decisions for certification at the end of grade eight. National and Sub-national Examinations in Ethiopia use a census approach to test all eligible students.

In this part of analysis, within 'quality of large-scale assessment and examination', evaluations against quality objectives for eight key areas have been made. Accordingly, in the NLSA program five key areas namely, assessment instruments, sampling, field

operations, data management, and data analysis were informed by different stakeholders to be achieved. Nevertheless, organizations and resources, assessment framework and reporting and disseminations were the three key areas described to be partly achieved in the NLSA program.

In the same way, five key areas that are, assessment instruments, field operation, data management, data analysis, and reporting and dissemination in the National Examination program were identified to be achieved. However, organization and resource, assessment frameworks were the two key areas suggested that improvements could be made in the National Examination.

In the Sub-National Examination, the two key areas specifically, field operations and reporting and dissemination were consulted by different stakeholders to be achieved. On the other hand, organization and resource, assessment framework, assessment instruments, data management and data analysis were the five key areas referred to be achieved partly in the Sub-national Examination program.

Based on the data gathering operation in the 'large-scale assessment and examination' programs, it was found out that there was no key area informed to be 'not achieved'.

In addition to the aforementioned evaluation part, the following recommendations for improvement for eight key areas have been forwarded. As a result, organization and resources was the common key area that required improvement in the three large scale assessment programs (i.e. in the NLSA, National Examination, and Sub-National Examination programs). Regarding organization and resource, it was found out that there have been limited capacities in the assessment management and staff, inadequate physical facility provisions to complete the diverse task in the assessment programs. Thus, it was recommended that assessment management and staff should be strengthened and upgrade their skills in the assessment and related field with short or long-term training in collaboration with government and donors. Added to this, adequate office rooms, essential software, scanners should be provided for assessment staff by the government in the three large scale assessment programs.

Assessment framework was also the common key area that was achieved partly in the three large scale programs. It was known that there was no assessment framework prepared for NLSA, National Examination and Sub-National Examination that provides a rationale for contextual data collected and specifies the design of the assessment program. Hence, it was recommended that the assessment framework should be prepared for the NLSA, National and Sub-National examinations programs.

Moreover, assessment instrument was the key area common to the NLSA and National examination programs and currently required no improvement. Conversely, it was recognized that assessment instrument was the key area that was achieved partly in the Sub-National program. It was found out that there was no piloting of test instruments in the Sub-National Examination to ensure the reliability, validity and fairness of the tests. Therefore, it was recommended that reviewing and pre-testing of test instrument ought to be in place in the Sub-National program.

Sampling was the key area that presently no improvement needed in the NLSA program. As the data gathering approach was census base, this key area was not an issue in the National and Sub-National programs.

Field operation was the common key area that was achieved in the three large scale

assessment programs.

Data management was an area that needed no improvement in the NLSA and National Examination programs. However, data management was partly achieved in the Sub-National program. It was identified that there was no data management manual for Sub-national Examination that ensure appropriate quality data structure and documentation. Thus, it was suggested that data management manual should be prepared for Sub-national Examination.

Data analysis was the key area that needed no improvement in the NLSA and National Examination programs. However, it was partly achieved in the Sub-National program. It was found out that there was no sound data analysis technique in the Sub-National Examination to provide valid and useful analytical inferences about the population of interest. Therefore, it was recommended that the sun-national education bureaus ought to provide scoring and data processing machines in the Sub-National Examination program.

Reporting and dissemination was the other common key area that needed no improvement in the National and Sub-National Examinations. However, it was known that the improvement could be made in this area in the NLSA program for effective use of assessment data and results to different stakeholder groups. Hence, it was recommended that the NLSA findings should be timely disseminated to different stakeholders according to their needs.

International assessment

As indicated in ESDP V plan, Ethiopia has a planned to join both Regional and International Assessments. But Ethiopia didn't join any Regional and International Assessment still now. So the analysis of international assessment against the quality indicators of ANLAS was not able to be conducted.

However, during stakeholder consultation, the participants emphasized its importance for monitoring the quality of our education system. Hence, it is recommended that Ethiopia should join SACMEQ from Regional Assessment and either TIMSS or PISA from international assessment to receive additional information about the education system from an international, comparative perspective and to learn from other systems of similar as well as different contexts.

Focus area 3.B: Quality of classroom assessment

Concerning the overall quality of classroom assessment practices, consultations with key stakeholders and document analysis indicated that there are gaps that require immediate actions in which case the seven key areas were partly achieved. Some of the gaps that have been identified include:

- problem of accessibility of the guidelines,
- shortage of resources and tools such as curriculum materials especially syllabi and teachers' guide,
- limited or no use of some assessment methods (example: interviews, questionnaires, peer- assessment, self-assessment),
- lack of 21st century skills as a package,
- weak mechanism of system level monitoring of the quality of classroom assessment and

- limited use of classroom assessment data for school interventions and improvement of students' learning.

Regarding guidelines, there is a problem of accessibility for most secondary and primary school teachers. Hence, the Ministry of Education and Regional Education Bureaus should be able to address the demand of schools by printing and distributing the guidelines. There is also a need for developing new guidelines that incorporate practical examples in each specific domains/ subject for both primary and secondary schools. The guidelines should also incorporate 21st skills.

The capacity building programs for teachers and school leaders are regularly conducted through pre-service and in-service programs. However, there is significant proportion of teachers and school leaders who did not get benefit from these programs. There is a lack of refresher and timely trainings. Thus, it is recommended that all teachers should get the opportunity of trainings through in-service programs. In addition, refresher & timely trainings (Hands- on trainings) on classroom assessment practices should be organized and conducted at national and sub-national levels. Similar trainings are required to be in place at the cluster resource centers for both secondary and primary schools. Awareness creation programs for teachers about the importance of classroom assessment practices should be made.

