



United Nations
Educational, Scientific and
Cultural Organization



**INTERNAL
OVERSIGHT
SERVICE**

Evaluation Office

EVALUATION OF THE LATIN AMERICAN LABORATORY FOR THE ASSESSMENT OF QUALITY IN EDUCATION

May 2020
IOS/EVS/PI 192



Office that commissioned the evaluation:

UNESCO Evaluation Office

Authors

Xavier Vanni C.

Juan Pablo Valenzuela B.

Members of the reference group

Silvia Montoya Director UNESCO Institute for Statistics

Margarita Peña Academic Vice-Chancellor of Jorge Tadeo Lozano University - Colombia

Margarita Zorrilla Professor and Researcher at the Department of Education of the Autonomous University of Aguascalientes - Mexico

Ancell Scheker Mendoza Director of Quality Assessment, Ministry of Education of the Dominican Republic and Coordinator of LLECE Dominican Republic

Luis Crouch Senior Economist International Development Group RTI

Jorge Manzi Director, Professor and Researcher of the MIDE UC Measurement Centre, Pontifical Catholic University of Chile

Ernesto Treviño Professor and Researcher, Faculty of Education, Pontifical Catholic University of Chile, and Director of the UC Centre for Educational Transformation

Sobhi Tawil and Maya Prince

Education Research and Foresight team, Education Sector, UNESCO

Evaluation period

September 2019 - May 2020

Evaluation Scope

Latin America and the Caribbean

May 2020

Original report in Spanish

Summary and Acknowledgements

SUMMARY

Over the past 25 years, the Latin American Laboratory for the Assessment of Quality in Education (LLECE) has led the implementation of a system for evaluating student learning that includes most of the Spanish-speaking countries of Latin America and the Caribbean, as well as Brazil. There is positive recognition from participating countries and multilateral and cooperative entities linked to the Laboratory. In spite of this, there are significant challenges relating to the usability and dissemination of results, as well as to achieving a greater impact on the processes of educational improvement and internationalization. It is essential to ensure financial sustainability, moving towards the financing of its operations by the participating countries. The most significant challenge faced in the medium term is the shift from a focus only on assessment management to one that also supports policies and capacities for improving countries' education systems, where the LLECE's linkage to the 2030 Agenda presents an extraordinary opportunity. The evaluation proposes eight recommendations on ensuring financial sustainability; strengthening the coordination team; devising a dissemination strategy; promoting the use of results for research; considering the inclusion of other countries in the region and reviewing periodicity.

ACKNOWLEDGEMENTS

The evaluation of the Latin American Laboratory for the Assessment of Quality in Education (LLECE) was requested by the UNESCO Office in Santiago de Chile (OREALC) and managed by Claudia Ibarguen, Head of the Evaluation Section of UNESCO's Internal Oversight Service (IOS/EVS). The evaluation was carried out by a team comprising Xavier Vanni and Juan Pablo Valenzuela, researchers from the Centre for Advanced Research in Education (CIAE) of the University of Chile, from November 2019 to April 2020. We would like to acknowledge the feedback received from the Reference Group (members listed above) as well as from Sobhi Tawil and Maya Prince from the Education Research and Foresight team of the Education Sector at UNESCO Headquarters. We would like to thank in particular Claudia Uribe, Director of the UNESCO Office in Santiago de Chile, and her team, Carlos Henriquez, Coordinator of the Latin American Laboratory for the Assessment of Quality in Education, and Carlos Cayuman, Education Measurement and Evaluation Specialist, who provided valuable contributions and support throughout the process. The external quality review of the report was carried out by Claudia Marcondes, an independent consultant.

List of Acronyms and Abbreviations

IDB	Inter-American Development Bank	PIRLS	Progress in International Reading Literacy Study
CIAE	Centre for Advanced Research in Education	PISA	Programme for International Student Assessment
CTAN	High-Level Technical Advisory Board	SACMEQ	Southern and Eastern Africa Consortium for Monitoring Educational Quality
CAE	Strategic Advisory Board	SERCE	Second Regional Comparative and Explanatory Study
ERCE	Fourth Regional Comparative and Explanatory Study	TERCE	Third Regional Comparative and Explanatory Study
IOS	Internal Oversight Service	TIMSS	Trends in International Mathematics and Science Study
IEA	International Association for the Evaluation of Educational Achievement	UIS	UNESCO Institute for Statistics
IIEP	UNESCO International Institute for Educational Planning	UNESCO	United Nations Educational, Scientific and Cultural Organization
LAC	Latin America and the Caribbean	UNICEF	United Nations of International Children's Emergency Fund
LLECE	Latin American Laboratory for the Assessment of Quality in Education		
SDG	Sustainable Development Goal		
OECD	Organisation for Economic Co-operation and Development		
UN	United Nations		
OREALC	UNESCO Office in Santiago de Chile; UNESCO Regional Bureau for Education in Latin America and the Caribbean		
PASEC	Programme for the Analysis of Educational Systems of the CONFEMEN (Conference of Ministers of Education of French-Speaking Countries) Countries		
PERCE	First Regional Comparative and Explanatory Study		

Table of contents

List of Acronyms and Abbreviations	4	Sustainability	32
List of Tables and Figures	6	Relevance of LLECE to the 2030 Agenda.....	32
Executive Summary	7	Financial sustainability	34
Recommendations.....	10	V. CONCLUSIONS.....	37
I. INTRODUCTION.....	16	VI. RECOMMENDATIONS	38
II. DESCRIPTION OF LLECE	16	Annexes	43
III. EVALUATION OBJECTIVES AND METHODS	17	A. Terms of reference	43
Context and scope of the evaluation	17	B. References.....	51
Evaluation objectives and questions.....	17		
Evaluation methods.....	18		
Limitations	19		
IV. RESULTS	19		
Relevance	19		
Relevance of the LLECE studies	19		
Progress of the LLECE studies over time	20		
Laboratory organization and management of studies	21		
Effectiveness	22		
Key impacts of LLECE	22		
Capacity building and collaboration	23		
Dissemination and release of results	24		
Efficiency	25		
Adequacy of LLECE's financial resources	25		
Comparability with other international assessment programs	28		

List of Tables and Figures

Table 1. Preliminary estimate of annual revenue for 2016 - 2021 by source of revenue (in thousands of dollars per year).....	26
Table 2. Preliminary estimate of annual expenditure for 2016-2021 by expenditure item (in thousands of dollars per year).....	27
Table 3. Additional complementary studies proposed inERCE.....	28
Table 4. Comparison of the characteristics of ERCE and of other international assessments.....	31
Figure 1. Number of publications mentioned each year in Google Scholar 2005-2018.....	25

Executive Summary

INTRODUCTION

- i. Over the past 25 years, the Latin American Laboratory for the Assessment of Quality in Education (LLECE) has led the implementation of a system for evaluating student learning that includes most of the Spanish-speaking countries of Latin America and the Caribbean, as well as Brazil, where the official language is Portuguese. LLECE had nonetheless not been the subject of a specific evaluation since 2009.
- ii. The general objective of the evaluation which is the subject of this report was to generate evidence and to understand clearly the value that participating in LLECE assessments has had, and still has, for Latin American countries; the objective was also to make proposals to enhance the impact of LLECE in the short and medium term. The evaluation questions focused on four main dimensions: relevance, effectiveness, efficiency and sustainability.
- iii. The evaluation used a mixed methodology, supplementing qualitative and quantitative information collection and analysis methods with the analysis of secondary information, document review, semi-structured interviews, questionnaire analysis and case studies. Semi-structured interviews were conducted with various key players, including LLECE country coordinators, UNESCO officials, coordinators of implementing institutions, and experts.

RESULTS

Relevance

- iv. In almost all the countries participating in LLECE there is high appreciation for the Laboratory's studies and confidence in their results. However, the studies' relevance varies among these countries and is highly dependent on the capacities and level of development of each country's assessment systems, as well as their participation in other international measurements such as the Programme for International Student Assessment (PISA) or the Trends in International Mathematics and Science Study (TIMSS).

- v. The LLECE studies have undergone significant improvement in technical and methodological terms over the last decade. Both the reliability and quality of the tests, as well as the rigour of the processes and analyses have made significant progress, reaching international standards in several aspects.
- vi. Curricular analysis is highly valued because it makes it possible to generate a more contextualized and fair assessment. Moreover, curricular analysis has a value in itself because it allows us to understand how various countries have assimilated the concepts of the 2030 Agenda, as well as to reflect on and initiate processes for updating national curricula.
- vii. The LLECE's tests and analyses of results include considerations of gender, socioeconomic status and the situation of indigenous students, but some countries consider that the approach is still limited. The challenge, therefore, lies in continuing to advance in this area, although this depends heavily on the priorities of each country.
- viii. All those interviewed consider that the management of the Laboratory has improved significantly in the last decade and that its institutional structure has been strengthened. In that regard, the change from a management model based on hiring individual consultants to one of working with specialized institutions as implementing partners is evaluated very positively because it has – 8 – generated more formal, systematic and high-quality work. The creation of the High-Level Technical Advisory Board (CTAN) is also appreciated because it has provided important guidance and brought credibility, validity and rigour to the studies.
- ix. However, most of those interviewed consider that, although the Laboratory team is very willing, committed, efficient and hard-working, it is too small to respond to the growing demands and needs of the countries and to be a counterpart of the implementing partners, especially in the more intense stages of the studies.
- x. The countries have a high appreciation of the participatory management of the Laboratory's studies, an approach which is significantly different from that of other international studies in which they participate. This participatory management approach is especially embodied by the assembly of national coordinators.

Furthermore, the fact that the studies are led by OREALC-UNESCO is highly valued by countries, since the UNESCO Office in Santiago de Chile is an institution that has a high degree of legitimacy in the region; this fact also makes it possible to generate regional dialogue in education and to connect with multiple international bodies.

Effectiveness

- xi. The LLECE studies have contributed to a better diagnosis of the reality of education with regard to learning achievements in the region, with more and better evidence. This has improved discussion of the education agenda and policies in the region, and more firmly established the subject of the quality and equity of learning achievements. In addition, it has increased participating countries' awareness of the strengths and weaknesses of their education systems and enhanced the culture of assessment and use of data for decision-making in education system improvement.
- xii. The findings of the LLECE studies have helped to place new issues and challenges on national education agendas, such as the importance of teaching skills, school attendance and early education, among others. However, with regard to concrete policies, the general view is that the results have had a rather limited and variable impact on education systems in the region.
- xiii. The most significant impact of LLECE has been the strengthening and improvement of the assessment systems of the countries in the region. The assessment framework and the methods for tool development and results analysis, *inter alia*, have been incorporated or adapted into the national assessment systems of many of the countries.
- xiv. Another important result of participating in LLECE has been capacity-building within national technical teams. The Laboratory is considered to be not only an assessment implementation project, but also a training space. In addition, the Laboratory has helped to establish a great many links and significant cooperation between countries on assessment issues.
- xv. The communication and positioning of the Laboratory constitute the area evaluated as the weakest. The general view is that no strategy has been developed to disseminate the results and findings of LLECE, and that no Laboratory narrative or targeted messages for different audiences have been provided. The results reach only the authorities and policymakers, while they are little known by the public, education stakeholders and researchers at the country level.

Efficiency

- xvi. LLECE was established with Inter-American Development Bank (IDB) support as its main source of funding. However, in subsequent cycles, donor contributions were significantly reduced, while UNESCO's contribution was maintained. Given that the countries of the region have improved – 9 – their level of development, ceasing to constitute one of the priority recipient regions for UNESCO's financial support, and given UNESCO's own increased financial difficulties, it has become essential for the countries themselves to bear most of the operational costs of LLECE.
- xvii. The greater complexity and quality of recent studies have increased the costs of the initiative without increasing country contributions. In addition, the lack of adequate budgetary planning and the scarcity of donors, as well as the one-year delay to the planned completion of the current study, have led to an increase in operational costs which, in the case of the Fourth Regional Comparative and Explanatory Study (ERCE; 2016-2021), will amount to a total of \$5.2 million, while the Third Regional Comparative and Explanatory Study (TERCE; 2010-2015) has incurred a cost of \$4.0 million. Notwithstanding the above, analysis indicates that the cost of the cycle per country is moderate when compared to the corresponding cost in other international assessment studies.
- xviii. The financial sustainability of subsequent LLECE studies must be based on the principle of participant self-financing of the baseline operation of the project, in which products common to all countries are defined and complementary products that require financing are defined. Furthermore, donor contributions must continue to be managed proactively with a view to financing new initiatives and development projects, and to supporting those few countries whose financial situation is very complex.
- xix. A comparison of LLECE and other learning assessment programs in which some Latin America and Caribbean Region (LAC) countries participate, such as PISA, shows significant advantages in the Laboratory's provision: its emphasis on the first years of schooling; the assessment of a variety of subjects in two different grades; the assessment of writing skills; assessment based on a shared curriculum; and also a considerably lower cost per cycle than that of other international assessments. It is important, however, to note that the active participation of some LAC countries in other international measurements poses important challenges for LLECE, with regard to clearly identifying its added value, addressing new country requirements

and advances in other measurement programs, and maintaining a competitive cost-to-results ratio.

Sustainability

- xx. LLECE plays a strategic role in the United Nations' 2030 Agenda for Sustainable Development. First, it has specific educational goals to achieve by 2030 and, second, OREALC is responsible for monitoring and supporting compliance in LAC countries. In this context, LLECE assessments will help to define the baseline of multiple indicators with ERCE and to carry out monitoring and assessment activities based on subsequent assessments.
- xxi. Thus, LLECE cooperates with the UNESCO Institute for Statistics (UIS), the Organisation for Economic Co-operation and Development (OECD) and the International Association for the Evaluation of Educational Achievement (IEA) in devising mechanisms to monitor the Sustainable Development Goal (SDG) 4 targets and in defining support and dissemination strategies to progress towards the goals of LAC countries. An example of this is the Rosetta Stone initiative, which will provide a scale for comparison between the IEA and LLECE tests, in order to have comparable indicators for children's educational achievement in LAC and the rest of the world.
- xxii. In order to progress towards achieving the goals of the 2030 Agenda, it is strategically important that LLECE be able to carry out an intermediate assessment (2025) and another assessment towards the end of the decade 2020-2030, in order to gather robust information to analyse educational target achievements and the results achieved by 2030. However, this is an extremely – 10 – complex challenge that requires a significant reduction in the length of each cycle of studies and a stable financing model that provides greater resources.

Conclusions

- xxiii. Four main conclusions may be drawn from the LLECE assessment findings. First, there is positive recognition from participating countries and multilateral and cooperation bodies linked to the Laboratory. Second, despite the above, there are significant challenges in relation to the usability and dissemination of results, and to achieving a greater impact on educational improvement and internationalization processes. Third, it is essential to ensure financial sustainability, thus moving towards

the financing of its baseline operations by the participating countries. Lastly, the most significant medium-term challenge is to move from focusing solely on assessment management to further supporting policies and building capacity to improve national education systems, and the linkage to the 2030 Agenda presents an extraordinary opportunity in that regard.

- xxiv. At the time of completing this report, the countries of the American continent are the worst affected by coronavirus disease (COVID-19) infections and deaths. The ravages of the most critical pandemic in recent decades are extending not only to the health sector, but also to all economic and social sectors, with consequences still difficult to predict, as it is unclear how circumstances will evolve in the coming months. Clearly, the education sector has been severely affected throughout the world. In Latin America and the Caribbean, several countries still have only distance education, which means that learning opportunities, especially for the most vulnerable, are seriously affected. This new context poses significant challenges for LLECE, beyond management and financial aspects, and especially with regard to the dimensions to be evaluated, in its agenda and priorities for the next assessment cycles. The challenge is thus raised this year of identifying how LLECE can become a valuable resource for the continent's education systems in the current context.

RECOMMENDATIONS

Recommendation 1:

Ensure short- and medium-term financial sustainability.

Recommendation 2:

Develop and implement a solid, coherent and systematic strategy for the dissemination of results which also takes into account proposals to improve country-led initiatives.

Recommendation 3:

Define the periodicity of the cycle of studies and its connection to the assessment of the goals of the 2030 Agenda.

Recommendation 4:

Enhance the relevance and use of results for countries.