Concerning provision of resources and tools to schools, shortages of curriculum materials particularly, teacher's guide and syllabi, were confirmed by key informants. Thus, it is recommended that education officials at each level need gathering/ data of all learners in the country, proper planning, timely printing and distribution of curriculum materials to all schools.

Assessment methods and content of assessments are areas that require improvement. As it was confirmed by key informants, only some assessment methods such as quizzes and teacher made tests are commonly used in most schools. However, assessment methods such as interviews, questionnaires, peer-assessment, use of standardized test and self-assessment are not commonly practiced. Regarding assessment content, there is a lack of inclusion of 21st century skills in the curriculum and/or lack of assessment of 21st century skills. To fill this gap, it is advisable to enrich the curriculum with 21st century skills methods such as critical thinking/reasoning, creativity/creative thinking, problem solving, metacognition, collaboration, communication and global citizenship, information literacy, ICT digital literacy and communication and incorporate these skills in the assessment system.

The other key areas that require improvement are monitoring mechanisms and use of assessment data. The study indicated that there is a weak system level mechanism of monitoring the quality of classroom assessment practices. The use of assessment data by teachers and education officials at each level to modify their teaching and make appropriate interventions respectively, is not satisfactory. Hence, it is recommended that strong system level mechanisms of monitoring the quality of classroom assessment should be established. With respect to use of assessment data, the capacity of teachers should be enhanced on how to use assessment data to modify their teaching and improve students' learning. Similarly, the concerned bodies at national and sub-national levels should use classroom assessment data for school interventions. Hence, classroom assessment practices should be used for teacher supervision, teacher performance evaluations, external moderation systems and regular reviews of the quality of education and awarding

government funding for research and development. With respect to use of assessment data, MoE, REBs and education officials at each level should enhance the capacity of teachers on how to use assessment data to modify their teaching and improve students' learning. Similarly, MoE, REBs and education officials at each level should establish systems for classroom assessment data utility of teachers to identify interventions and modifying their teaching practice, school leaders to develop goals and programs for school improvement and to report to the district officials that manage and support schools.

Exhibit 6 lists the evaluation category, aspects of improvement, and recommendations identified for each of the three ANLAS focus areas.

Exhibit 6: Summary of recommendations

| Key areas | Evaluation Category | Aspects for improvement | Recommendations for improvement |
|--|---|---|--|
| Focus area 1: Context of the assessment system | | | |
| <p>1. Legislation or policy The assessment system is guided by legislation or policy.</p> | <p>2. Partly achieved (Improvements are required)</p> | <ul style="list-style-type: none"> • The policy frameworks, Manuals and guidelines that guide the assessment system are not enacted by education law. • There is inconsistency between the policy provision and practices, particularly, there is inconsistency of the criteria set in policy documents that students are required to score a minimum of /50%/achievement level with that of the Norm referenced approach used by the grading system. • Lack of subject specific assessment guidelines at school level | <ul style="list-style-type: none"> • Federal Education law should be developed to strengthen, resolve disputes and protect the assessment system • The Education policy objectives should be revised and reformulated in a way to react the identified gaps • Consistent evaluation and grading system with the criteria set in the education policy must be applied in public examinations. • Up-to-date subject specific assessment guidelines and standards must be addressed to all schools. |

| Key areas | Evaluation Category | Aspects for improvement | Recommendations for improvement |
|---|---|--|--|
| Focus area 1: Context of the assessment system | | | |
| <p>2. Leadership</p> <p>The government demonstrates senior leadership and political will in support of the assessment system. The importance of the assessment system is recognized by all key stakeholder groups.</p> | <p>2. Partly achieved (Improvements are required)</p> | <ul style="list-style-type: none"> • Lack of appropriate management system that can respond to the increasing number of subjects in national examination particularly in Mother Tongue Languages. • Some officials are assigned to areas of technical expertise by political decision and not based on competence which induce unenthusiastic impact to deliver the intended mission. • The Ministry has given little attention to the assessment system especially in funding, disseminating, communicating & utilization of assessment results. | <ul style="list-style-type: none"> • Number of subjects for national examination must be fixed nationally in order to maintain fairness and quality assessment. • Government should establish merit base system of hiring education officials so that they can give technical and professional support in the area. • The Ministry of Education should establish strong monitoring and support system for assessment Programs on funding, disseminating, communicating & utilization of assessment results. |
| <p>3. Funding</p> <p>The government provides sufficient and stable funding for the assessment system.</p> | <p>2. Partly achieved (Improvements are required)</p> | <ul style="list-style-type: none"> • Inadequacy of budget for the design Implementation, analysis, printing and dissemination practices. • Lack of budget for assessment capacity building and infrastructure | <ul style="list-style-type: none"> • The government should allocate Appropriate budget for effective assessment practices, capacity building & technology infrastructure for the assessment system |
| <p>4. Institutional arrangements</p> <p>The government has institutional arrangements in place for designing, implementing, analyzing</p> | <p>1 (Achieved)</p> | | |