Recommendation 5:

Promote the use of the results for academic research purposes.

Recommendation 6:

Continue to strengthen the LLECE coordination team.

Recommendation 7:

Facilitate the inclusion of more of the region's countries in LLECE.

Recommendation 8:

Broaden the focus of the Laboratory's work in the medium term, bringing it from the coordination of a regional assessment system to that of a system which also plays a role in capacity-building and educational improvement policies.

Management response

General management response

The evaluation carried out of the Laboratory has been an extremely useful exercise, which helps to review elements of context, progress points, pending tasks, strengths, and needs in order to plan our strategic action and implementation together with our value proposition for the countries. There is an evident alignment between the diagnosis of the evaluation team and several of the problems and opportunities that the members of the Laboratory and the participating countries had already identified and expressed in previous instances.

From the direction of OREALC, the technical team of the Laboratory sees with relative optimism the fact that several of the recommendations are aligned with strategies that have already begun to be implemented: steps towards strengthening financial sustainability, seeking attractive and effective dissemination strategies and the assurance of a good implementation of the studies, to maintain the high technical standards that have been achieved, are all elements that go along this line. On the other hand, the evaluation identifies pending challenges, which should be at the center of the conversations that are foreseen in the short and medium term. Management and support are still needed to ensure a greater flow of resources, strengthening strategic alliances and close relations with collaborators. Work should continue on positioning the Laboratory, both in front of broad audiences and more specialized public, keeping always to provide useful information for educational systems and communities as the first priority. Finally, mechanisms should be sought to ensure the periodicity of the study, within a possible framework of implementation and without sacrificing technical standards of the study.

In the light of these results, the main priorities are:

1. Keep progressing towards the financial sustainability of the Laboratory, by closing the agreements with countries for the term of this assessment but also for the cycles to come, and continue working for the sustainability of projects with the help of key allies.
2. Strengthen and deepen the communication strategies in order to have a real impact on public policy and be useful for the education communities.
3. Strategically look for initiatives that allow the Laboratory to expand its role in order to deliver increasingly pertinent information to the countries, not only when evaluating, but also in order to influence educational processes.
4. Finally, it is necessary to look for support, alliances, projects and sources of financing that allow obtaining resources to strengthen the work carried out for more than 25 years and to publicize the public value, capacity building, information and innovations, which is generated in alliance and participation of 18 countries in the region.

Recommendations

Recommendation 1:

Secure the short and medium term financial sustainability

Responsibility of: UNESCO HQ and UNESCO Santiago Office

Management response

Accepted

With the following remark:

Although the LLECE / OREALC team agrees with the need to ensure the financial sustainability of the Laboratory's work, the responsibility for its financing depends entirely on the member countries that participate in it.

Only when evaluations of lessons learned would be considered a strategic line of work of UNESCO, with the approval and commitment of member countries, this effort will achieve sustainability in the medium and long term. In this sense, its inclusion into the next medium-term strategy would be an important step on this direction.

However, the Laboratory is advancing numerous strategies to achieve financial viability in the short term - until the end of the current cycle the next year,

1. Extension of agreements in force with LLECE member countries until 2021 and increase in fees for the period 2020-2021: The agreements in force with each country currently cover the activities until 2020. Due to the launch of the results of the ERCE 2019 that will be held in mid-2021, we are working with each country on the extension of the agreements until next year, to generate a regular income base that covers the full implementation period of the study until the release of its results. Additionally, this extension of the current agreements includes an increase in the participation fee for the 2020-2021 period. The new participation fee is in line with a more realistic LLECE situation, in which the lack of external donors makes it necessary for member countries to finance the activities and products required for ERCE 2019 to successfully complete its regular cycle.
2. Search for donors and / or projects to prepare the studies and thematic documents associated with ERCE 2019: We are currently searching for institutions that can collaborate technically and financially in the development of thematic studies based on ERCE 2019 data. There are work proposals presenting the background of each thematic study for the institutions that might be interested in collaborating on these activities. Along these lines, we have made progress on the alliance with Unicef that will contribute to the last stage of the project.

In the medium term, the aforementioned actions could be the basis of a new operating model of the LLECE for the next ERCE cycles, in which the participation fees of the countries will make it possible to finance the base operation of the ERCE during a complete regular cycle, while the search for technical / financial partners of the Laboratory will be oriented, mainly, to carry out studies derived from the ERCE or other initiatives that lead to possible innovations and / or improvements.

Recommendation 2:

Design and implement a systematic, robust and articulated dissemination strategy of the results

Responsibility of: UNESCO HQ and UNESCO Santiago Office

Recommendation 3:

Define the periodicity for the cycle of studies performed by ILLECE and its articulation with the evaluation of the goals of the Agenda 2030

Responsibility of: UNESCO HQ and UNESCO Santiago Office

**Accepted
With the following remarks:**

The implementation of this recommendation is subject to the effectiveness of the financing strategy proposed in the previous point. However, important efforts are being made in this regard: With the support of the LLECE National Coordinators, the dissemination strategy will be worked on during the first semester of 2021 and will include as many elements identified and proposed in this Evaluation as possible.

For example, in line with the recommendation to «actively use social networks and improve the website», a new website of the Laboratory has been launched, which so far contains the main results of the TERCE and the Curricular Analysis of the ERCE 2019. Progress will also be made in a systematic dissemination plan, which is not limited to the delivery of final results but also of the results of the curriculum analysis, the first stage of the ERCE studies, as a prelude to the delivery of results, something that has not been done for previous studies. As a preparatory experience for the development of a more solid and comprehensive communication base, during the last months a strategy was designed, led by the LLECE coordination and implemented together with the participating countries, for the launch of the Curriculum Analysis carried out at the end of July. Despite the current pandemic context, this strategy made concrete and visible progress.

**Accepted
With the following remarks:**

The definition of contributions from the organization and the participating countries is essential for financial viability. Although this is one of the main elements for the future planning of LLECE, the current status, the financial model and implementation pace of the cycle by the countries make it difficult to set dates with certainty in the short term. It would be necessary first, to modify the financing model of the Laboratory to achieve long-term economic support, and second, to see the technical and economic feasibility of the countries for its implementation.

Recommendation 4:

Increase the relevance and use of the results by the countries

Responsibility of: UNESCO HQ and UNESCO Santiago Office

Recommendation 5:

Potentiate the use of the results on research by the academic sector

Responsibility of: UNESCO HQ and UNESCO Santiago Office

Recommendation 6

Continue strengthening the coordination team of LLECE

Responsibility of: UNESCO HQ and UNESCO Santiago Office

Accepted

The following measures are considered for the short term:

1. National reports of the Curriculum Analysis of ERCE has been developed for the first time.
2. It is expected to produce national reports for the delivery of ERCE 2019 results with all the countries of the Laboratory, which will be available at the regional launch of the study results.

The results of the ERCE will be disseminated through the new website and social networks, with a focus on visualizing the results and understanding all audiences of the results.

Accepted

Some of the recommendations proposed in the Evaluation are beginning to be implemented, for example the case with the launch of the ERCE Curricular Analysis 2019.

Currently, documents are being translated into English. Now the website has a section that contains the main results of the TERCE and the Curricular Analysis of the ERCE 2019. All the documents produced by the Laboratory will also be made available on this website for easy access.

Depending on the availability of funds, the option of holding seminars or workshops with academics and researchers that are oriented to the production of results from the ERCE 2019 data will be evaluated.

Not accepted

It is not feasible given the current LLECE budget. The coordination team manifested its concerns several times regarding the imbalance in the workload.

The priority is to achieve financial sustainability to successfully complete the regular cycle of the ERCE 2019, which includes ensuring the job stability of the current team.

Subsequently, as the financial model of the Laboratory allows, the possibility of strengthening the coordination team will be evaluated.

Recommendation 7

Open the possibility of including other countries of the region to LLECE

Responsibility of: UNESCO HQ and UNESCO Santiago Office

Recommendation 8

In the medium term, extend the work focus of the Laboratory from the coordination of an evaluation system to including a role on capacity building policies and improvement of education standards.

Responsibility of: UNESCO HQ and UNESCO Santiago Office

Not accepted

It is not feasible given the current LLECE budget and the requirements necessary for its implementation.

At this time, the study reaches almost all the Spanish-speaking countries of the Region, with the sole exception of Venezuela, in addition to the participation of Brazil. Brazil requires specific adaptations of the ERCE development process due to the language difference. The support of the UNESCO Office in Brasilia and the opportunity to have people in the Brazilian national team who can understand Spanish, has allowed the participation of Brazil. The inclusion of more countries, although it is one of the interests indicated, implies analyzing the following elements:

1. Economic feasibility of working with countries with another language. The implementation of the evaluation in another language brings with it increased costs in the work with implementing partners, coordination, translation, meetings, among others.
2. Adaptation of instruments to other languages
3. Work with national teams
4. Technical feasibility of working with countries with small populations.

Due to the above reasons, the inclusion of more countries from the Region is difficult to contemplate in the current scenario and will be subject to the improvement of the state and financial model of the Laboratory. The Roseta Stone project in alliance with the UIS, mentioned in the evaluation, and which seeks to connect the ERCE results with those of other regional learning evaluations, takes a step in the direction towards a greater opening of the Laboratory to other countries.

Accepted**With the following remarks:**

The results of the LLECE feed the work developed at the other units of OREALC. The Laboratory is part of an organization in which each unit carries out its own tasks but ensuring synergy and complementary such that as a whole, it aims to advance school improvement and reach the goals of the 2030 Agenda.

The current context with the pandemic, opens an opportunity to investigate an evaluation system with new components such as the intermediate towards a diagnostic and progress evaluation, and the formative one.

As a preliminary step, the efforts will focus on getting the support of the UIS to search for possible donors and to make synergies in the establishment of this platform in more countries. In addition, during the pandemic, the Laboratory's Technical Coordination has worked to prepare and publish various short thematic documents with a focus on recommendations for educational systems and communities.

I. INTRODUCTION

II. DESCRIPTION OF LLECE

1. The Latin American Laboratory for the Assessment of Quality in Education (LLECE) of the UNESCO Regional Bureau for Education in Latin America and the Caribbean (OREALC) is the unit responsible for monitoring the progress of pupils in the region through large-scale assessment studies. Since its establishment in 1994, the Laboratory has provided a regional reference and framework for consultation and cooperation among countries in the field of educational assessment, technical support for the training and capacity-building of teams responsible for national measurement and assessment systems, and a source of access to information available to countries to promote the development of empirical-evidence-based education policies. The Laboratory emerged at a time in which many countries in the region were introducing education reform without sufficient and relevant information for its development and lacked a critical mass of resources to assess the quality of education. At that time, only a few countries in the region were carrying out large-scale learning measurements.
2. LLECE emerged at a time when developing countries in various regions started to coordinate efforts to implement transnational assessment systems, thus reflecting challenges that extend beyond Latin America and the Caribbean. The Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) is the organization most similar to LLECE, which has implemented four assessment cycles since 1995, with the participation of 15 English-speaking African countries in its latest cycle. Nevertheless, unlike LLECE, SACMEQ assesses only sixth-grade pupils and only in reading and mathematics. Moreover, its funding relies on substantial contributions stemming from international cooperation, especially from the Netherlands. Additional initiatives include the Programme for the Analysis of Educational Systems of the CONFEMEN Countries (PASEC), which involves the participation of Francophone African countries. However, it coordinated independent national assessments between 1991 and 2012, and changed its strategy to make data comparable among the various participating countries only since 2014 (UNESCO, 2019). This programme assesses the performance of second- and sixth-grade pupils in reading and mathematics. A recent initiative is the SEA-PLM (Southeast Asia Primary Learning Metrics) programme, which, for the first time in 2019, is coordinating the assessment of 11 Southeast Asian countries.
3. For more than 25 years, LLECE has worked in cooperation with most Latin American and Caribbean countries on comparative studies that measure the learning achievement of primary school pupils in the third and sixth grades. These are devised and implemented with the active participation of the countries in the region through an assembly of national coordinators from the participating countries. The UNESCO Office in Santiago de Chile has a small technical team that coordinates the studies and engages specialized implementing partners and technical institutions that carry out the key processes of the studies.
4. Since 2011, LLECE has also had the support of the High-Level Technical Advisory Board (CTAN), composed of four internationally recognized experts with expertise in standardized assessments and four countries representing the assembly within the body. CTAN provides expert advice, strategic guidance and recommendations to inform technical decisions on issues such as sampling, test design, item analysis, secondary analysis of associated factors and report preparation. Recently, the Laboratory also established the Strategic Advisory Board (CAE), composed of a group of education experts and other strategic project partners. The purpose of the Board is to advise on the Laboratory's strategic direction and its financial sustainability in the medium term.
5. The First Regional Comparative and Explanatory Study (PERCE, 1997) involved 13 countries and assessed reading and mathematics in the third and fourth grades. After PERCE, two new complete cycles of the study were undertaken: the Second Regional Comparative and Explanatory Study (SERCE, 2006) and the Third Regional Comparative and Explanatory Study (TERCE, 2013), which began to assess the third and sixth grades, including sixth-grade science and writing. In addition, countries can add a national module, in which they may include items related to particular elements of their curricula, either in academic areas or other areas of development. The tests are based on each country's national learning objectives, so that each edition of the study begins with an exhaustive analysis of the school curricula in all participating countries.

6. The Fourth Regional Comparative and Explanatory Study (ERCE, 2019) is currently under way and its results will be issued in 2021. It involves 18 countries¹ and assesses the performance of third- and sixth-grade pupils in mathematics, reading and writing, and of sixth-grade pupils exclusively in natural sciences. Furthermore, a set of questionnaires is used for pupils, teachers, families and school principals; they seek to explain pupil learning outcomes through context variables and associated factors. This version of the study also includes a new module for measuring socio-emotional skills in sixth-grade pupils; it is focused specifically on awareness and appreciation of the other, self-regulation and self-management.

¹Argentina, Bolivia (Plurinational State of), Brazil, Chile, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru and Uruguay

III. EVALUATION OBJECTIVES AND METHODS

Context and scope of the evaluation

7. Over the past 25 years, LLECE has led the implementation of a system for measuring pupil learning that involves most Latin American and Caribbean countries. However, the Laboratory has not undergone an individual evaluation since 2009, when the University of Chile was responsible for this task (University of Chile, 2009). As a result, this evaluation will provide a review of its contribution and role in the area of quality education and policies in Latin America in the past decade and will highlight future challenges and opportunities. The evaluation will contribute to strategic decisions and inform future action for the implementation of the Laboratory's work.
8. The evaluation will be both retrospective – providing an overview of the Laboratory's performance over the past 10 years – and prospective – giving rise to recommendations on ways in which the Laboratory may better achieve its future objectives. Some questions relating to the evaluation of the Laboratory from the perspective of the participating countries may take longer to explore than the time frame of this evaluation. The focus of the evaluation is neither the effectiveness of the implementation of the assessments in each country nor a technical evaluation of the quality of LLECE methods or measurement tools.

Evaluation objectives and questions

9. The overall objective of the evaluation was to generate evidence and provide an in-depth understanding of the value – past and present – which participating in LLECE assessments holds for Latin American countries, and to make proposals to enhance the impact of LLECE in the short and medium term.

10. In addition, the evaluation was aimed at identifying the contributions of LLECE to: (i) providing quality evidence on student learning and its explanatory factors; (ii) generating information to guide educational policy decisions at the national level among participating countries; (iii) promoting regional dialogue on curricular aspects of the subjects assessed and on the quality of education; and (iv) building professional and technical capacity in educational assessment among participating countries.
11. In addition, the evaluation will: (i) evaluate the activity of OREALC's LLECE over the past decade, analysing its progress and identifying its strengths and weaknesses; (ii) analyse the consistency between the activity of LLECE and UNESCO's goals and mission, exploring synergies with various UNESCO programmes; (iii) evaluate the Laboratory's contribution to the 2030 Agenda, specifically SDG 4; (iv) examine the funding model and mechanisms, both within UNESCO and the participating countries, taking into account their sustainability and future projections.
12. The evaluation questions focused on four main dimensions: relevance, effectiveness, efficiency and sustainability.²

Evaluation methods

13. The evaluation involved multiple methods, combining qualitative and quantitative information collection and analysis methods through secondary information analysis, document and budget review, semi-structured interviews, questionnaire analysis and case studies.
14. With regard to document review, an analysis was made of: the reports of the previous two studies evaluating the Laboratory's work (2008 and 2009); the reports of TERCE and SERCE; the minutes of the meetings of national coordinators of the participating countries, of the High-Level Advisory Technical Board (CTAN) and of the Strategic Advisory Board (CAE); internal Laboratory documents; and the budgetary background of the Laboratory.
15. Semi-structured interviews were held with various key stakeholders participating at different levels in the development of the Laboratory, including: the Director of OREALC; the LLECE/OREALC Coordinator and Technical Team; the national coordinators of all the participating countries, except those of the Plurinational State of Bolivia and Honduras, with whom it proved impossible to conduct the interview; the Head of

² The details of the questions are contained in Annex 1.

the Education Research and Foresight program, UNESCO; the Director of the UNESCO Institute for Statistics (UIS); a member of the Laboratory's Strategic Advisory Board (CAE); the coordinators of the Measurement Centre of the Pontifical Catholic University of Chile, a TERCE and ERCE implementing partner; the former Director of the Centre for Comparative Education Policy at Diego Portales University, which was a TERCE implementing partner; a researcher at the Centre for Advanced Research in Education (CIAE) who worked as an implementing partner for the curriculum component of ERCE; and a PISA expert from the OECD.

16. Furthermore, some indicators were developed to compare results and processes with those of other large-scale learning measurement systems, such as the OECD's PISA3 assessment and IEA's TIMSS and Progress in International Reading Literacy Study (PIRLS) assessments.⁴ An analysis of academic productivity was also carried out based on the results of PERCE, SERCE and TERCE, using academic search engines.
17. In addition, two in-depth case studies were conducted, which complemented the more cross-sectional findings of the study. The countries selected – with their agreement – were Ecuador and the Dominican Republic, in view of their active participation and the relevance of LLECE in these cases. Both countries were visited by one of the study consultants, who interviewed various key stakeholders in the school system, including the assessment director, the LLECE Coordinator, Ministry of Education authorities, researchers and academics, teachers' union leaders, and representatives of education sector NGOs.
18. A short online questionnaire was initially planned for the national coordinators, but it was discarded given that, shortly before the start of the evaluation, the Laboratory coordination team had sent the national coordinators a questionnaire with open-ended questions on the implementation of ERCE and its strengths, impacts and challenges. Accordingly, it was considered inadvisable and unnecessary to send another questionnaire in such a short time. The replies to the questionnaire sent by the Laboratory were also considered in the analysis for this report.

³ PISA is a study undertaken by the OECD that assesses how education systems prepare their students to apply knowledge and skills to tasks relevant to their current and future lives. This study is conducted every three years among 15-year-old students. The assessment is administered through computer- or paper-based tests that assess knowledge and skills with regard to reading, mathematics and natural sciences.

⁴ TIMSS is an international assessment of the learning achievements of fourth- and eighth-grade learners of mathematics and science, carried out every four years. These tests are conducted by IEA, which also conducts the PIRLS assessment, targeting fourth-grade pupils.

Limitations

19. The greatest limitation for the implementation of the study lay in time constraints and the time of year, which made it difficult to schedule several of the interviews, which meant that a few could not be conducted. Moreover, the rotation of national coordinators in some countries led to difficulties in reviewing the direction taken and changes initiated by LLECE studies, as well as their impact on national education systems and policies.

IV. RESULTS

Relevance

Relevance of the LLECE studies

20. In almost all the countries participating in LLECE, there is a high level of appreciation for the Laboratory's studies and confidence in their results. However, the studies' relevance varies among countries; relevance is highly dependent on the capacity and level of development of the assessment systems of each country, as well as on participation in other international measurements such as PISA or TIMSS. Interviewees were of the view that one of the strengths of the LLECE studies is that they allow for longitudinal monitoring of student learning grounded in the regional reality. Moreover, LLECE involves several countries that do not participate in other international studies⁵; LLECE also provides the only international assessment system for primary education in a number of countries.⁶ In addition, it is very useful and relevant to the 2030 Agenda and the evaluation regarding the goals established therein. Some of the experts interviewed consider the LLECE system to be the most developed regional assessment system in the world, compared to other systems such as SACMEQ and PASEC, and also value the fact it has continued uninterrupted over the past three decades.
21. Based on the interviews and materials reviewed, countries value at least six characteristics of the LLECE studies.
- (i) The assessments are carried out at the primary school level, namely, in the early stages of schooling, where education policies can have a greater impact. This presents a significant advantage compared to other international measurements such as PISA, which assesses learning outcomes in 15-year-olds.
 - (ii) The progress of the outcomes in each country are made available and can be compared with that of countries in the region with similar characteristics.
 - (iii) In contrast to international assessments such as PIRLS and TIMSS, which operate in a more globalized context and do not take into account the specificities of LAC, the LLECE tests have a common regional curriculum base and thus allow for a more accurate, context-specific assessment.
 - (iv) The form of governance is participatory and decisions are taken collectively by the participating countries, which generates a sense of ownership and commitment to the studies.

⁵ Bolivia (Plurinational State of), Cuba, El Salvador, Nicaragua

⁶ Argentina, Bolivia (Plurinational State of), Brazil, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru and Uruguay

(v) The innovations achieved through TERCE, and especially ERCE, have facilitated the incorporation of many of the advances in other assessments and secured the international recognition of LLECE by the other assessment bodies.

(vi) The fact that assessment is led by OREALC is highly valued by the participating countries, since this institution has a high degree of legitimacy in the region while also promoting regional dialogue on education and liaising with a variety of eminent international multilateral bodies renowned for their technical expertise. For all these and other reasons, almost all the national coordinators interviewed explicitly expressed their countries' interest in continuing to participate in the LLECE studies.

Progress of the LLECE studies over time

22. All the stakeholders interviewed, both national coordinators and experts, pointed out that the techniques and methods used in LLECE studies have improved significantly over the past decade. Both the reliability and quality of the tests and the rigour of the processes and analysis have progressed significantly, thus meeting international standards in several respects. Among other points, interviewees highlighted the incorporation of updated methods, both in the development of the tests and the analysis of results and associated factors, such as the item response theory (IRT) model; an increase in the number of assessed components, such as the writing test (since SERCE) and the socio-emotional skills module incorporated into ERCE; the inclusion of open-ended questions and the use of items to assess them; the development of protocols for the different processes; and improved comparability of the progress of outcomes.
23. In contrast, interviewees also highlighted some technical aspects and processes that would need to be reviewed or improved. Several national coordinators and experts argued that sampling is unnecessarily extensive and therefore very costly, owing to the stratification criteria of the sample used, thereby suggesting that the criterion of school size should be reviewed. This is more noticeable among countries participating in other international measurements, in which smaller sample sizes are examined. Moreover, there are a number of discrepancies between some national coordinators and the implementing partner regarding the information capacity of the tests. Four coordinators noted that the tests would make little distinction at the lower end of performance distribution – which, for countries with a very high proportion of students at the lower end of the distribution, is of little informative value and use for

understanding results. The implementing partner, however, considers that the tests provide adequate information capacity in the low-score region. Thus, the experience of the PISA for Development (PISA-D) test is very valuable, since it was aimed at expanding the coverage of the PISA assessment in nine developing countries (including five from LAC), where special emphasis was placed on covering a wider range of performance levels at the lower end of the student achievement distribution.

24. Some national coordinators suggested that there was little time between the pilot test and the final test, which did not allow for a detailed review of the results in order for countries to make comments and suggestions. Additionally, they suggested the need to continue improving the manuals and protocols developed, to continue standardizing the different processes and to have greater clarity on how each procedure should be carried out. Furthermore, almost all coordinators stated that while the platform and software had improved compared to those used in TERCE, they still posed difficulties for use by national team professionals.
25. The curriculum analysis is highly valued by almost all the stakeholders interviewed, especially the national coordinators. Moreover, many countries consider it to be one of the great strengths of the study because it facilitates a more accurate, context-specific assessment. It also has the potential to identify the weakest learning aspects in each country, thus facilitating the development of public policy priorities to improve educational opportunities for primary school children. Almost all the national coordinators are of the view that this shared curriculum framework would be very beneficial for the legitimacy of LLECE studies. They consider that without this common framework it would be very difficult to agree on what the focus of the assessments and expected standards should be and, in this connection, other international tests are completely exogenous to the reality of the LAC countries. Moreover, curriculum analysis is valuable in itself, because comparing national curricula helps to understand how different countries have integrated into their curricula the concepts of the 2030 Agenda and the remaining challenges, and also to reflect on and initiate processes for updating national curricula with enhanced evidence.
26. Nevertheless, most of the experts and a couple of national coordinators argued that while curriculum analysis is important, it should focus on key common skills so that an in-depth analysis is not necessary for each cycle of studies. Moreover, about half of the national coordinators, as well as the experts interviewed, were critical of some of the criteria and procedures used to carry out the curriculum analysis. Key considerations included: (i) that only the content and skills covered in the grade in which the tests

are carried out should be considered – excluding the content and skills covered in all previous grades – failing which, some bias would be introduced into the assessment, since different content is delivered in different grades in the participating countries; and (ii) that the way to determine what is ultimately considered common content, to be assessed in the tests, is not sufficiently clear to countries, and that therefore, communication in that regard needs to be improved.

27. Some of those interviewed argued that some problems arise from the fact that the tests have a common curriculum, namely, that the introduction of curriculum changes at the national level weakens the basis for making comparisons over extended periods and that the curriculum analysis process extends the duration of each study cycle. Lastly, the great majority of those interviewed consider that the enormous effort that is made with curricular analysis is wasted, while at the same time countries receive very little feedback in this aspect, especially for making public policy decisions.
28. It was decided that the new innovation module for the measurement of socio-emotional skills would be applicable only to sixth-grade pupils, after the tool was piloted in 2018. The aspects assessed are awareness and appreciation of the other (empathy, acceptance of schoolmates belonging to different social groups), self-regulation and self-management (self-perception regarding behaviour and effort, and impulsiveness and tolerance to frustration). This innovation is highly valued, since it provides for a more comprehensive assessment and highlights the importance of developing socio-emotional skills, the factors that influence them and the relationship with the outcomes of the learning achievement tests. However, several national coordinators consider that in view of its recent incorporation into ERCE, the conceptual framework for this assessment cannot be sufficiently discussed and further comments and suggestions cannot be made. Accordingly, they consider that if this module for the measurement of socio-emotional skills is effective, it should be incorporated definitively into LLECE, but needs to be discussed at greater length and in greater depth.
29. Lastly, the consideration of specific and/or vulnerable groups and populations was generally given little emphasis in the interviews. However, the coordinators of four countries noted that the consideration of vulnerable groups and indigenous peoples was still limited, owing to the limited translation and linguistic adaptation of the tests. In addition, gender, socioeconomic level and indigenous peoples are considered in the analysis of both results and associated factors. The general view is that it is important to make further progress in this area, but that it depends heavily on the ability and willingness of each country to invest more resources in it, in order, for example, to have

an oversample of some populations and/or to address some of these groups through the national module.

Laboratory organization and management of studies

30. The coordinators and experts are of the view that compliance with timetables and deadlines has improved in ERCE compared to previous studies. However, half of them stated that significant improvements were still required and that the LLECE studies did not provide a detailed and complete schedule of all the stages and processes from the beginning of each cycle. In the interviews, the coordinators also referred to the timetable for the northern region, where the tests are administered first, noting that the schedules for the development and revision of tools are very tight and that, consequently, it would be necessary to consider a change in this regard.
31. All those interviewed consider that the management of the Laboratory has improved significantly in the past decade and that its institutional structure has been strengthened. In that regard, the shift from a management model based on hiring individual consultants to a model of cooperation with specialized institutions as implementing partners is considered very positive by almost all countries and experts, which has facilitated more formal, systematic and high quality work. While the capacity and work of the implementing partners is evaluated as positive, there is general criticism about their high level of concentration in Chile and in few institutions. It is considered very important to further diversify the implementing partners and to review the obstacles or elements that might hinder other institutions in the region from participating. Some experts and representatives suggested that this lesser diversity could be partly explained by the fact that, in several countries, the greater technical capacity linked to measurement processes is located in government bodies rather than in universities or private institutions, and also by the fact that the financial resources available would not be appealing enough to attract a number of other institutions.
32. One management problem raised by a large group of national coordinators concerns the possibility to respond to requests or queries from countries with regard to the application, processing and analysis of results, which leads to difficulties such as decision-making in the absence of the necessary information and to delays that result in very tight deadlines and increased costs. A concrete example was the correction of open-ended questions, a new and complex issue for many of the countries. The general view is that, while the Laboratory team is very willing, committed, efficient and hard-

working, it is too small (only three support staff members, two assistance staff members and the Coordinator) to meet the growing demands and needs of the participating countries and to be a counterpart to the implementing partners, especially in the most intensive stages of the project. During the evaluation interviews, the need to increase the size, technical capacity and stability of the team was identified as a challenge.

33. Many of those interviewed highly value the assembly of national coordinators. This is considered to be a distinctive feature of LLECE, compared to other institutions conducting international studies in which countries consider that they do not have much say in decisions and a very vertical relationship is established between the institution conducting the study and each country, with no overarching discussion bodies. A greater sense of involvement and influence is important, as it further promotes the validation and use of the study results and findings. Nevertheless, some of the experts interviewed argued that one disadvantage of this governance structure is slower decision-making as compared to that related to other international studies.
34. The High-Level Technical Advisory Board (CTAN), which was established during TERCE but expanded its functions and activities significantly during ERCE, is highly valued by all stakeholders. The perception is that CTAN has contributed valuable guidance and observations to the various stages and processes of the study, and that it has provided credibility, validity and rigour to the decisions made and to the studies in general. In addition, the work of CTAN has helped to generate higher-level technical dialogue with the Laboratory team, with the implementing partners, and in the discussions and decisions of the assembly of national coordinators.
35. While the establishment of the new Strategic Advisory Board (CAE) – composed of ad honorem representatives of international bodies with a strong commitment to the present and future development of LLECE – is still considered positive, many national coordinators consider its function and relevance unclear and find it difficult to identify the expected outcomes providing a basis for its evaluation. Furthermore, both the national coordinators and some experts consider that the OREALC authorities should be more involved in LLECE. It seems that they have not been so present historically and that improving the functioning and impact of the results and findings would require a more political and systematic dialogue on LLECE with the ministry of education authorities in each country, rather than only at the national coordinator level.

Effectiveness

Key impacts of LLECE

36. Overall, there is some consensus that the LLECE studies have improved the diagnosis of the educational reality of the region and of each participating country, especially those with a lower level of development. This enhanced evidence has helped to improve the discussion of the education agenda and policies in the region, and to strengthen the impact of the theme of quality and equity of education, with a focus on learning achievements. In addition, it has generated greater awareness in the countries of the strengths and weaknesses of their education systems, as well as informed the design of some national policies and the recommendations of OREALC/UNESCO. In this sense, a recurrent idea that emerged in many of the interviews is that the LLECE studies have been highly beneficial in promoting an enhanced culture of assessment and use of data among education system stakeholders in the region.
37. However, their impact varies considerably among countries and depends largely on the importance attached by each country to the LLECE studies. In about one third of the countries, the results of LLECE studies have been used to define the objectives of their national education plans or medium- and long-term commitments, and/or to report on the results and quality of their school systems. For example, in the Dominican Republic, the indicators of the National Development Strategy Act (long-term political agreement in which education is one of the key areas) are based on LLECE study results. In El Salvador and Ecuador, a similar observation is made with regard to the national education plan objectives and, in the case of Panama, the LLECE study results have provided an important input to the development of the National Commitment to Education.
38. The findings of the LLECE studies have contributed to placing new issues and challenges on national education agendas, including the importance of teaching skills, school attendance and early education, the consequences of repetition, and problems related to climate and violence in many schools. Nevertheless, with regard to concrete policies, many of the interviewees are of the view that the results have had a rather limited and variable impact among countries. In the Dominican Republic, for instance, they have been used as a basis for some policies such as increased school time and early education provision, and policies addressing climate problems at school. In the case of Nicaragua, the results of TERCE provided the basis for teacher-training policies

and the school retention and attendance programme. However, in other countries such as Brazil, Peru, Colombia, and Chile, the results have had little effect and no major impact is observed in terms of concrete education policies. This may be due to, among other reasons, the systematic conduct of national tests and the level of development of their national measurement systems, and to their participation in other international measurements with a greater public impact.