| Key areas | Evaluation Category | Aspects for improvement | Recommendations for improvement |
|--|--|---|--|
| Focus area 1: Context of the assessment system | | | |
| and using data from various learning assessments. | | | |
| <p>5. Capacity to use assessment data</p> <p>The government has the capacity to use data from learning assessments for evidence-based decision making in education policy and practice.</p> | 2. Partly achieved (Improvements are required) | <ul style="list-style-type: none"> • There is skill gap in utilization of different software's for analyzing, interpreting of Assessment data. • Limited Capacity of the curriculum development & implementation department in utilization of assessment findings for formative evaluation • Structural gap of EMIS in MoE regarding to documentation and dissemination of data results. | <ul style="list-style-type: none"> • Continuous training on utilization of software applications must be given for, experts in NEAEA. • Intensive awareness and capacity building programs must be given for education officials and/or experts of curriculum department on the importance and utilizing the assessment results for formative evaluation. • The Ministry of education need to develop a system that enhance the capacity of EMIS in a way, information is efficient, effective and accessible to all key stakeholders |
| <p>6. Public engagement</p> <p>The public understands, is engaged in, and supports the assessment system.</p> | 2. Partly achieved (Improvements are required) | <ul style="list-style-type: none"> • Not all key stakeholders have a common understanding on the purpose & importance of National Learning Assessment. • Communication gap in disseminating assessment results. | <ul style="list-style-type: none"> • MoE should give an Intensive orientation and awareness for key stakeholders on the purpose of national assessments. • MoE/NLA department should sufficiently communicate the assessment practices and results via printing and non-printing medias |

| Key areas | Evaluation Category | Aspects for improvement | Recommendations for improvement |
|---|--|---|---|
| Focus area 2: Coherence of the assessment system | | | |
| <p>1. Structure of the education system</p> <p>The assessment system provides performance data in key learning domains and relevant contextual data, at key stages of primary and secondary school education, and for relevant levels of the education system.</p> | 1. (Achieved) | | |
| <p>2. Education policy priorities</p> <p>The assessment system provides relevant data to inform education policy priorities. The assessment programs that form part of the assessment system are aligned to jointly provide the relevant data to inform education policy priorities.</p> | 2. Partly achieved (Improvements are required) | <ul style="list-style-type: none"> 21st century skills of learning domains are indicated in the curriculum framework. However, they are not properly addressed in the assessment systems since both NLA and national/regional examinations are using only multiple-choice test items which do give the room for assessing performance skills. Data from NLSAs and classroom assessments are not well aligned to inform education policy priorities. | <ul style="list-style-type: none"> 21st century skills such as problem-solving skills, critical thinking, communication skills and team work in the learning domains should be more incorporated in the assessment system using a constructed response test items in the future. EMIS should device a mechanism to collect and organize classroom assessment data to align with NLSAs to inform the system for evidence-based decision. |
| <p>3. Learning standards and curriculum</p> <p>The learning domains/</p> | 2. Partly achieved (improvements) | <ul style="list-style-type: none"> Gap in teachers' skills to align classroom assessment with the curriculum standards (competencies) | <ul style="list-style-type: none"> Intensive capacity building should be provided for teachers and school leaders on classroom assessment techniques and its alignment with curriculum standards |

| Key areas | Evaluation Category | Aspects for improvement | Recommendations for improvement |
|--|--|---|---|
| Focus area 2: Coherence of the assessment system | | | |
| subjects in the assessment system are aligned with official learning standards and/or curriculum. | are required) | <ul style="list-style-type: none"> Continuous classroom assessment is not currently implemented as part of teachers' instruction. There are competencies that are not assessed in large scale assessments like listening, speaking, practical activities. | (competencies) <ul style="list-style-type: none"> Instructional materials should be developed in such way that they incorporate more guidance for continuous classroom assessment NEAEA and Curriculum Development and Implementation Department should device a way to assess 21st century skills. |
| 4. School education workforce development Training and capacity-building programs for the school education workforce are aligned with the legislation or policy framework for the assessment system, and official learning standards and/or curriculum that guide the assessment system. | 2. Partly achieved (improvements are required) | <ul style="list-style-type: none"> There is a gap in alignment between training and capacity-building materials with national official documents like curriculum documents. Lack of using assessment data for capacity building programs School principals and teachers do not consider continuous classroom assessment as part of instructional process rather assumed as summative and additional task. This is due to lack of capacity and attitude to assessment Lack of monitoring and evaluation to check the impact of trainings provided on assessment programs | <ul style="list-style-type: none"> Teacher training institutions and MOE should prepare classroom assessments training materials that aligned with the official documents. Capacity building materials should be incorporate the findings of the assessment data. Teachers and educational leaders' development directorate should work on enhancing capacity building and raise awareness of teachers and school leader on assessment programs. MoE should device mechanism for monitoring and evaluating the impact of training on assessment programs. |
| 5. Use of data Assessment data is used by various key stakeholders | 2. Partly achieved (improvements) | <ul style="list-style-type: none"> Lack of timely dissemination of assessment results to various key stakeholders. | <ul style="list-style-type: none"> NEAEA should improve accessibility of data that address the needs of various stakeholders timely. |

| Key areas | Evaluation Category | Aspects for improvement | Recommendations for improvement |
|--|---------------------|---|---|
| Focus area 2: Coherence of the assessment system | | | |
| for evidence-based decision making in education policy and practice. | are required) | <ul style="list-style-type: none"> Lack of accountability for the assessment findings Limited use of assessment data for intervention | <ul style="list-style-type: none"> MoE should establish policies and accompanying rules and regulation for accountability; MoE should provide training/awareness creation on using data for various stakeholders. |