39. In some countries, LLECE studies have influenced or served as an input for curriculum changes, ranging from changes or adjustments in content (Nicaragua) and curriculum structure (Uruguay) to the incorporation of new approaches to skills development, as in the case of the Dominican Republic.
40. One of the difficulties in increasing the impact of the information and findings arising from the studies lies in a significant gap or weak link often observed between the units implementing the assessments and the national authorities and ministry of education policymakers, who devise and implement national policies. This difficulty is also observed in relation to other education stakeholders in the countries, with whom the distance is generally even greater. Another difficulty is that, in general, national assessment bodies have little capacity to conduct an in-depth analysis of results and even less capacity to translate findings into policy proposals. In fact, several countries were critical of the LLECE global performance reports, arguing that they provide an insufficient understanding of results, and that more tailored reporting at the national level is required, to identify more effectively the national priorities and policies that need to be improved. Additionally, several interviewees suggested that the analysis of associated factors would have been very focused on structural and demographic factors, which are certainly considered relevant, but that insufficient emphasis has been placed on teaching and leadership practices – although teaching practices will be included for the first time in the ERCE analyses.
41. The most significant impact of LLECE has been to strengthen and improve the assessment systems of the countries in the region. The assessment framework, the methods for tool development and results analysis, the dimensions and grades assessed, the use of open-ended questions, and the method of application, inter alia, have been incorporated or adapted into the national assessment systems of many of the countries. For example, Costa Rica and Peru are implementing or developing national writing tests based on the LLECE experience. As for Brazil, the Laboratory experience inspired it to develop a science test, and it is currently working on the development of a writing test. In the Dominican Republic, national census tests began to be introduced in third

and sixth grade, and in Nicaragua – which did not have a national assessment system – national pilot tests are being conducted in language, mathematics and science. For many countries, LLECE has also provided an important benchmark for verifying the validity and reliability of their national tools. As one of the interviewees brought to our attention, many country assessment units and departments have grown and/or been strengthened alongside the development of LLECE, which is precisely one of the Laboratory's secondary objectives.

42. Nevertheless, LLECE's promotion of internal capacity-building for the design, implementation and use of tests assessing student learning in many LAC countries, as well as the growing participation of countries in the region in other international tests, have led to an increased demand for technical quality and usability of the Laboratory's tests. As a result, the participating countries recognize as highly valuable the improvement achieved in the TERCE – and especially ERCE – assessments, but expect even further progress in future LLECE assessments, thereby raising significant technical, management and funding challenges for the years to come.

Capacity building and collaboration

43. Experts and coordinators from most countries reported that another effect of participation in LLECE has been the capacity building of the national technical teams. The Laboratory is considered to be not only an assessment implementation project, but also a training space, which was strongly supported by the latest studies. The interviewees consider that the Laboratory offers multiple opportunities for learning new and more robust methodologies for assessments and results analysis. The national coordinators find the workshops that are held to be of high quality and relevance. Workshops have moved from a focus on understanding and developing the reports to a broader approach that includes capacity building for the exploitation of national data and the analysis of associated factors, inter alia. Incidentally, as in other aspects, the effects in this area are heterogeneous depending on the assessment capacities of the countries, as well as on the stability of their technical teams.
44. However, one difficulty in this area is the restriction of some countries' own resources for participating in training courses; thus, in some cases, they have not been able to participate in all the workshops, or not all the professionals of the teams whose presence was required have been able to attend. Moreover, there have not been any online training sessions that could have reduced costs. In addition, it was pointed out

that training and workshops should use strategies and actions to provide differentiated support according to the capacity level of the countries, or when there are team rotations due to changes in governments, which, in LAC, is relatively frequent. Likewise, some coordinators suggested that introductory workshops could be held in the days prior to regular training for those countries with technical teams that had lesser capacities in certain aspects. Furthermore, they posited that there was no collaboration with other UNESCO institutions in the region, such as the UNESCO International Institute for Educational Planning's Office in Buenos Aires (IIEP-UNESCO, Buenos Aires), which would allow for the development of a broader strategy for capacity building and amplify the Laboratory's impact.

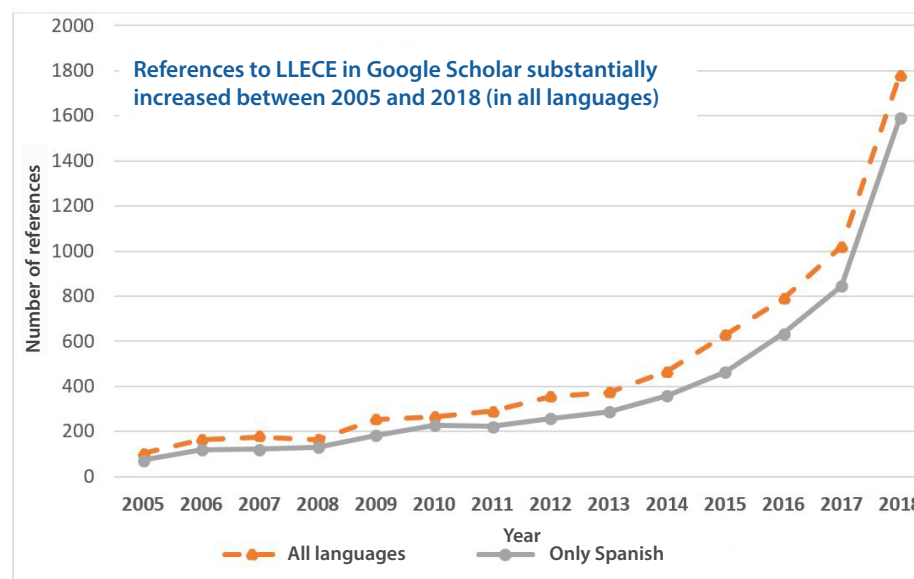
45. On the other hand, all interviewees point out that the Laboratory has allowed for a great deal of linkage and collaboration between countries, not only with respect to LLECE, but much more broadly, on assessment issues to improve or to develop their national assessment systems, for instance through reports of national assessment results that reach schools and families. Furthermore, they consider that one of the great strengths of LLECE is the exchange between countries and the formation of a network, even a certain community, between coordinators and specialists from high-confidence countries. This has given rise to frequent and smooth communication to share strategies and to clarify doubts about the different processes, as well as some missions and technical support between countries.
46. However, the national coordinators are critical of the fact that these interactions have taken place on an informal and bilateral basis rather than as a strategy developed by the Laboratory and consider that it would be very important for the Laboratory to assume a more formal and active role in this regard. To date, these have been individual country initiatives, but there is no platform to communicate or exchange materials or strategies, nor is there an internship programme to discover and learn from the experiences of other countries.
47. Several of the experts interviewed, and some national coordinators, consider that LLECE could offer countries other advice or products in terms of assessments, curricula and the use of information for policies. Moreover, a specific challenge is that the focus is moving from managing the implementation of assessment cycles to building a Laboratory that also prioritizes capacity building for school improvement and training in and the dissemination of strategies to carry out assessment cycles.

Dissemination and release of results

48. One of the views that generated the most overwhelming consensus among the interviewees was that the Laboratory's communication and positioning was very weak, which was one of its greatest and most urgent challenges. The general perception is that no dissemination strategy has been developed for the Laboratory, nor a description of it, nor clear messages to deliver the results for each cycle. Dissemination has been almost exclusively limited to the public delivery of results and the release of rather technical reports containing the results and analysis at the end of each assessment cycle. Dissemination seems to have been forgotten by the Laboratory, neither prioritized nor given sufficient resources.
49. Dissemination efforts and strategies have been mainly devoted to what each country can do, as many have limited capacities and/or can count on only a lacklustre involvement of education policy authorities in this task. In several countries, communication is left to the agencies or units within the ministries in charge of the assessments, which often do not have the capacities, time, and/or resources to design and implement more comprehensive and robust communication strategies.
50. The results typically reach only authorities and policymakers, but are little known to citizens, schools, teachers and even country researchers, which supports the observations of the national coordinators. Although the results are reported publicly in the countries, are on the websites of OREALC/UNESCO, the assessment agencies and/or the ministries, and generate some debates in the press and among the authorities, they ultimately do not have the same visibility or impact as the results of other international tests such as PISA.
51. It is indicated that the results reports produced are considered to be of high quality, but are complex and very long for school actors, and even for policymakers, so that when they do have an influence, they do so only at the macro level, but not at the micro level of schools and classroom and/or school management practices. The limited capacities of most school actors at different levels also have an effect when it comes to interpreting the results of assessments and translating them into concrete actions. Among the materials on the LLECE website, the most user-friendly and highly valued are the reports on teaching contributions, which precisely address specific issues for those involved in school systems.

52. The Laboratory has not had a communication plan to position the findings and results in the media, much less to disseminate them on social media, nor have the Laboratory's coordinators devoted much time to positioning LLECE at the LAC level or in other international spaces. No strategy or products targeted to different audiences (authorities, policymakers, school actors, citizens in general) have been developed either. A communication plan was designed within the framework of TERCE, but, for various reasons, it was not possible to implement it.
53. On the other hand, regarding local dissemination, it is critical that all countries succeed in producing their national reports, both on the results of learning tests and on associated factors and curricular analysis, an undertaking that has, to date, been voluntary and has not been achieved in all countries. In this regard, although all expressed their interest in developing national reports, several expressed the need for greater support from the Laboratory to do so, which is taken into account in the planning of ERCE.
54. Finally, many of the national coordinators pointed out that few academics knew about or used the LLECE databases and that the Laboratory has not developed strategies to promote their use either. Moreover, in a significant number of countries, there are few researchers capable of analysing the data. Despite the above, in recent years, there has been a growing number of academic and popular publications that refer to the results of LLECE assessments. While in 2010 there were only 266 publications, by 2015, these had more than doubled and by 2018, they had increased sevenfold, to 1,780. This rapid increase in recent years is a testament to the relevance of LLECE's results for generating new knowledge, even taking into account that almost 90% of these publications are in Spanish, while publications in other languages such as English, the main language of academic publications, are still scarce. These results reflect the high future potential of the research associated with the ERCE results, but there is still the challenge, mentioned by the experts interviewed, of translating some of the materials into English for a wider dissemination of the Laboratory's results.

Figure 1. Number of publications mentioned each year in Google Scholar 2005-2018



Publications mentioned each year in Google Scholar with the terms «SERCE» or «TERCE» and «UNESCO» (and «educación» or «education»)

Efficiency

Adequacy of LLECE's financial resources

55. The evolution of LLECE's funding structure, created by UNESCO in 1994, has been successful in that it has generated a sustainable initiative with enormous projections, which is still in place after 25 years of operation. Initially, 64% of its total funding came from a contribution of \$1.8 million from the Inter-American Development Bank (IDB), 26% from the 13 countries participating in PERCE, and the remaining 10% from UNESCO. In the following cycles, the initial contribution from IDB was used up and donor contributions were proportionally lower. LLECE went to being financed mainly by contributions from participating countries. The resources provided by UNESCO and OREALC have remained very limited, especially those for support for the central coordination of LLECE, development projects and some limited capacity-building initiatives.

56. For TERCE (2010-2015)⁷, with 15 participating countries, as well as the Mexican state of Nuevo León, one of the main efforts made by the LLECE coordination team was the search for donations, since IDB was no longer financing the operational part. Although contributions had been obtained from six external institutions, the sum of those resources totaled only \$688,000 (one third of what IDB had contributed to PERCE) because the economic improvement of most of the countries in the region had reduced their vulnerability and thus their eligibility for receiving resources from cooperation agencies to finance the operational costs of the initiative.
57. The analysis of the financial background and the reports submitted by LLECE on this subject, as well as the interviews with the actors linked to the Laboratory, show that, in recent years, LLECE has not had an appropriate budget plan. Furthermore, it has not been sufficiently clear about the financial resources required for a new cycle of the study, resulting in a deficit in the final implementation stage of ERCE. In this regard, the countries consider that there has been little information on the initiative's original budget, expected sources of funding or the annual information that would have allowed for early adjustments to be made to the different income and expenditure items.
58. Lack of appropriate budget planning, over-reliance on donor contributions and the shorter duration of the current assessment cycle led to the misguided definition of a policy based on low annual country contributions, leaving the previous cycle's country contributions almost unchanged. Annual contributions to ERCE funding were \$15,000 for countries with lower payment capacities, \$19,000 for medium-level countries and \$23,000 for higher-capacity countries. This is despite the fact that, since TERCE, the complexity and quality of the studies had increased, which was further deepened during ERCE, implying a higher cost. The estimated cost of the full TERCE cycle was \$4 million, while for ERCE, it was \$5.2 million (LLECE/ERCE, 2019a), while no progress had been made in designing a self-financing policy for the baseline costs of implementing the study.
59. Table 1 shows that if the average contribution made by the countries between 2016 and 2019 is maintained for the years 2020 to 2021, a deficit of around \$850,000 will be generated; this sum is essential to completing the current ERCE cycle. This amount is similar to that of the reduction in contributions from external entities and UNESCO itself, a situation that seems unlikely to be reversed in the future, as UNESCO's overall resources have also been reduced and the contributions that UNESCO makes from its regular programs are concentrated in countries with lower incomes than those in

⁷ It was not possible to obtain information for SERCE or detailed background information from previous studies.

LAC. On the other hand, the accumulated experience of LLECE itself shows how risky it is to use donor contributions to finance the regular operational management of the Laboratory.

Table 1. Preliminary estimate of annual revenue for 2016 - 2021 by source of revenue (in thousands of dollars per year)

REVENUE	2016	2017	2018	2019	2020	2021
1. Country contributions to LLECE ^a	\$ 415	\$ 338	\$ 435	\$ 393	\$ 838	\$ 800
2. Contributions from external institutions ^b	-	-	\$ 132	\$ 153	\$ 26	-
3. Contributions from UNESCO regular program	-	-	\$ 48	\$ 84	\$ 50	\$ 35
4. Contributions from UNESCO – OREALC	\$ 110	\$ 146	\$ 395	\$ 40	\$ 43	\$ 26
5. Indirect (in kind) ^c contributions	\$ 60	\$ 60	\$ 106	\$ 125	\$ 112	\$ 72
6. OREALC staff support	-	-	\$ 40	\$ 40	\$ 40	\$ 40
TOTAL REVENUE	\$ 585	\$ 544	\$ 1.156	\$ 835	\$1.109	\$973

Source: OREALC/UNESCO: Minutes of the preliminary estimates for the LLECE budget for 2016-2021.

For 2020 and 2021, revenue is estimated based on the new proposed contributions, with some adjustments for 2021.

This refers to contributions from external entities within the framework of collaboration agreements or support for specific projects linked to ERCE, especially from the Spanish Agency for International Development Cooperation (AECID), the United Nations Children's Fund (UNICEF) and UIS.

These are estimates of the value of non-monetary contributions from various actors (collaborating entities, the participating countries themselves) for the organization of workshops, coordinator meetings or technical support.

Table 2. Preliminary estimate of annual expenditure for 2016-2021 by expenditure item (in thousands of dollars per year).

EXPENDITURE	2016	2017	2018	2019	2020 ^e	2021 ^e
1. Implementing partners	-	\$ 162	\$ 741 ^a	\$ 280	\$ 190	\$ 236
2. LLECE coordination staff	\$ 223	\$ 245	\$ 273	\$ 286	\$ 304	\$ 304
3. Participation in meetings	\$ 19	\$ 33	\$ 43	\$ 50	\$ 52	\$ 60
4. Preparation of meetings	\$ 60	\$ 60	\$ 60	\$ 60	\$ 60	\$ 60
5. Additional technical support	\$ 17	\$ 165	\$ 34	\$ 35	\$ 203	\$ 306
6. Miscellaneous expenses (equipment, mail, etc.)	\$ 0	\$ 0	\$ 26	\$ 40	\$ 67	\$ 67
7. UNESCO HQ administrative costs	\$ 29	\$ 24	\$ 30	\$ 27	\$ 59	\$ 56
8. OREALC staff support	-	-	\$ 40	\$ 40	\$ 40	\$ 40
TOTAL EXPENDITURE	\$ 349	\$ 689	\$1.248	\$ 819	\$ 975	\$1.123

Source: OREALC/UNESCO: Minutes of the preliminary estimates for the LLECE budget for 2016-2021.

a: Provisioned expenses for the payment of contributions in 2019 are included.

e: Estimated expenses

60. In this context, LLECE proposed to make up this deficit by increasing the 2020 contributions to \$50,000 per country, as well as requesting an extension of ERCE until 2021, including an additional contribution of this same amount. Given the increase in time and contributions required, it was agreed that LLECE would deliver two new products that would add further value to the participation of all ERCE countries. The first is a national final results report and the second, a national analysis of the curricular framework, identifying priority areas for the future work of educational teams and increasing the impact of the study in each country.

61. This new budgetary framework will mean that the average contribution per country will be of \$179,000 for the entire ERCE cycle, still far below the costs of other learning measurement programs in which LAC countries participate (see table 4 below). It will also mean that the contributions of other entities to the total cost of the cycle will still amount to almost 40% of the total cost. Despite the fact that most national coordinators understand the financial reasons for this decision, and that at the assembly at the end of 2019 (LLECE/OREALC, 2019) it was agreed to support efforts to enable their respective countries to approve the increase, this unexpected change in the budgetary framework at such a stage of advancement presents various complex difficulties for the countries. Among the main difficulties is inter-institutional coordination within each country, as the authorities that approve this additional funding do not necessarily belong to the institutions where the coordinators work. Moreover, some mention the complexity of modifying an international agreement in which the initial deadlines for the delivery of certain committed results will not be met and the resources requested increased. This situation has generated a degree of uncertainty regarding the sustainability of ERCE itself, as some authorities see that without the additional resources, the results of the study under development might not be available.
62. In this sense, some country coordinators point out that OREALC/UNESCO should have first brought this kind of situation up directly with the political authorities of the respective countries, ministers or vice-ministers of education, rather than beginning with the coordinators and sending a written communication, given that it is a political and administrative problem rather than a technical one. In some cases, it is mentioned that an additional increase will alter the total expenditure already made by the countries to implement the study and their current fiscal conditions.
63. A second component that accounts for the insufficient planning of ERCE is that, during this assessment cycle, multiple opportunities and products were committed to generate greater added.

Table 3. Additional complementary studies proposed in ERCE

Name	Description
1 Innovation module	Analysis of the tool and generation of the results report
2 Contributions to teachers	Series of four publications, one for each area evaluated
3 Report on teachers	Thematic report with an additional analysis on variables regarding teacher issues, associated with student performance
4 Standard setting	Workshops to define cut-off scores for the 2019 ERCE performance levels
5 Correction of open questions	Centralized support from MIDE UC Measurement Centre in coding open-ended questions and the writing test

64. A third factor that exacerbates the problem of financing ERCE is the extension of the planned time frame, which was increased from five to six years, implying additional operating costs. This delay also implies postponing the delivery of the committed results by one year, whereas the political and technical authorities of the countries require them for the discussion and evaluation of public policies in the sector or the identification of priorities for those policies. Nevertheless, the effective six-year ERCE cycle has been a considerable improvement compared to previous cycles. PERCE thus had a complete cycle that lasted nine years before the publication of its results and SERCE and TERCE, seven years.
65. In 2019, LLECE's coordination team and the entities participating in the Strategic Advisory Board (CAE) made a remarkable effort to face this financial situation, seeking alliances with different cooperation entities that could finance the innovative components of the study. However, in a medium-term perspective, the regular financing of a new assessment cycle must be settled to determine which of these products will become essential and common to all participating countries, and which will be financed by each country's baseline contribution

Comparability with other international assessment programs

66. Large-scale standardized education assessments have been used in comparative studies for 60 years, but their generalization and global discussion of results have mainly developed in the last two decades (Brunns et al., 2019). In this context, LLECE is recognized as the entity that provides the most important systematic assessments for the countries of LAC, both because of its seniority (dating to before the creation of PISA by the OECD) and because it represents the standard for participating countries. It is also recognized among its peers for its growing technical legitimacy and the valuable mechanisms for participation and collaboration among the Laboratory's member countries. Finally, most of the national coordinators and experts interviewed emphasized its enormous potential for the discussion of public policies to improve schools in the continent.
67. Although there are more international experiences of large-scale assessments, we compared some of the attributes of ERCE with those of four other assessment alternatives in which LAC countries could participate.
- These are:
- The Program for International Student Assessment (PISA), developed by the OECD;
 - PISA for Development (PISA-D), an OECD proposal for less developed countries that have not previously participated in PISA;
 - Trends in International Mathematics and Science Study (TIMSS);
 - Progress in International Reading Literacy Study (PIRLS): both TIMSS and PIRLS are developed by the International Association for the Evaluation of Educational Achievement (IEA).
68. ERCE stands out for its high coverage in LAC, although only among Spanish- and Portuguese-speaking countries (Brazil in the latter case), as it has never included either the English- or French-speaking countries of the region. Unlike the four tests mentioned, ERCE is of a regional nature. The others are not specific to a continent or language.

69. Gradually, an increasing number of LAC countries have adopted the PISA test. 14 of the 18 countries participating in ERCE recently participated in PISA or PISA-D. Only the Plurinational State of Bolivia, Cuba, El Salvador and Nicaragua do not participate in PISA. However, LAC participation is very low in IEA assessments, with Chile being the only country in the region to participate in both.
70. One of the main advantages ERCE has over PISA, as pointed out by the national coordinators, is its focus on primary school students (third and sixth grades), since this allows for the early identification of challenges and achievements in educational policies by country, while PISA and PISA-D involve the participation of 15- to 16-year-old students in their assessments, enrolled mainly in secondary education.
71. Notwithstanding the above, the PISA and PISA-D assessments are also valued by several of the coordinators interviewed, as they provide ERCE with complementary information, particularly with respect to secondary school students. For example, both assessments provide relevant evidence of Sustainable Development Goal (SDG) 4.1 of the 2030 Agenda. One of the most noteworthy innovations of PISA-D is that it includes a module for dropout students; early leaving of school is quite common among LAC countries, but not necessarily among primary school pupils, as most children regularly attend school throughout their first years of education.
72. The interviewees also indicated additional advantages of ERCE compared to PISA, such as the work done by LLECE to construct all implementation stages in a participatory manner, capacity building for national technical teams and the fact it responds to a regional curricular proposal, rather than one that is externally defined and that reduces the contextualized identification of its own challenges.
73. Additionally, costs are relatively higher for PISA and PISA-D than for ERCE. While the average contribution of the countries participating in ERCE will amount to \$179,000 over the entire cycle (including the adjustments proposed for 2020-2021), the cost of PISA, for a country that has not previously participated, is \$275,000, including support in the different processes and the national report⁸, which is included in contributions to LLECE. In other words, PISA costs \$96,000 more than ERCE.
74. These differences of cost between ERCE and PISA also account for the higher added value of the Laboratory, as, with fewer financial resources, ERCE includes valuable additional results such as the curricular analysis described previously and the assessment of two educational levels (third and sixth grades), while PISA assesses only one. In addition,
- ERCE also assesses students' writing skills, while PISA does not.
75. An important lesson LLECE drew from the PISA-D experience is that when large-scale tests are applied to vulnerable countries that have fewer technical capacities, an extensive cycle is required for their implementation. While PISA has three-year cycles, PISA-D⁹ requires five years, reflecting the fact that the coordination stages for the signing of agreements, alignment and initial technical support are more extensive for these countries than for the average country participating in PISA. Another observation was that implementation costs for the most vulnerable countries can require much more financial resources than those of countries that are not vulnerable or that have more developed technical capacities. In some cases, the contributions of the countries participating in PISA-D were even double the contributions of countries participating in the regular PISA program.
76. Despite these differences and the advantages of ERCE over PISA, several interviewees noted that PISA had several valuable features, as did the institution responsible for it, the OECD. The first is PISA's international and public validation: PISA enjoys a high degree of recognition and its national evidence and inter-country comparison, which involves both countries having a similar level of development and the most developed countries, have a significant impact on public policy discussion. The second characteristic is its higher level of planning and its adequate compliance with deadlines for different milestones, such as technical support and timely clarification regarding the national teams' technical and management doubts during the implementation period. A third advantage is technological innovation, which has made it possible to include an interactive assessment for interested countries, which improves the assessment. A fourth advantage is the incorporation of new assessment themes, including new topics such as social-emotional skills and financial literacy, some of which are mandatory in assessments, and others of which are optional for interested countries; there is even the designing of methodologies that allow for longitudinal studies, as in the case of Canada.
77. A fifth advantage is the quality and diversity in the delivery of results for each cycle. Great value is placed on the user-friendliness of the reports, the drawing of general conclusions and the identification of factors and conditions for improvement, but also the generation of specific background information for each country, with the national reports of the PISA-D participants, published one year after the collection of the data, being particularly important. It was also mentioned that the gradual delivery of these results is useful. It is not only the delivery of the

⁸ For new countries, the cycle financed is of four years, although the results are delivered within three years, since the fourth encompasses the production and publication of other results such as the specific assessment reports for each subject.

⁹ No formal information on the costs of the initiative was obtained. However, this initiative benefited from contributions from ten development agencies and cooperation agencies and three private entities. These contributions partially financed the direct costs of some countries, the development of some national reports, the training workshops and technology to implement the project at the national level (OECD, 2017).

comparative report that is relevant, but also the development, in the following years, of specific reports and specialized webinars and seminars, so that the analysis and dissemination of results occur together with the implementation of a new assessment cycle in the field.

78. A sixth advantage is the easy public access to the results, reports, databases and technical manuals, as well as the fact they are in different languages, thereby promoting research on and the continuous dissemination of the results obtained. Publication in English is a critical attribute, since most of the research is carried out in English. This would enable the extension of the Laboratory's scope and the expansion of inter-institutional relationships for the development of technical and financial collaboration projects.
79. A final advantage is the institutional strategy of international dissemination of the results adopted by PISA's highest authorities; this applies to visits and presentations of the main results in various countries, as well as to webinars or other communication mechanisms. At the same time, there is a much closer institutional relationship between the OECD and PISA, and greater recognition, than between UNESCO's Headquarters in Paris and LLECE/OREALC. Thus, the priority challenges lie in making progress in each of the areas identified, especially in the dissemination of the results of the current cycle, as well as in the preparation of the next assessment cycles.
80. However, despite massive participation in PISA and PISA-D on the part of the countries assessed in ERCE, one of the coordinators concludes that, notwithstanding PISA's greater national and international visibility as compared to that of ERCE, both assessments suffer from similar weaknesses: implementation and dissemination run parallel to the design, assessment and implementation of public policies on education. In that regard, LLECE could strive towards the distinctive objective of shifting from a focus on assessment management to a focus on the development of capacities and strategies for educational improvement and decision-making regarding public policy.
81. On the other hand, TIMSS and PIRLS together have many attributes similar to those of ERCE. These two tests assess students at levels close to those of ERCE (fourth and eighth grade in TIMSS and fourth grade in PIRLS), as well as similar disciplines: TIMSS assesses math and science, while PIRLS assesses reading, but not writing. Similarly, both assess a curricular framework which is common to all participating countries, but which in TIMSS and PIRLS is not collaboratively constructed, as it is in ERCE. In addition, as is the case for ERCE, these tests are done on paper and not by computer, although this will change in the next cycle.
82. The experience of IEA, with its two tests, can reveal some important lessons for LLECE. The first is that since these two tests take place in four-year (TIMSS) and five-year (PIRLS) cycles, it is possible that ERCE will improve its implementation capacity in order to take place in five-year cycles, as initially planned; however, this also requires that all countries rigorously keep to their respective schedules, something which has not always been accomplished for LLECE. Still, a much greater reduction is not feasible, because of both the heterogeneity of the participants and the greater

complexity of the assessments (levels and areas) and the need to develop a shared curricular framework.

83. In addition, IEA simplifies the development of the shared curricular framework, which makes it possible to reduce costs and time for the implementation of its assessments. For OREALC and LLECE member countries, this is a subject for urgent discussion. On the other hand, IEA is at the forefront of assessment innovation. For example, in the current cycles, it proposes a migration to interactive assessment (which does not exclude the paper alternative); this makes it possible to improve presentation and the interaction between the tool and students, which in turn makes it possible to improve the validity of assessments, the identification of factors and conditions associated with the associated learning, and which therefore will affect the future design of public policies for educational improvement. Several LLECE countries identify progress in this area as a priority; this is especially so among those countries which have systematically participated with other international tests and which acknowledge that the ERCE tests are at a technical disadvantage, as there is no discussion regarding the incorporation of these innovations.
84. IEA employs another good practice: it enables participating countries to include additional grades in the reading assessment, at an additional cost of 50% of what is required for fourth-grade students, as well as different sub-samples for the purpose of including regional or group-specific representations of interest to a country. These are characteristics which, as noted by some countries participating in ERCE, should be further developed in upcoming LLECE assessments.
85. Finally, a very important attribute of the IEA assessments is their close link to research opportunities based on their results. These opportunities are promoted in multiple ways, including the development of an academic journal and agreements with publishers regarding technical publications, as well as through the development of strategic alliances to provide support services and receive contributions from donors, and through inter-institutional collaborations such as the Rosetta Stone project itself, which is explained further below.
86. However, part of the differences identified between IEA and LLECE can be explained by the enormous disparity between the financial resources of the two institutions. While IEA assesses about 60 countries every cycle, LLECE only assesses 18. If we consider together both assessments, in pencil-and-paper format, for two grades, one ERCE cycle costs around \$289,000 per country; meanwhile, TIMSS-PIRLS would go up to \$835,000, without even taking into account the assessment of students' writing skills, which is not covered by PIRLS. On the other hand, while the government of each country participating in ERCE pays nearly \$179,000 for the cyclical assessment, in the case of TIMSS-PIRLS, the amount contributed is the total estimated expenditure, which means that each country pays 4.5 times the average current cost of ERCE.

Table 4. Comparison of the characteristics of ERCE and of other international assessments

	ERCE	PISA	PISA-D	TIMSS	PIRLS
Entity responsible	OREALC/UNESCO	OECD	OECD	IEA	IEA
Number of participating countries or entities (<i>LAC countries</i>)	18 (2019) (18)	79 (2018)(10)	9 (2018)(4)	64 (2015)(2)	57 (2021)(2)
Levels assessed	3 rd and 6 th grade, primary	15-year-old students	15-year-old students (and young dropouts)	4 th and 8 th grade, primary	4 th grade, primary
Areas assessed	Mathematics Reading Writing Science (6 th grade only) Socio-emotional skills (6 th grade)	mathematics Reading Science Additional skill ^c	Mathematics Reading Science	Mathematics Science	Reading
Assessment cycle	2016-2021 (6 years)	2019-2022	2014-2019 (5 years, but period from agreement to results lasts 3 years)	2019 (4 years)	2018-2022 (5 years)
Payment per cycle (USD\$; per country, for expenses of institution responsible); does not include national implementation costs	178.812 ^a	220.500 ^d - 275.000 ^e	267.000 – 746.000 ^f	472.500 ^g - 619.500 ^h	241.500 ⁱ - 393.750 ^j
Actual cost per cycle (USD\$), without national implementation costs	288.991 ^b	220.500 ^d 275.000 ^e	s/i	472.500 ^g 619.500 ^h	241.500 ⁱ 393.750 ^j
Funding from country participating in evaluation	61,9% ^a	100%	100%	100%	100%

^a Preliminary figures: This 2016-2021 cost takes into account the estimates provided by LLECE, which include a significant increase in the contributions of participating countries for 2020-2021; the proportion of financing from country contributions went from 37.6% in 2018 to 75.6% in 2020. According to this assumption, the percentage of total LLECE expenditure covered by the contributions of the ERCE participants would come to an average of 61.9% for the period, for which reason this is the estimated cost amount for each participating country.

^b Preliminary figures: Amount which reflects the total cost per LLECE participant, including resources from other.