| Key areas | Evaluation Category | Aspects for improvement | Recommendations for improvement |
|---|--|---|--|
| Focus area 3.A: Quality of large-scale assessments and examinations | | | |
| <i>[Program 1 – National Large-Scale Assessment (NLSA)]</i> | | | |
| 1. Organization and resources There is an assessment team with dedicated staff that is appropriately skilled and adequately resourced to complete the diverse tasks associated with key assessment program areas. | 2. Partly achieved (Improvements are required) | <ul style="list-style-type: none"> Currently, there is limited assessment capacity of management, experts and staff (in the area of assessment - related fields like psychometrics and special need) responsible for the NLSA At present there is inadequate physical facilities provision in the NLSA program. | <ul style="list-style-type: none"> Assessment management body, experts and staff in the NLSA program should strengthen their skills by training in the field of assessment and related fields be it short or long term with the help of government and donors. Office rooms, different essential software, scanners should be provided for assessment staff in the NLSA program by the government. |

| Key areas | Evaluation Category | Aspects for improvement | Recommendations for improvement |
|--|---|--|---|
| Focus area 3.A: Quality of large-scale assessments and examinations | | | |
| <p>2. Assessment framework</p> <p>There is documentation that clearly describes the knowledge and skills to be assessed in the key learning domains/subjects, provides a rationale for the contextual data collected and specifies the design of the assessment program. The document is made available to individuals/groups who are involved in assessment instrument development, key stakeholders and the public.</p> | <p>2. Partly achieved (Improvements are required)</p> | <ul style="list-style-type: none"> At present, there is no assessment framework for NLSA program. But there is a policy framework for NLSA program. | <ul style="list-style-type: none"> Assessment framework should be prepared for NLSA programs by NLSA for future use. |
| <p>3. Assessment instruments</p> <p>Quality assurance mechanisms are in place to ensure the test and any contextual data collection instruments are reliable, valid and fair.</p> | <p>1. (Achieved)</p> | | |

| Key areas | Evaluation Category | Aspects for improvement | Recommendations for improvement |
|--|---------------------|-------------------------|---------------------------------|
| Focus area 3.A: Quality of large-scale assessments and examinations | | | |
| <p>4. Sampling</p> <p>The sample, through the use of scientific sampling methods, helps to guarantee appropriate and estimable levels of statistical precision and validity in the interpretation of assessment results.</p> | 1. (Achieved) | | |
| <p>5. Field operations</p> <p>Quality assurance mechanisms are in place to ensure field operations are standardized, monitored and documented, to ensure that the data are collected under the same conditions, independent of the administration context, in an efficient and secure manner.</p> | 1. (Achieved) | | |
| <p>6. Data management</p> <p>Quality assurance mechanisms are in place to</p> | 1. (Achieved) | | |

| Key areas | Evaluation Category | Aspects for improvement | Recommendations for improvement |
|--|--|--|--|
| Focus area 3.A: Quality of large-scale assessments and examinations | | | |
| ensure the final database is free from discrepancies and errors, appropriately structured and documented. | | | |
| <p>7. Data analysis</p> <p>Technically sound and appropriate data analysis techniques are used to provide analytical results that permit valid and useful inferences about the population(s) of interest. Analytical results are fully documented and reproducible.</p> | 1. (Achieved) | | |
| <p>8. Reporting and dissemination</p> <p>Appropriate products and approaches to reporting and dissemination are tailored to the different stakeholder groups and promote appropriate and effective use of the</p> | 2. Partly achieved (Improvements are required) | <ul style="list-style-type: none"> The dissemination strategy is not effective to address the use of NLSA data and results. The NLSA findings could not be disseminated to different stakeholders on time. | <ul style="list-style-type: none"> Dissemination strategy should address the effective use of assessment data and results to different stakeholder groups. The NLSA findings should be timely disseminated to different stakeholders according to their needs. |

| Key areas | Evaluation Category | Aspects for improvement | Recommendations for improvement |
|--|---|--|--|
| Focus area 3.A: Quality of large-scale assessments and examinations | | | |
| assessment data and results by those groups. | | | |
| <i>[Program 2 – National Examination]</i> | | | |
| <p>1. Organization and resources</p> <p>There is an assessment team with dedicated staff that is appropriately skilled and adequately resourced to complete the diverse tasks associated with key assessment program areas.</p> | <p>2. Partly achieved (Improvements are required)</p> | <ul style="list-style-type: none"> Presently, there is limited assessment capacity of management, experts and staff (in the area of assessment - related fields like psychometrics and special need) responsible for the National Examination. At present there is inadequate physical facilities provision in the National Examination program. | <ul style="list-style-type: none"> Assessment management body, experts and staff in the National Examination program should strengthen their skills by training in the field of assessment and related fields be it short or long term with the help of government and donors. Office rooms, different essential software, scanners should be provided for assessment staff in the National Examination by the government. |
| <p>2. Assessment framework</p> <p>There is documentation that clearly describes the knowledge and skills to be assessed in the key learning domains/subjects, provides a rationale for the contextual data collected and specifies the design of the assessment program. The document is made available to individuals/</p> | <p>2. Partly achieved (Improvements are required)</p> | <ul style="list-style-type: none"> Currently, there is no assessment framework for National Examination program. However, there is a policy framework for National Examination program. | <ul style="list-style-type: none"> Assessment framework should be prepared for National Examination program by National Examination program for future use. |

| Key areas | Evaluation Category | Aspects for improvement | Recommendations for improvement |
|--|---|-------------------------|---------------------------------|
| Focus area 3.A: Quality of large-scale assessments and examinations | | | |
| groups who are involved in assessment instrument development, key stakeholder groups and the public. | | | |
| 3. Assessment instruments Quality assurance mechanisms are in place to ensure the test and any contextual data collection instruments are reliable, valid and fair. | 1. Achieved | | |
| 4. Sampling The sample, through the use of scientific sampling methods, helps to guarantee appropriate and estimable levels of statistical precision and validity in the interpretation of assessment results. | As the National Examination followed census approach to collect data, this key area was not applicable. | | |

| Key areas | Evaluation Category | Aspects for improvement | Recommendations for improvement |
|--|---------------------|-------------------------|---------------------------------|
| Focus area 3.A: Quality of large-scale assessments and examinations | | | |
| <p>5. Field operations</p> <p>Quality assurance mechanisms are in place to ensure field operations are standardized, monitored and documented, to ensure that the data are collected under the same conditions, independent of the administration context, in an efficient and secure manner.</p> | 1. (Achieved) | | |
| <p>6. Data management</p> <p>Quality assurance mechanisms are in place to ensure the final database is free from discrepancies and errors, appropriately structured and documented.</p> | 1. (Achieved) | | |
| <p>7. Data analysis</p> <p>Technically sound and appropriate data analysis techniques are used to provide analytical results</p> | 1. (Achieved) | | |

| Key areas | Evaluation Category | Aspects for improvement | Recommendations for improvement |
|--|--|---|--|
| Focus area 3.A: Quality of large-scale assessments and examinations | | | |
| that permit valid and useful inferences about the population(s) of interest. Analytical results are fully documented and reproducible. | | | |
| 8. Reporting and dissemination Appropriate products and approaches to reporting and dissemination are tailored to the different stakeholder groups and promote appropriate and effective use of the assessment data and results by those groups. | 1. (Achieved) | | |
| <i>[Program 3 – Sub-National Examination]</i> | | | |
| 1. Organization and resources There is an assessment team with dedicated staff that is appropriately skilled and adequately resourced to complete the diverse | 2. Partly achieved (Improvements are required) | <ul style="list-style-type: none"> There is currently limited assessment capacity of management, experts and staff (in the area of assessment - related fields like psychometrics and special need) responsible for the Sub-National Examination At present there is inadequate physical facilities | <ul style="list-style-type: none"> Assessment management body, experts and staff in the Sub-National Examination program should strengthen their skills in the field of assessment and related fields by training be it short or long term with the help of national states, federal government and donors. Office rooms, different essential software, scanners should be provided for assessment staff in the Sub- |

| Key areas | Evaluation Category | Aspects for improvement | Recommendations for improvement |
|--|--|---|--|
| Focus area 3.A: Quality of large-scale assessments and examinations | | | |
| tasks associated with key assessment program areas. | | provision in the Sub-National Examination program. | National Examination by the national states. |
| <p>2. Assessment framework</p> <p>There is documentation that clearly describes the knowledge and skills to be assessed in the key learning domains/subjects, provides a rationale for the contextual data collected and specifies the design of the assessment program. The document is made available to individuals/groups who are involved in assessment instrument development, key stakeholder groups and the public.</p> | 2. Partly achieved (Improvements are required) | <ul style="list-style-type: none"> Currently, there is no assessment framework for Sub-National Examination program. | <ul style="list-style-type: none"> Assessment framework should be prepared for Sub-National Examination program by Sub-National Examination program for future use. |
| <p>3. Assessment instruments</p> <p>Quality assurance mechanisms are in place to ensure the test and any contextual data collection</p> | 2. Partly achieved (Improvements are required) | <ul style="list-style-type: none"> At present there is no piloting of test instruments in Sub-National examinations | <ul style="list-style-type: none"> Reviewing and Pre-testing of test instruments ought to be in place in the Sub-National Examination program to ensure the tests are reliable, valid and fair. |

| Key areas | Evaluation Category | Aspects for improvement | Recommendations for improvement |
|--|--|-------------------------|---------------------------------|
| Focus area 3.A: Quality of large-scale assessments and examinations | | | |
| instruments are reliable, valid and fair. | | | |
| <p>4. Sampling</p> <p>The sample, through the use of scientific sampling methods, helps to guarantee appropriate and estimable levels of statistical precision and validity in the interpretation of assessment results.</p> | <p>As the Sub-National Examination program followed census approach to collect data, this key area was not applicable.</p> | | |
| <p>5. Field operations</p> <p>Quality assurance mechanisms are in place to ensure field operations are standardized, monitored and documented, to ensure that the data are collected under the same conditions, independent of the administration context, in an efficient and secure manner.</p> | <p>1. (Achieved)</p> | | |

| Key areas | Evaluation Category | Aspects for improvement | Recommendations for improvement |
|---|--|--|---|
| Focus area 3.A: Quality of large-scale assessments and examinations | | | |
| 6. Data management Quality assurance mechanisms are in place to ensure the final database is free from discrepancies and errors, appropriately structured and documented. | 2. Partly achieved (Improvements are required) | <ul style="list-style-type: none"> Currently, there is no data management manual in the Sub-National Examination program. | <ul style="list-style-type: none"> Data management manual should be developed to ensure the final database is free from error and appropriately documented in the Sub-National Examination program. |
| 7. Data analysis Technically sound and appropriate data analysis techniques are used to provide analytical results that permit valid and useful inferences about the population(s) of interest. Analytical results are fully documented and reproducible. | 2. Partly achieved (Improvements are required) | <ul style="list-style-type: none"> At present there is lack of marking, scoring and data processing machines in the Sub-National Examination program. | <ul style="list-style-type: none"> The Sub-National education bureau ought to provide marking; scoring and data processor machines that help provide valid and useful inferences of the target population in the Sub-National Examination program. |
| 8. Reporting and dissemination Appropriate products and approaches to reporting | 1. (Achieved) | | |