^c Every period, PISA contemplates the assessment of an additional skill (in 2018, overall competence was tested, and for 2021, creative thinking will be tested). It also offers some optional additional assessments, for which additional financial contributions are required from countries (e.g., financial literacy in 2018).

^d PISA: Payment by the country to participate for the first time in a four-year cycle (base cost).

^e PISA: If support, such as data analysis and reporting, is included for the preparation and implementation stages.

Assumption: Official data are reported in euros, for which reason they have been converted to dollars at an exchange rate of 1 euro to 1.1 United States dollars.

^f PISA-D: These are the reported contributions made respectively by Paraguay and Ecuador to PISA to participate in the initiative. As is the case in all standardized assessments, these amounts do not include the national.

^g TIMSS: Cost of participating on both levels evaluated using a paper support. The country can only participate in one level so the cost has to be half of what it is indicated.

^h TIMSS: When countries can participate on a numeric evaluation, it is asked an additional fee.

ⁱ PIRLS: It is possible to contract the evaluation for additional grades and its costs represents 50% of the cost for fourth grade.

^j PIRLS: It refers to the digital format which allows to deliver richer results.

Sustainability

Relevance of LLECE to the 2030 Agenda¹⁰

87. LLECE plays a strategic role in the 2030 Agenda for Sustainable Development, agreed by 193 countries with the United Nations to achieve 17 major goals by 2030. Within this framework, Sustainable Development Goal (SDG) 4 («ensure inclusive and equitable quality education and promote lifelong learning opportunities for all») comprises the targets related to education. This entails progress from achievements in educational coverage to a greater commitment to quality, equity and inclusion in educational opportunities.
88. Specifically, SDG target 4.1 states, «ensure that all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and Goal-4 effective learning outcomes» (United Nations, 2015). OREALC is the entity responsible for monitoring and supporting LAC countries as they work toward achieving SDG 4. In particular, LLECE studies make it possible to generate data for monitoring SDG 4.1.1 indicators (specifically, indicators 4.1.1 [a] and 4.1.1 [b])¹¹ for the 18 ERCE participants [2016-2021]), as LLECE offers the only regional database with national and comparable coverage for primary school students.
89. In this regard, all the interviewees consider LLECE to be strategic, since not only does it have the background information with which to establish a baseline, using ERCE 2019 and previous TERCE results, but it also will facilitate monitoring throughout the decade with subsequent agreed assessments.
90. At present, the dialogue which LLECE and OREALC have with LAC countries is not based on progress toward fulfilment of 2030 Agenda commitments; rather, it focuses on three areas: i) the development of monitoring and reporting mechanisms; ii) the generation of technical capacities for this development; and (iii) making use of the ERCE 2019 reports, to be published in 2021, to include a section on SDG 4.
91. For the countries participating in ERCE, the primary importance which LLECE holds with regard to the 2030 Agenda resides in its provision of information critical to the monitoring of SDG 4.1 indicators and in the fact that LLECE has enabled a shared discussion among the countries in

10 The connections between comparative learning assessments and the 2030 Agenda are discussed in detail in the UNESCO report entitled *The Promises of Large-Scale Learning Assessments: Acknowledging Limits to Unlock Opportunities* (<https://unesdoc.unesco.org/ark:/48223/pf0000369697>).

11 SDG indicator 4.1.1 (a) refers to ensuring quality and equitable education for boys and girls in grades 2 and 3, in both language and mathematics, and 4.1.1 (b) refers to achieving this target by the end of primary school.

order to develop common indicators. To achieve this objective, the discussion on methodologies, the baseline and advances requires LLECE coordination which goes beyond the work conducted by the assembly of coordinators, since in several countries, the 2030 Agenda is managed by institutions not directly associated with that work. The assembly of national coordinators, which met in Guatemala in late 2019, confirmed the commitment to the 2030 Agenda and LLECE's responsibility to support countries in the monitoring of SDG indicators 4.1.1 (a) and 4.1.1 (b), which is transversally supported by the participating countries. For some, such as Cuba and the Plurinational State of Bolivia, this is even the only source of internationally comparable information for the purposes of monitoring.

92. However, the results of the ERCE tests are not sufficient for adequate monitoring of SDG target 4.1 indicators. This is because of a still-unresolved technical issue arising from the fact that there has yet to be a definition of the minimum level of learning which is to be used for establishing the 2019 baseline and for monitoring in the lead-up to 2030¹². To resolve the matter, LLECE, together with the UNESCO Institute for Statistics (UIS) and IEA, has been participating in the Rosetta Stone project, financed with resources from the Bill and Melinda Gates Foundation. The aim of the project is to make it possible to compare ERCE results with those of different standardized tests worldwide. This experience is one of the best examples of the strategic coordination which should be promoted more among LLECE and other UNESCO institutions. UNESCO Headquarters should strengthen its institutional and communication links with LLECE, highlighting more clearly the strategic role which UNESCO assigns to LLECE with regard to educational improvement processes in the region and to the generation of learning at the global level.

12 Currently, half of the countries concerned do not have comparable indicators for assessing the baseline and progress with regard to SDG target 4.1 indicators (Bruns, Akmal and Birsall, 2019).

The Rosetta Stone projet

The Rosetta Stone projet is an effort to compare internationally the progress of SDG target 4.1 indicator 4.1.1 based on different standardized tests applied around the world. It is led by the International Association for the Evaluation of Educational Achievement (IEA) with regional study coordination centres.

The projet has two main objectives. The first is to link ERCE and other regional assessment studies in Africa, Asia and Oceania which are conducted at the end of primary education with the TIMSS and PIRLS global assessments.

The second is to create a correspondence table which makes it possible to compare regional assessment results with regard to the TIMSS mathematics and PIRLS reading scales.

The first step is being taken with the regional studies of ERCE and the Program for the Analysis of Educational Systems of the CONFEMEN Countries (PASEC, which includes 10 countries). ERCE's participation is financially supported by the Bill & Melinda Gates Foundation and managed by the UNESCO Institute for Statistics (UIS).

By establishing a correspondence table (based on three or four countries from the regional studies) it will be possible to compare the results of all the countries participating in each regional assessment associated with the table. This way, students who achieve the minimum reference levels in all the countries assessed will be available in 2021, making it possible to repeat the linking exercise for the ensuing assessments.

Source: Presentation by Oliver Neuschmidt y "El proyecto de vinculación Rosetta. Midiendo el progreso global hacia el ODS 4.1".

93. LLECE plays a broader role than simply providing future information for the monitoring of these indicators; it also actively participates in the Global Alliance to Monitor Learning (GAML), in coordination with UNESCO and other international entities, such as the Organisation for Economic Co-operation and Development (OECD), IEA and the UIS, participating in the initiative's regular meetings, discussing documents and guidelines, and taking part in concrete initiatives. In this arena, it participates in the improvement of the mechanisms for global monitoring of SDG 4 indicators, as well as in the formulation of support and outreach strategies for progress in achieving LAC countries' targets. Within this framework, LLECE, along with the OECD and IEA, agreed that the public reports of the comparative assessments should include a section on countries' progress with regard to SDG 4. In addition, since 2018, it has participated in the Rosetta Stone project, which reflects international recognition of the technical quality and relevance of the information produced by the Laboratory's work.

94. LLECE's strategic role will become utterly critical throughout the rest of the 2020s. That is because the decade is all the time which is left for achieving the Sustainable Development Goals of the 2030 Agenda; it is also because many countries do not have enough timely data for defining and assessing the targets of the 2030 Agenda. This very lack of data brings with it a difficulty in adequately informing decision-makers, reducing the likelihood of meeting the targets committed to (United Nations, 2019).
95. In addition, LLECE can help to assess the progress achieved with regard to several other 2030 Agenda indicators, beyond target 4.1, such as target 4.5 and target 4.7. The first concerns the elimination of disparities in access and performance among different groups of students who are vulnerable for reasons including gender, lower socioeconomic status, disability, indigenous ethnicity and the effects of conflict. The target 4.7 indicators refer to the acquisition of knowledge and skills needed to promote sustainable development.¹³ There is also the possibility of helping with the monitoring of target 4.a indicators (which concern school environments, equipment and conditions) and target 4.c indicators (which concern teacher training support and qualification conditions).
96. This potential for expanding LLECE collaboration in the monitoring of SDG 4 beyond target 4.1 is relatively unknown among the countries concerned, although the coordinators from Panama and Cuba are promoting progress in this regard. To this end, they propose incorporating these targets as a strategic objective promoted by UNESCO and LLECE/OREALC into both the design of the ERCE 2019 results report and subsequent cycles.
97. It is in the context of the 2030 Agenda that an enormous challenge arises not only for LLECE and OREALC, but also for liaising with the rest of UNESCO's institutions. The end of the current ERCE cycle, as well as the prospects for supporting LAC countries in reaching their targets by 2030, brings with it the need to implement two additional ERCE cycles in the 2021-2030 period. One for monitoring progress and the other for delivering the final achievements related to the indicators, not only for target 4.1, but also for all targets with a high potential for treatment with robust and comparable evidence.
98. At its meeting in Guatemala in October 2019, the assembly of national coordinators agreed to work towards implementing these two cycles. Despite this fact, such implementation involves enormous challenges. The first challenge is that of

¹³ Specifically in: i) global citizenship education and ii) education for sustainable development, including gender and human-rights mainstreaming in national education policies, curricula, teacher education and student assessment

considerably shortening the implementation period for the two cycles, since it would require completing both within a nine-year period; the current cycle is six years long (2016-2021).¹⁴

99. Another challenge consists in the technical discussion of the shared curriculum framework which is implemented in the course of each assessment cycle. Although this is an aspect of the program which is highly valued by the participating countries, it involves a time-consuming process and could only be carried out as part of some cycles, especially if the duration of the cycles is reduced. In addition, there should also be discussion about its current construction and the adequate representativeness of the disciplines involved, since currently, the results could be comparable in time, but they require improved identification of the skills and knowledge which students should have for the educational levels assessed.
100. This discussion is linked to the definition of the minimum performance levels expected and comparable with those of other tests (Rosetta Stone project). This is because the achievement of target 4.1 requires the definition of the minimum performance level on an international scale. This involves a modification of the levels used in the delivery of the TERCE results in the previous cycle; this is due to the estimation that these minimum levels could be above TERCE level 2 (defined as the adequate level), but below level 3. A simple solution which would clarify this change would be to have the ERCE-TERCE comparison reports deliver the results with both definitions of the student achievement levels, facilitating the transition to the definition of the new levels.
101. A third challenge consists in ensuring the Laboratory's financial sustainability and identifying the differentiated technical support and conditions which the countries would require in order to address the objective of implementing two cycles in what remains of this decade.
102. Finally, progress must be made towards generating more evidence for the discussion of public policies and improvement strategies within each country. With regard to the definition of a sequential plan for the analysis and dissemination of the ERCE 2019 results, said results cannot conclude with the delivery of the products committed to (curricular analysis, results report, and associated factors report).

¹⁴ For curricular tests such as TIMSS and PIRLS, the cycles are four and five years long respectively. In these cases, there are curricular frameworks which are not developed in such a participatory manner, making it possible to incur lower costs and require less time for advancing to subsequent stages. At the same time, since most of the participating countries are developed countries, they can afford higher financial costs and support their respective technical teams.

103. Thus, the planning of the new LLECE cycles must be compatible with the generation of new products and the in-depth discussion of the results of the ERCE cycle. Otherwise, all efforts would be concentrated only on the implementation of the studies and not on the studies' usability for the improvement of school systems. LLECE's enormous potential for assisting countries with increasing the value of assessment results necessarily implies sufficient and assured funding from the countries concerned and from the potential complement of donor entities ensuring its sustainability; otherwise, it is better to limit expectations only to the basal products of each assessment cycle.
104. Thus, it seems relevant that the discussion regarding the new cycles and the technical, management and financial aspects entailed should also be undertaken by UNESCO with the corresponding authorities of the region's ministries of education, and not only at the meeting of the assembly of coordinators (which, it has already been agreed, will take place in the second half of 2020). This would strengthen institutional recognition of LLECE/OREALC, in addition to demonstrating LLECE's strategic value for UNESCO's priorities in the new decade.

Financial sustainability

105. study concludes that the most relevant challenges facing LLECE include its short-term financial sustainability and the redesign of its medium-term financing structure. Said redesign applies both to the new LLECE cycles and to the Laboratory itself, if it intends to evolve, in the medium term, into an entity which goes beyond the implementation of the assessment cycle. These challenges were identified as urgent by several of the country coordinators, as well as by all the experts and institutional representatives we interviewed.
106. LLECE put forward a proposal to resolve ERCE's chronic deficit, which was financed with contributions from OREALC¹⁵/UNESCO, related entities and other donors until 2019, but which is not sustainable given the reduced resources of UNESCO itself. This proposal, which would make it possible to complete the current assessment cycle by means of additional contributions from participating countries, is uncertain. This is because at the end of this report there were different situations between countries, which were mostly explained by the administrative difficulties involved in changing the international agreement for the project, but which were also the result of political,

¹⁵ The contributions made by OREALC/UNESCO to finance these deficits have even entailed advancing resources for the years 2020 and 2021, illustrating the urgent need for additional contributions from the corresponding countries for the conclusion of the current assessment cycle.

social, or financial changes in some countries and the difficulties associated with the coronavirus disease (COVID-19) pandemic in recent months. For example, in the cases of Chile and the Plurinational State of Bolivia, social and political problems prevented the administering of the tests at establishments in 2019, and consequently, prevented ERCE from having the basic information needed for assessment. In Argentina, because of the recent change of government, under absolutely critical financial circumstances, it was not yet possible to ensure the contribution of additional resources.

107. There is no denying the commitment of the LLECE country coordinators to ensuring the proper closing of the current assessment cycle. This is reflected in the commitment established at the assembly meeting in Guatemala in late 2019 to manage their countries' formal commitments with the additional resources. It is also reflected in the creation of a commission of six member countries to identify proposals to help resolve the financing issues of the current cycle, as well as to make better use of the tremendous opportunities for better products and results from ERCE. However, in almost all cases, the coordinators do not directly resolve these financial issues.
108. The lesson learned from this experience is that LLECE needs better budget planning, systematic monitoring of accounts related to said planning, and the provision of timely financial information to the corresponding member countries. All these commitments were recently agreed upon between the LLECE coordinating officials and the member countries. They are indispensable for strengthening institutional relations, as well as for the planning of a new assessment cycle.
109. In addition, a relevant conclusion stemming from this experience is that the future sustainability of the LLECE cycles must be financed mainly by the regular contributions agreed upon with the participating countries, as is the case of other international assessments such as PISA, TIMSS or PIRLS. For this, the planning of activities and the extension of assessment cycles absolutely must take place; otherwise, there will be an impact on the costs and usability of results with regard to the national authorities. Also indispensable is proper and accurate estimating of the budget required to implement the study, which explicitly defines the products which each country will have as a result of the contributions in question, as well as the additional resources which would be required of each country if it chose to include additional products which LLECE could provide, as is currently the case in the PIRLS tests.
110. LLECE/OREALC cost projection should be analysed in relation to the additional functions required; in addition, the sources of its funding should be identified. In the context of the

regular operational management of the ERCE cycle, it seems essential to strengthen the LLECE central coordination team, introducing greater specialized technical capability, such as that afforded by a communication and outreach unit. It is also essential to consider the provision of resources in order for the main results and resources of LLECE to be available in other languages. Also needed is the inclusion of resources in order for the Coordinator to play a more active role in LLECE's international presence, especially in the countries participating in the Laboratory. This additional investment should be included in the regular funding provided by the countries. Since this evaluation is not the referent for determining this cost for a new cycle, consideration should be given to setting the minimum annual contribution from each country at \$50,000 in actual value, based on a five- to six-year cycle. This goes beyond the continuity of efforts to increase the contributions of donors and collaborators.