| Key areas | Evaluation Category | Aspects for improvement | Recommendations for improvement |
|--|---------------------|-------------------------|---------------------------------|
| Focus area 3.A: Quality of large-scale assessments and examinations | | | |
| and dissemination are tailored to the different stakeholder groups and promote appropriate and effective use of the assessment data and results by those groups. | | | |

| Key areas | Evaluation Category | Aspects for improvement | Recommendations for improvement |
|---|--|--|---|
| Focus area 3.B: Quality of classroom assessments | | | |
| 1. Guidelines There are national or sub-national level documents that provide guidelines or recommendations for classroom assessment. | 2. Partly achieved (Improvements are required) | <ul style="list-style-type: none"> • Problem of accessibility of the centrally developed classroom assessment guidelines or manuals for schools • Lack of classroom assessment guidelines with practical examples specific to each subject at primary and secondary schools. | <ul style="list-style-type: none"> • The Guideline is required by all secondary & primary schools. Thus, the concerned bodies at the national and sub-national levels should try to address the demand of schools by printing and distributing the guidelines. • There is also a need for developing guidelines that incorporate practical examples in each specific domains /subject at both primary and secondary schools. The guidelines should also incorporate 21st skills as a package. |
| 2. Training | 2. Partly | <ul style="list-style-type: none"> • Lack of refresher and timely training focusing on | <ul style="list-style-type: none"> • Refresher & timely Trainings (Hands- on trainings) on |

| Key areas | Evaluation Category | Aspects for improvement | Recommendations for improvement |
|---|--|---|--|
| Focus area 3.B: Quality of classroom assessments | | | |
| Training programs are provided to (future and current) teachers and school leaders to build their capacity to assess students' learning. | achieved (Improvements are required) | <p>classroom assessment practices.</p> <ul style="list-style-type: none"> lack of training in the importance of classroom assessment | <p>classroom assessment practices should be organized and conducted at national and sub-national levels. Similar trainings are required to be in place at the cluster resource centers for both secondary and primary schools.</p> <ul style="list-style-type: none"> Awareness creation for teachers about the importance of classroom assessment practices <i>should be made</i>. |
| 3. Resources and tools Teachers use a variety of resources and tools for classroom assessment. | 2. Partly achieved (Improvements are required) | <ul style="list-style-type: none"> Shortage of curriculum materials (syllabi, teachers guide, textbooks) | <ul style="list-style-type: none"> Adequate funding is required for printing and distribution of curriculum materials. Education offices at each level need gathering/predicting data of all learners, proper planning, timely printing and distribution of curriculum materials to all schools. Thus, both federal and regional governments should allocate sufficient budget to alleviate this problem. |
| 4. Assessment methods Teachers use multiple assessment methods in order to support valid and reliable assessment of students' learning. | 2. Partly achieved (Improvements are required) | <ul style="list-style-type: none"> There are limited or no usage of multiple methods especially interviews, standardized tests, student self-assessment, questionnaires by teachers in both primary and secondary schools. | <ul style="list-style-type: none"> Teachers should be trained on how to use different assessment methods. To do this, both teacher training institutions and regional Education Bureaus should work together to enhance the capacity of teachers on classroom assessment practices. |
| 5. Assessment content The knowledge and skills assessed are clearly | 2. Partly achieved (Improvements are required) | <ul style="list-style-type: none"> The 21st century skills were not sufficiently incorporated in the curriculum materials The knowledge and skills to be assessed in the Art | <ul style="list-style-type: none"> 21st skills should be incorporated as a package in curriculum. The knowledge and skills to be assessed in the Art |

| Key areas | Evaluation Category | Aspects for improvement | Recommendations for improvement |
|--|---|--|---|
| Focus area 3.B: Quality of classroom assessments | | | |
| defined. | | <p>are not clearly defined in primary schools</p> <ul style="list-style-type: none"> ICT is not part of the school curriculum in primary schools. | <p>should be clearly defined.</p> <ul style="list-style-type: none"> ICT should be part of the school curriculum in primary schools. |
| <p>6. Quality assurance</p> <p>There are formal mechanisms at the school and national/sub-national levels to monitor the quality of classroom assessment practices.</p> | <p>2. Partly achieved (Improvements are required)</p> | <ul style="list-style-type: none"> Monitoring mechanism is not intending on promoting the quality of classroom assessment practices. But rather, it is for the sake of evaluating teacher's performance. Weak system of monitoring the quality of classroom assessment practices | <ul style="list-style-type: none"> Strong system level of monitoring the quality of classroom assessment practices should be established by MOE. The monitoring mechanism should aim at promoting the quality of classroom assessment practices in schools. It should not be for the sake evaluating teacher's performance. |
| <p>7. Use of assessment data</p> <p>Data from classroom assessment is used to improve teaching and learning.</p> | <p>2. Partly achieved (Improvements are required)</p> | <ul style="list-style-type: none"> Lack of awareness of teachers in identifying learning gaps and modifying their teaching practices. Lack of appropriate interventions by using classroom assessment data. | <ul style="list-style-type: none"> Awareness creation for teachers on how to use the assessment data to identify learning gaps and modify their teaching practices The school leaders and local educational authorities should plan and implement appropriate interventions based on the data from classroom assessment. |

4 Using the ANLAS findings

ANLAS is designed to inform the development and implementation of improvement strategies as part of the wider education sector planning process. This section first explains the national education sector planning process in Ethiopia, and then describes how the ANLAS findings are planned/envisaged to be used to inform education sector planning.