111. All the coordinators interviewed noted the major progress achieved during the current ERCE assessment, its enormous potential, and LLECE's relevance with regard to the generation of national technical abilities in the area of assessment processes, collaborative work among the countries of the region, and the enormous expectations regarding its increased impact on processes for the improvement of school systems. However, in a context where most of the costs of baseline assessment operation are, in future, borne by the countries concerned, there is an emergence of expectations or proposals to maintain, among certain countries, high interest in participating in future assessments, especially when most of the countries have as their reference the benefits and costs of participating in PISA. Consequently, a structural change in the financing system entails a discussion between UNESCO and the highest education authorities of each of the various countries in the region.
112. The suggestions mentioned by the countries include, in addition to changes in planning and financial management: (i) a further reduction of the gap there exists with regard to other assessment programmes, especially in terms of format innovation, allowing progress towards more interactive models between the tool and students; (ii) the carrying out of planning and the assessment cycle; (iii) the establishment of a priority focus on other skills, furthering the work initiated in ERCE; (iv) the enhancement of the complementary functions of LLECE: capacity development, technical support, promotion of collaboration between countries, and learning from best practices; and (v) much greater efforts to disseminate and ensure the potential usability of results. However, most of these proposals entail significant additional costs, which need to be compatible with a new financing framework.

113. Since most participants have some connection to PISA, some interviewees compared the new contributions for 2020-2021 to LLECE with those required for participation in the OECD assessment. In that respect, it is very important to emphasize the differences between the two tests as regards coverage of the levels assessed and dimensions, as well as the national teams' capacity development and collaborative networking. LLECE should, however, also make a proposal to address the technical challenges cited above to continue to replicate the best practices of the other assessments in future LLECE cycles. For example, an interactive computer-based methodology could be developed for assessment responses for those countries interested in this approach and willing to cover the associated costs.
114. In addition, the Laboratory must ensure that most LAC countries can and will continue to participate in LLECE. This is not only because of its status as a public asset and its development of better school-improvement opportunities for all its members, but also because the number of participating countries is a variable critical to the achievement of economies of scale in the operational costs of ERCE and in the estimation of the cost per country. Unlike the PISA, TIMSS or PIRLS tests, there is not much room in ERCE to increase the number of participating countries. Including the countries not currently participating would mean including English- and French-speaking Caribbean countries, which would be an additional development project separate from what currently comprises regular LLECE operations.
115. Given countries' diverse financial capabilities, it seems advisable to maintain differentiated conditions for some less-developed countries. However, the corresponding group should be very small, doing away with the existence of three groups. To the extent possible, the funding for this smaller contribution should come from LLECE/OREALC itself or from a multilateral entity.
116. This way, the agreements between LLECE and each country could be differentiated. All would share a common basis, but additional costs would be defined if an agreement is made to include supplementary support requested by a given country. Some countries mention this request, although it is not clear whether their respective ministerial institutions would provide the financial support for it. The PISA experience should serve as a reference, since for the entire cycle it offers countries a base cost which increases by €45,000 if there is a request for additional assistance with preparation and implementation, as well as with data analysis and national reporting.
117. Consideration could also be given to the development of training programs in addition to those contemplated in the regular program. For example, El Salvador indicated that in the next agreement, it would like to see resources added for this purpose. There could also be a definition of resources differentiated by the countries interested in having LLECE assist them with the analysis of additional or specific components, for example, in the case of national modules. This way, LLECE could generate different support services which could be self-financed and which could generate complementary resources for the Laboratory, ensuring progress towards responsibility beyond the management of each of the ERCE cycles, many of which could be offered to more countries than those participating in the assessment cycle.
118. Finally, there is the development of innovations and programs for cooperation with different international or multilateral institutions, including UNESCO's own institutions. These initiatives should be financed with contributions from these donor entities or with resources additional to the regular contributions of the participating countries. The Rosetta Stone project stands as an emblematic experience, as do the collaboration with UNICEF on curricular studies at the national level and the collaboration with the Inter-American Development Bank (IDB) to strengthen the work of LLECE with the participation of the CTAN. The institutional experts interviewed noted that this last source of funding offers tremendous opportunities, but that it is essential that LLECE make progress in generating specific projects, which are linked to its agenda and to the areas in which it proposes to advance strategically. It should also disseminate its main products in English. In addition, UNESCO should increase its institutional support for LLECE and strengthen international strategic recognition of LLECE.
119. Finally, there is an inter-institutional dilemma regarding the priorities which LLECE should address in the coming years. On the one hand, implementation of the monitoring of compliance with SDG 4 of the 2030 Agenda has the support of UNESCO, for which reason the ability to carry out two additional ERCE cycles between 2022 and 2030 is highly relevant. However, this is a complex challenge, not only because the actual length of the assessment cycles is six years or more. Even if the cycles were indeed reduced to five years, as initially planned for the current ERCE cycle, the results of the second assessment would not meet the deadline for determining the achievements made. A second reason is that most of the countries' technical teams mention the urgent need for LLECE to make greater efforts to analyse, disseminate and learn from the results of the current cycle (2016-2021). However, in order to make progress with regard to this objective, LLECE needs to allocate more leadership and management resources, as well

as technical and financial support in the 2021-2023 period, which makes this difficult to reconcile with the management of two cycles in just a nine-year period.

120. An alternative which could help to reconcile the two objectives would be to have UNESCO propose to the ministers of the region an agreement for the entire remainder of the period. Under this agreement, the countries would commit specific resources for the first cycle and would commit to discussing the resources for the second period; however, the analysis of curricular consistency would be conducted only in the first period and not in the second. For their part, UNESCO and its associated entities could commit additional resources in order to move towards 2030 Agenda monitoring, including the identification of several targets in addition to 4.1, whose uptake among participating countries could be supported by the work of LLECE.

V. CONCLUSIONS

121. The analysis carried out allows us to come to four main conclusions regarding the evaluation of the Latin American Laboratory for the Assessment of Quality in Education (LLECE) of the UNESCO Office in Santiago de Chile (OREALC). The first is that the Laboratory is highly regarded by the participating countries and by the multilateral and cooperative entities which have had a historical relationship with it. The participating countries have pointed out multiple reasons for the relevance and importance of LLECE. These include its longevity, its focus on multiple dimensions and on the first years of schooling for children, its participatory governance mechanisms, the development of technical assessment capacities, and its collecting of evidence for the generation of better quality policies.
122. LLECE has systematically improved the quality of its processes and of the products it generates, putting it on a level with the most prestigious large-scale education assessment institutions worldwide. These actions are complemented by the establishment of permanent technical bodies and international support bodies, such as the CTAN, which improve the international technical standards of assessments. This growing international recognition has given the Laboratory a growing role in the 2030 Agenda, since it enables the monitoring and assessment of the commitments made by the countries of the region, as well as anticipating the initiative's enormous impact throughout the rest of the decade.
123. However, LLECE also faces multiple challenges with regard to increasing the impact on national education systems and to supporting the design and assessment of countries' policies and strategies for improvement. One such challenge which stands out resides in substantially improving the communication and dissemination of results. Another, in liaising more with research centres and public institutions which design educational policies. In addition, there are challenges associated with the Laboratory's raised profile and internationalization, which imply, among other things, the establishment of alliances with other international assessment networks.

On the other hand, the active participation of most LAC countries in other international tests, especially PISA and PISA-D, implies a context of continuous improvement of their processes, methodologies, dissemination strategies and inter-institutional partnerships, as well as the clear identification of the comparative advantages with respect to these other measuring tools.

124. In addition, financial sustainability poses a great challenge, both for the last period of ERCE and for the coming cycles of the studies. This requires an improvement in budget planning and in the financing model, which involves, among other things, a greater contribution from countries, as well as greater accountability. In this regard, the various initiatives proposed and implemented by the LLECE coordination team since 2019 are going in the right direction and respond to the challenges identified; however, they require prompt resolution which will make it possible to ensure the sustainability of the initiative.
125. Although LLECE has, in the end, achieved a process of maturity and learning in the management of assessment cycles, it is important that it move forward in transforming into a regional entity which not only has a main focus on assessment, but also focuses on supporting policies and generating capacities for improving countries' education systems. The 2030 Agenda 2030 affords an extraordinary opportunity for this, since UNESCO is the main organization responsible for follow-up of SDG 4, which relates to education, and since LLECE has the potential to become the main resource for its monitoring and assessment in the countries of Latin America and the Caribbean, as well as a referent for other developing countries.
126. Lastly, although it does not fall within the scope of this study, it is important that, as part of its strategic planning, the Laboratory analyse international changes and trends in the field of assessment, as well as the effects of COVID-19 on education systems and students' learning achievements, since both phenomena will affect the demand for assessment in Latin America and the Caribbean. Likewise, it is necessary to reflect on the possibility of incorporating other modalities and/or tools for the assessment of dimensions of the quality of education which are considered valuable, but which are not part of LLECE measurements, as well as tools which can be used by countries to obtain information on their students' progress more frequently.

VI. RECOMMENDATIONS

1. **Ensure short- and medium-term financial sustainability.** This entails advancing in the construction of a financing model based on the rule that the basal products and processes of the Laboratory's studies must be fully, or at least largely, funded by countries' dues, as is the case of the international studies reviewed. Given the differences in countries' financial capacities, it would be advisable to identify which countries find themselves in complicated circumstances; funding for these countries could be subsidized by a donor entity. Since such elements should be financed by the countries concerned, LLECE could also, through differentiated agreements based on individual countries' interests and needs, offer complementary products or advisory assistance in the form of national modules, regional sub-samples, more detailed analysis of results and policy proposals, or additional training, for example.

Furthermore, innovation and research projects should be financed with complementary resources from UNESCO or its related entities, or from multilateral organizations and/or international donors. This entails identifying specific projects of interest to donors and proactive management on the part of the LLECE coordination team and the CAE, as well as strong support from UNESCO Headquarters and other UNESCO entities, such as the UIS, with potential donors. In this context, raising the level of dialogue on LLECE's results and findings to the political sphere in the various countries would facilitate the securing of funds, the ensuring of government commitment, and the achievement of impact on policies.

Finally, this new model also requires improving the Laboratory's budget management and planning processes, as well as the delivery of more detailed information and greater accountability towards countries with regard to the use of resources.

2. **Design and implement a solid, coherent and systematic strategy for the dissemination of results.** This requires rethinking and planning strategies for dissemination and public discussion in much greater detail and from the beginning of each cycle. This dissemination plan should consider various strategies and products for disseminating results among the different target – 41 – audiences, namely, at

least: political and sectoral authorities, management officials, teachers, families, the media and researchers. However, the communication strategy must establish not only the target audience, but also the message which is to be communicated and the purpose of its communication. It requires producing an inspiring communicational account of the studies. In this regard, developing a solid dissemination strategy should include at least: i) identification of the key messages to be communicated; ii) development of shorter and more thematically focused reports; iii) use of different types of formats for the dissemination of findings, such as infographics, video briefs, webinars on specific topics, and seminars; iv) active use of social networks and improvement of the Laboratory website; v) development of alliances with other public and private foundations or institutions for the purpose of disseminating the study so as to facilitate the holding of, in different countries, seminars, webinars and forums aimed at publicizing the study and generating debate on the results.

However, the analytical plan must be sequential and systematic. It cannot conclude only with results-delivery reports (curricular analysis, results report and report on associated factors). This requires defining additional products related to dissemination and the establishment of evidence and resources for decision-making regarding public policy, as is systematically done by other assessment programs, such as PISA. This entails developing, in the years following the delivery of the aggregate and national results, reports specific to each discipline assessed or to specific topics (gender perspective in discipline assessments, for example), all of which is associated with great interest and value addition during the period in which countries' technical teams are preparing and implementing a new assessment cycle. These complementary efforts should be planned. The financial and human resources needed for their application should be ensured.

Certainly, this absolutely requires the investment of more resources in the dissemination plan. It is also essential that LLECE coordinators and OREALC/ UNESCO authorities devote more time to dissemination and positioning efforts in the participating countries, in other UNESCO spaces and institutions, and in other regional and world education forums. In the latter two cases, it is essential that the reports also be published in English and some in Portuguese, given that, thus far, they have only been published in Spanish, which severely limits their dissemination beyond the region. This change is strategic, not only because one of the largest participating countries is Portuguese-speaking, but also because dissemination in English would allow the Laboratory to expand its dialogue globally; it would also

facilitate dialogue with donors and partners in the development of education in developing countries.

3. Define the periodicity of the LLECE study cycle and its connection to the assessment of goals of the 2030 Agenda.

The results of the study show that there is broad consensus on the critical importance of establishing a clear and definitive frequency. This would allow both the LLECE coordination team and the countries to formulate a medium-term strategic plan which would prevent having to start from scratch every cycle; it would also enable better management of the various processes involved, since there would be clarity as to their stages and length. However, there is a challenging situation with regard to this; the General Assembly of the United Nations and the countries of the region have requested that two cycles of the study be conducted before 2030 in order to provide information on the achievement of the Sustainable Development Goals of the 2030 Agenda. Under current circumstances, this situation appears to be extremely complex because it would mean significantly reducing the length of earlier cycles to less than five years. This is already considered by some countries to be a short time for certain processes. Moreover, the Laboratory's previous efforts in this area have not achieved the time reduction anticipated. This is compounded by the effects which COVID-19 will have on the availability of countries' budgetary resources, as well as on the education sector's assessment and research agenda, a situation which will gradually be elucidated as the pandemic emergency progresses and the magnitude of its consequences are understood.

Thus, vital to addressing this challenge is the commitment of the countries and the Laboratory; in this regard, at least the following is necessary: i) the establishment and improvement of joint planning of both study cycles; ii) a formal commitment by the countries and the LLECE coordination team to strictly comply with the timetables for each stage, as well as a follow-up platform for each milestone, which will make it possible to monitor progress and delays; iii) the revision of all current processes of study implementation so as to identify possible ways of reducing time, such as sharing the curricular analysis between the two study cycles, for example (this would mean that an in-depth analysis takes place only in the first cycle, while in the second, this analysis is simply updated); iv) agreement, with countries' authorities, on a new stable and adequate financing model for the next two cycles of the study; and v) reaching of a formal commitment so that countries participate in the next two cycles of the study by the second half of 2021.

Given LLECE's enormous potential for monitoring and assessment of the SDG 4 indicators, as well as the progress made with other relevant international organizations, it is very important that UNESCO Headquarters and its related entities provide institutional support to OREALC and to the Laboratory much more decisively with regard to the identification of financing mechanisms for the implementation of an initiative such as the one previously mentioned, since this undertaking would be impracticable if it relied only on resources contributed by countries.

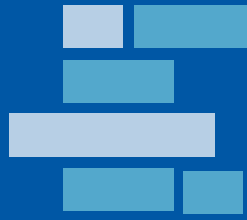
4. **Enhance the relevance and use of results for countries.** Key to this is ensuring that all countries manage to produce their national reports on both results and associated factors, as well as providing further guidance on the use of results for policy purposes and for improving pedagogical or management issues in schools. This involves providing more support to countries, especially those with less capacity for and experience with the preparation of national reports, by means of, for example, the design of a common template and the corresponding data. This would facilitate the development of national reports so that all that countries need to do is adapt or adjust the template according to their needs and their desired emphasis. In addition, the analysis of associated factors should give greater consideration to teaching practices, leadership practices, and schools' internal processes. Likewise, studies which are more qualitative could be conducted in order to develop and identify good practices and examples in the region which help to achieve the improvements expected. Finally, in order to make results more accessible to schools, a simpler, shorter, and more comprehensive report format could be established for school principals and teachers. It could be easily reproduced by each country; it would include some guidance on the consequences of the findings for teachers and principals, as well as on the implications and limitations of the results. In addition, guidance and initiatives could be developed in order to help countries to increase capabilities among local education authorities, school principals and teachers. These capabilities would relate to interpreting and employing the findings and results of the international and national studies, as well as to developing the assessment capacities of said authorities, principals and teachers themselves
5. **Promote the use of the results for academic research purposes.** To this end, we propose holding seminars with academics and researchers, holding assessment-technique training workshops, developing virtual workshops or webinars for

researchers, creating a fund for research (for which resources would have to be identified), and facilitating the publication of this research in a special edition of the journals of UNESCO or of other international organizations, such as the Organization of Ibero-American States for Education, Science and Culture (OEI). It would also be advisable to publicize initiatives being developed by some countries to enhance the usability of the Laboratory's data in national research. This, for example, is the case of the recent creation of the National Agenda for Educational Research (ANIE) in Ecuador, which brings together 27 public, academic and research organizations and organizations supporting educational development in the country; the ANIE promotes, coordinates and conducts quality research on the sector. This will facilitate the availability and usability of educational assessment information bases. In addition, ensuring that the publications and materials for data use are both easily accessible and issued in English is critical to encouraging academic and university centres of excellence worldwide to participate in results-related research and analysis. Finally, it would be worthwhile to create and manage, on the OREALC website, an archive containing all the publications associated with the studies, as this would also facilitate use for research and encourages new projects.