As indicated in the current Education and Training Policy (ETP) of 1994, the four medium-term Education Sector Development Programmes (ESDP) had been developed and implemented. The current ESDP V is the fifth medium-term plan which serves as the central strategy document for educational development in Ethiopia from 2015/16 to 2019/20. To monitor and improve the quality education, an attempt has been made to measure education outcomes through robust assessment system since ESDP IV.

The ANLAS model could help as theoretical foundation to adapt and adopt for the Ethiopian education in general and the assessment system in particular. Besides, the ANLAS is used for education planning processes such as Education sector analysis, Education sector planning (ESP), and endorsement of ESP, ESP implementation, ESP monitoring and sector reviews, and Evaluation.

Therefore, this ANLAS diagnostic study will be utilized to develop the next ESDP VI which will have been implemented after 2020. The ANLAS report is also used as a source document for curriculum revision, teacher training institutions, and teacher continuous professional development (CPD). Exhibit 7 provides details of the timeframes for the current and next education sector plan.

Exhibit 7: Education sector planning process

| | |
|--|--------------------------|
| Current Education Sector Plan (ESDP V) | 2015/16 – 2019/20 |
| Next Education Sector Plan (EDSP VI) | 2020/21 – 2024/25 |
| Relevant education sector planning stages to consider ANLAS recommendations | 2020/21 |

Appendix 1: Focal point and national team

Exhibit 8: The National Team

| Name | ANLAS Role | Organization |
|-----------------------------|----------------------|--|
| Mesaye Demessie Zeleke | Focal Point | Deputy Director General, NEAEA |
| Yilikal Wondimeneh Demissie | National Team Member | National Learning Assessment Directorate, NEAEA |
| Arega Mamaru Yewore | National Team Member | National Examination Development and Administration Directorate, NEAEA |
| Nega Gichile Bongasse | National Team Member | National Curriculum Development and Implementation Directorate, MOE |
| Abraham Mengistu Sertse | National Team Member | Planning and Resource Mobilization Directorate, MOE |
| Talefe Eshete Awoke | National Team Member | Teachers and Educational Institutions Leaders Licensing Directorate, MOE |
| Aregawi Gidey Mesele | National Team Member | National Learning Assessment Directorate, NEAEA |
| Bekele Geleta Ayana | National Team Member | National Learning Assessment Directorate, NEAEA |
| Effa Gurmu Bati | National Team Member | National Learning Assessment Directorate, NEAEA |
| Getachew Abebe Tsigie | National Team Member | National Examination Development and Administration Directorate, NEAEA |
| Belay Endeshaw Gizaw | National Team Member | National Examination Development and Administration Directorate, NEAEA |
| Robel Getachew Worku | National Team Member | National Examination Development and Administration Directorate, NEAEA |

Appendix 2: Stakeholder and document mapping tables

Mapping table II: Overview per key stakeholder group

Exhibit 9: Mapping table II: Overview per key stakeholder group

| Key stakeholder group | Focus area 1 Context | Focus area 2 Coherence | Focus area 3.A Quality of large-scale assessments and examinations | Focus area 3.B Quality of classroom assessment |
|--|-------------------------|---------------------------|---|---|
| National or sub-national level officials | | | | |
| • Senior Curriculum Development and Implementation Experts | ✓ | ✓ | ✓ | ✓ |
| • Senior Educational Assessment Experts | ✓ | ✓ | ✓ | ✓ |
| • Senior Examination Development Experts | ✓ | ✓ | ✓ | ✓ |
| • Senior Examination Data Processor Experts | ✓ | ✓ | ✓ | ✓ |
| • Regional and Zonal Educational Experts | ✓ | ✓ | ✓ | ✓ |
| • Senior planning and Resource mobilization Expert | ✓ | ✓ | ✓ | ✓ |
| Training program providers | | | | |
| • Teachers from College of Teachers Education | ✓ | ✓ | ✓ | ✓ |
| School leaders and teachers | | | | |
| • Senior Teachers from primary schools | ✓ | ✓ | ✓ | ✓ |
| • Senior Teachers from secondary schools | ✓ | ✓ | ✓ | ✓ |
| • Primary and secondary school principals | ✓ | ✓ | ✓ | ✓ |
| Development Partners | | | | |
| • World Bank Senior Education Expert | ✓ | ✓ | ✓ | ✓ |
| • DIFD Senior Education Expert | ✓ | ✓ | ✓ | ✓ |
| • UNICEF Senior Education Expert | ✓ | ✓ | ✓ | ✓ |

Mapping table III: Overview per document

Exhibit 10: Mapping table III: Overview per document

| Document | Focus area 1 Context | Focus area 2 Coherence | Focus area 3.A Quality of large-scale assessments and examinations | Focus area 3.B Quality of classroom assessment |
|---|-------------------------|---------------------------|---|---|
| Education and Training Policy, MOE, 1994 | ✓ | ✓ | ✓ | ✓ |
| Curriculum Framework for Ethiopian Education (KG-Grade 12, MOE, 2010) | ✓ | ✓ | ✓ | ✓ |
| Policy Framework for National Large-scale Assessment, NEAE,2012 | ✓ | ✓ | ✓ | ✓ |
| Policy Framework for National Examinations Assessment, NEAE,2012 | ✓ | ✓ | ✓ | ✓ |
| Ethiopia Early Grade Reading Assessment/ English, USAID/RTI, 2010 | ✓ | ✓ | ✓ | ✓ |
| Ethiopia Early Grade Reading Assessment/ Mother Tongue, USAID/AIR, 2010,2014,2018 | ✓ | ✓ | ✓ | ✓ |
| Early Grade Mathematics Assessment: Baseline study report, NEAEA, 2014 | ✓ | ✓ | ✓ | ✓ |
| National Learning Assessment of Grade 4 & 8 Reports, NEAEA,2004,2008,2012, 2016 | ✓ | ✓ | ✓ | ✓ |
| National Learning Assessment of Grade 10 & 12 Reports, NEAEA, 2010, 2014, 2018 | ✓ | ✓ | ✓ | ✓ |
| Education Statistics Annual Abstract, MOE, 2016,2017 | ✓ | ✓ | ✓ | ✓ |
| Contextual data Questionnaires (EGRA, EGMA,NLA), NEAEA, 2014,2017 | ✓ | ✓ | ✓ | ✓ |
| National Examinations students Application Form, 1979-2018 | ✓ | ✓ | ✓ | ✓ |
| Education Sector Development Programme V (ESDP V), MOE, 2015 | ✓ | ✓ | ✓ | ✓ |
| General Education Quality Improvement Package (GEQIP), MOE, 2008 | ✓ | ✓ | ✓ | ✓ |
| General Education Quality Improvement Package for Equity (GEQIP- | ✓ | ✓ | ✓ | ✓ |

Analysis of National Learning Assessment Systems
Ethiopia Country Report

| Document | Focus area 1 Context | Focus area 2 Coherence | Focus area 3.A Quality of large-scale assessments and examinations | Focus area 3.B Quality of classroom assessment |
|---|-------------------------|---------------------------|---|---|
| E), MOE, 2017 | | | | |
| NEAEA Annual Reports, 2018 | ✓ | ✓ | ✓ | ✓ |
| Syllabus for each subject by grade level, MOE, 2008 | ✓ | ✓ | ✓ | ✓ |
| Table of Specifications, NEAEA (n.d) | ✓ | ✓ | ✓ | ✓ |
| Item development Manual, NEAEA, 2011 | ✓ | ✓ | ✓ | ✓ |
| Item Development Evaluation Checklist, NEAEA, 2016,2018 | ✓ | ✓ | ✓ | ✓ |
| National school classification framework, MOE, 2013 | ✓ | ✓ | ✓ | ✓ |
| General Education Inspection Framework, MOE, 2014 | ✓ | ✓ | ✓ | ✓ |
| Continuous Assessment Manual for First cycle Primary (1-4), MoE & UNICEF, (n.d) | ✓ | ✓ | ✓ | ✓ |
| Competency based assessment for learning, UNICEF, 2017 | ✓ | ✓ | ✓ | ✓ |
| Continuous Assessment and how to use it, USAID, 2006 | ✓ | ✓ | ✓ | ✓ |
| Assessment, Measurement and Evaluation in Primary Schools (Course Module), MOE,2014 | ✓ | ✓ | ✓ | ✓ |
| Assessment, Measurement and Evaluation in Primary Schools (Course Module), MOE,2014 | ✓ | ✓ | ✓ | ✓ |
| Continuous Professional Development for Primary and Secondary Teachers, Leaders and Supervisors, MOE, 2010, 2011,2012 | ✓ | ✓ | ✓ | ✓ |
| School Annual Reports,2018 | ✓ | ✓ | ✓ | ✓ |
| Education Sector Development Programme IV (ESDP IV), MOE, 2010 | ✓ | ✓ | ✓ | ✓ |
| Education Road Map (draft), 2018 | ✓ | ✓ | ✓ | ✓ |
| Final READ Trust Fund Report 2008-2015, World Bank, 2015 | ✓ | ✓ | ✓ | ✓ |

References

- Australian Council for Educational Research Centre for Global Education Monitoring, and UNESCO Institute for Statistics. "Principles of Good Practice in Learning Assessment." ACER-GEM and UIS, 2017.
- Clarke, Marguerite. "What Matters Most for Student Assessment Systems: A Framework Paper: Systems Approach for Better Education Results (SABER) Student Assessment Working Paper No. 1." Washington, DC: World Bank, 2012.
- Council of Ministers Regulation No.260/2012. 2012. "National Educational Assessment and Examinations Agency Establishment Council of Ministers Regulation No.260/212." *Federal Negarit Gazeta - No.15*. Addis Ababa: Berhanina Selam Printing Enterprise, January 26.
- Cresswell, John C. "System-Level Assessment and Educational Policy." Melbourne: ACER, 2017.
- Education strategic center (ESC), MoE, 2018-2030, June 2018 " Ethiopian education development roadmap: an integrated executive summery".
- Global Partnership for Education. "Concept Note: Analysis of National Learning Assessment Systems (ANLAS)". Washington, DC: Global Partnership for Education, 2018.
- MOE. 2018. *Education Statistics Annual Abstract*. Addis Ababa: MoE. Annual Educational Statistics Abstract, Addis Ababa: MOE.
- MoE, 2016-2020, Educational Development Sector Plan V (EDSP V), Ethiopian Federal Democratic Government, MoE
- Network for African Learning Assessments (NALA). "Learning Assessments Systems Evaluation Framework (draft).", n.d., http://www.adeanet.org/adeapmp/sites/default/files/activities/learning_assessments_systems_framework_v18.pdf
- The Federal Democratic Republic of Ethiopia National Commission Central Statistics Agency. 2016. *National Strategy for the Development of Statistics (2016-2020)*. Strategy Plan, Addis Ababa: CSA.
- Transition Government of Ethiopia. 1994. *Education and Trainig Policy*. Addis Ababa: MOE.