6. **Continue to strengthen the LLECE coordination team.** The study's findings show that it is absolutely necessary to expand the Laboratory's central coordination team of professionals in order to respond appropriately to countries' multiple needs, monitor the work of implementing partners, and address the challenges and tasks of which the coordination team is in charge. In particular, priority must be given to strengthening the team's specialized technical capability, as well as its communication and outreach capabilities. In addition, it is necessary to create working conditions and opportunities for professional development which enable greater team stability. This would require estimating the resources needed for financing in the coming study cycles.
7. **Facilitate the inclusion of more of the region's countries in LLECE.** Although there has been no explicit request from other Caribbean countries to join the Laboratory so far, LLECE should be attentive to other countries' interest in joining. This absolutely requires ensuring the financing and strengthening of the Laboratory's technical and management capabilities. In particular, LLECE could explore the international cooperation community's interest in supporting the inclusion of Haiti, the region's most vulnerable country, since its participation would facilitate the monitoring and assessment of the Sustainable Development Goals

of the 2030 Agenda and the designing of strategies for supporting processes for continuous improvement of its school system.

- 8. Broaden the focus of the Laboratory's work in the medium term, bringing it from the coordination of a regional assessment system to that of a system which also plays a role in capacity-building and educational improvement policies.** The many requests from participating countries and the suggestions from the experts show that there is an opportunity for LLECE to broaden its scope of action and to add increased value based on the very data collected. This involves progress in helping countries to lead educational improvement processes, establishing the generation of professional and technical capacities, the development of evidence-based policies, and the dissemination of best experiences. Moreover, it involves moving toward a more systemic understanding of the multiple sources of information available to the countries of the region. Unlike in the 1990s, most countries now have internal systems for the assessment of their education levels. As well, gradually, several countries have been participating in PISA, which focuses primarily on secondary education. This initiative entails the challenge of establishing differentiated governance structures for LLECE and implementing the regional study. It also requires much greater financial and institutional support from UNESCO Headquarters and other UNESCO entities, in order to generate mechanisms for dialogue with the highest educational authorities in each country, in addition to maintaining institutional mechanisms for technical participation on the part of the countries concerned. In this regard, the COVID-19 situation can provide an opportunity to broaden the scope for assessing the skills and knowledge acquired by students and schools; such broadening would be undertaken by coordinating the Laboratory's efforts with those of other international organizations and focusing on addressing the pandemic's effects on learning in LAC education systems and on working towards the achievement of the goals of the 2030 Agenda.



Annexes

A. Terms of reference

B. References



Annexes

A. Terms of reference

Background

1. The Latin American Laboratory for the Assessment of the Quality of Education (LLECE) or “Laboratorio.” was established in 1995 by UNESCO’s Regional Bureau of Education for Latin America and the Caribbean (OREALC/UNESCO Santiago). Its mission is to carry out regional large-scale student assessment studies.¹⁶ Since its inception, member countries support LLECE by providing a financial contribution and participating in the different phases and implementation of the study through their Ministry of Education’s evaluation offices or evaluation institutes. OREALC/UNESCO has a small technical team that coordinates the study and engages with technical and implementing partners involved in the different aspects of the study. Since 2011, LLECE is also supported by a High-level Technical Advisory Board (CTAN)¹⁷, that provides expert advice, strategic guidance and recommendations to inform technical decisions on issues such as sampling, test design, analysis of items, secondary analysis of associated factors, and reports. Recently, OREALC also established a Strategic Advisory Board for the Laboratorio, formed by a group of education experts and other strategic partners.¹⁸ The purpose of this Board is to advise OREALC on the strategic direction of the Laboratorio and on its financial sustainability for the coming years.
2. The Laboratorio emerged in the context of the 1990s, a period in which many countries in the region were adopting educational reforms without sufficient and relevant information for their design and lacking a critical mass of resources to measure education quality. In this period, only a few countries in the region were conducting educational assessments.¹⁹
3. As a framework for regional cooperation in educational quality assessment, the

¹⁶ Large-scale learning assessments (LSLAs) are a form of national or cross-national standardized testing that provide a snapshot of learning achievement for a group of learners in a given year and in a limited number of learning domains. *The promise of large-scale learning assessments, Acknowledging limits to unlock opportunities*, UNESCO 2019.

¹⁷ The CTAN consists of a group of world-renowned experts in educational assessment research such as Eugenio Gonzalez (ETS), Ralph Carstens (IEA), Wolfram Schulz (ACER) and Harvey Sanchez (Aleph).

¹⁸ Participants include people from the UNESCO Institute for Statistics (UIS), the Inter-American Development Bank (IDB), the World Bank and UNICEF

¹⁹ http://www.unesco.org/new/fileadmin/MULTIMEDIA/FIELD/Santiago/pdf/Kit_TERCE.pdf

Laboratorio has as an overall goal to measure student achievement and study the factors that relate to learning as a means to inform educational policy and, ultimately, the quality of education. The Laboratorio strives to reach this overall goal by:

- a. coordinating the regional comparative studies including curriculum analysis of participating countries, construction and piloting of test items, development of survey instruments, test application and data collection, data analysis and reporting.
 - b. developing education assessment capacities in Ministries of Education or Evaluation Institutes; and
 - c. serving as a space to generate ideas, disseminate advances and foster regional discussions around education quality.
4. For UNESCO, learning assessment refers to a wide range of methods and tools used to evaluate, measure and obtain valid, reliable data about the context, resources, processes and results of learning. ²⁰ It is about gathering information from multiple sources on what learners know and what they can do with what they have learnt. It helps compare results with set standards and objectives and to identify discrepancies and their possible causes.²¹ Learning assessments can serve diverse purposes and needs, among them: inform teaching and learning practices, provide evidence for policy development and programme design, and ensure accountability for results.
 5. Since 1995 the Laboratorio has produced three comparative and explanatory regional educational assessments (PERCE 1997, SERCE 2006 and TERCE 2013) A fourth one, the Regional Comparative and Explanatory Study (ERCE 2019) is currently ongoing. The regional assessments are based on the national learning goals of each country. Each edition of the study begins with a thorough analysis of the school curriculum of all participating countries in the areas and grade levels that are part of the assessment. This in turn helps establish a common conceptual framework on the type of questions asked of primary school students across the participating countries.

²⁰ https://en.unesco.org/system/files/private_documents/learning_assessments_brochure-eng-web.pdf

²¹ https://en.unesco.org/system/files/private_documents/learning_assessments_brochure-eng-web.pdf

6. The Laboratorio is the most important and most ambitious student performance assessment study in Latin America. The regional studies have sought to guide educational debate towards the challenges faced in the region while also providing information on the gap in the achievement levels between students in different contexts, highlighting situations of inequity.
7. The regional studies have been:
 - PERCE (1997)²². Obtained the first comparable data on learning achievement. Implemented in 13 countries²³ for students in the third and fourth grades in mathematics and reading. It represented a first step towards the creation of a culture of testing and accountability in Latin America.
 - SERCE (2006)²⁴. It focused on skills in mathematics, reading, writing and natural sciences²⁵ for students in the third and sixth grades. The SERCE covered 16 countries and the Mexican State of Nuevo León.²⁶
 - TERCE (2013)²⁷. The third edition included the participation of 15 countries and the Mexican state of Nuevo León.²⁸ It assessed the performance of students in third and sixth grades in mathematics, reading and writing, and natural sciences.²⁹ Through the questionnaires of associated factors, the study included an emphasis on the use of information and communication technology (ICT) and its association with student learning, and the implementation of national modules for those countries interested in analyzing issues of particular interest for their own realities.
 - ERCE (2019). The fourth study, currently ongoing, includes the participation of 18 countries and assesses the performance of students in third and sixth grade in mathematics, reading and writing, and science. Moreover, this version of the study includes a new module for the measurement of socioemotional skills among six graders, specifically focused on self-control and empathy.
8. The cycle of the ERCE is aligned with the three lines of work described in paragraph 3. The different phases of the study involve an active participation of the national teams located in the countries' agencies and ministries. In order to ensure that ERCE's implementation runs smoothly and complies with the quality standards expected of an international large-scale assessment, the Laboratorio collaborates closely with national teams. In the context of this collaboration, LLECE organizes capacity development workshops for members of national teams as a means to install capacities in national assessment systems.³⁰
9. The Internal Oversight Services (IOS) of UNESCO evaluated LLECE in 2008, as part of a larger evaluation on Measuring Learning Achievements. The LLECE section of the report evaluated UNESCO's support to the LLECE along nine dimensions.³¹ Some of the main findings from the evaluation were that: i. The Laboratorio had led well the implementation of the SERCE, integrating improvements and innovations based on learning from the PERCE process. ii. The introduction of a Technical Consultative Committee was positive. iii. The members of the national network assessed highly the support from OREALC although OREALC rated low the support from UNESCO headquarters. iv. There were differences among countries in the degree of support they required to build their assessment capacities.
10. The current operating costs of the LLECE amount to approximately US\$1 million/year partially financed through the contributions of the participating countries, which pay on a sliding scale an average of US\$ 19,000 per country per year. UNESCO headquarters does not allocate resources for LLECE, so mobilization of extrabudgetary resources from other donors is a main task of the coordination team. Donors have included the UNESCO Institute for Statistics (UIS), UNICEF, the Spanish Cooperation Agency (AECID) and the Fundación Santillana.
11. The Laboratorio is a mainstay of educational assessment in Latin America. This year will mark its 25th year of activity. Some countries have participated in all assessments and now have longitudinal data allowing appraisals going back almost 3 decades. Other countries have joined recently (for the fourth phase) and are in the process of generating a first batch of assessment results, such as the case of Bolivia.

²² First Regional Comparative and Explanatory Study.

²³ Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, Honduras, México, Paraguay, Perú, Dominican Republic, and Venezuela. There were a total of 54,589 test-takers in language and 54,417 in mathematics

²⁴ Second Regional Comparative and Explanatory Study

²⁵ The science assessment in all versions of ERCE is applied only on sixth graders.

²⁶ Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, the Dominican Republic, Uruguay and Venezuela. The Mexican state of Nuevo León also joined as a sub-national entity

²⁷ Third Regional Comparative and Explanatory Study

²⁸ Argentina, Brazil, Chile, Colombia, Costa Rica, Dominican Republic, Ecuador, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru and Uruguay

²⁹ TERCE had the largest number of students of the three studies. A total of 195,752 students participated; of which 100,752 were in third grade and 95,000 in were sixth graders.

³⁰ See Annex 2 for a list of the capacity development workshops organized in 2018.

³¹ (1) administration and management; (2) information; (3) communication; (4) technical; (5) material; (6) finances; (7) human resources development; (8) organizational development; and (9) legislation and regulations.

2030 Sustainable Development Agenda

12. SDG 4 promotes inclusive and equitable quality education. Specifically, Target 4.1 looks to ensure that, by 2030, “all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes.” This target is monitored by indicator 4.1.1 to which LLECE studies provide data for the Latin American Region.
13. UNESCO Santiago is responsible for leading and monitoring the progress of the implementation of the roadmap designed for Latin America and the Caribbean towards the achievement in the region of the SDG4, detailed in the Education 2030 Agenda. High-level government officials at a Regional Meeting of Ministers of Education of Latin America and the Caribbean recognized this role in 2017.³²
14. Large-scale learning assessments can provide data to better understand if students are learning, and if learning is equitable across all groups in society, which is key data for reporting on indicator 4.1.1. Learning assessments offer insights into progress made and challenges still present.
15. In Latin America, ERCE is the most representative learning assessment in primary and it is the only tool to provide comparative data across countries in these grades. In its fourth cycle, ERCE will provide data for indicators 4.1.1.a and 4.1.1.b for 18 countries.
16. In addition, the Laboratorio is currently developing a project with the International Association for the Evaluation of Educational Achievement (IEA) called “Rosetta Stone,” with the objective of generating comparative data between ERCE and IEA’s studies.³³ This will allow for a considerable increase in the number of countries in the world that have comparable data for indicator 4.1.1.

³² E2030: Education and skills for the 21st century, Buenos Aires, Argentina, January 24 – 25, 2017.

³³ The name of the project refers, literally, to a conversion table. In this project, the table will allow to “translate” the scores of ERCE into scores of IEA’s two studies, the Trends in International Mathematics and Science Study (TIMSS) and the Progress in International Reading Literacy Study (PIRLS) or vice versa. This project is being funded by Bill & Melinda Gates Foundation, thanks to the efforts made by UIS.

Rationale for the Evaluation

17. For the past 25 years, the Laboratorio has been leading the implementation of a large-scale student assessment that covers the majority of the countries of Latin America. Notwithstanding, the Laboratorio has not been the subject of a dedicated evaluation since 2009³⁴. The evaluation will allow taking stock of its contribution and role in the area of quality education in the Latin American region and highlight future opportunities and challenges. The evaluation will be an input into strategic decisions and inform future actions for the implementation of the Laboratorio’s work.

Purpose and Scope

Objectives

18. The overall objective of the evaluation is to assess and better understand the value or worth that participating countries bestow on the LLECE.
19. The evaluation will also assess the contributions of the Laboratorio in terms of (i) providing reliable data on student learning and its explanatory factors; (ii) providing information to guide policy decisions at the national level; (iii) promoting a regional dialogue on curriculum and quality education; (iv) building capacities in national teams on student assessment
20. In addition, the evaluation will:
 - Assess the Laboratorio’s operation during the past 25 years, analyzing its evolution and identifying its strengths and weaknesses.
 - Examine the Laboratorio’s alignment with UNESCO’s goals and mission, exploring synergies with different programs within UNESCO.
 - Assess the extent to which its activities and purpose contribute to the accomplishment of the 2030 Agenda’s SDG 4 in general, and target 4.1 in particular.
 - Examine LLECE’s business model, funding mechanisms and country and headquarters contributions.
 - Review LLECE’s contribution to the global reporting of indicators of SDG4 including its recent development of the collaboration project with IEA.

³⁴ In 2008 an external evaluation was conducted by the University of Colorado and by the Universidad de Chile in 2009.

Use

21. The evaluation will provide evidence to a number of stakeholder groups. Among them is the recently formed Strategic Advisory Group, supporting strategic direction to the LLECE. Other potential users of the evaluation are the Education Sector in UNESCO Headquarters, UNESCO's Executive Board, national technical teams and OREALC itself. The evaluation will provide different outputs (evaluation report, summaries, presentations) based on each audience's specific needs.

Scope and Evaluation questions

22. The evaluation will be both retrospective and prospective, providing a synopsis of the performance of the Laboratorio in the past 25 years and recommendations on improvements to best achieve the Laboratorio's objectives in the future. (Considering the implementation of the four ERCE studies, PERCE, SERCE, TERCE and the ongoing ERCE 2019)
23. The evaluation will not focus on the effectiveness of individual countries in implementing the national assessments nor is it a technical evaluation on the soundness of the LLECE's educational assessment methodology or tools.
24. Key evaluation questions could include the following

Relevance

- a. To what extent is the Laboratorio known, valued and trusted by the countries in the region as a useful tool to strengthen their assessment systems and gather information to guide their policies? (perceptions of value from countries)
- b. To what extent is the Laboratorio an educational assessment process that is adequate and responds to the needs of the participating countries? (has responded to actual needs)
- c. To what extent is the Laboratorio consistent with the priorities established in the 2030 Education Agenda? (aligned with the UN Agenda)
- d. To what extent has the LLECE integrated considerations of gender and other vulnerable and minority groups that are a priority for participating countries? (aligned with gender priority)

Effectiveness

- e. What have been the most salient and overarching results of the LLECE?
- f. To what extent have the outputs from Laboratorio (the three regional and comparative studies) influenced educational programming and/or policy decisions? (effective in influencing policy and decisions)
- g. To what extent has the Laboratorio developed capabilities in the national evaluation teams of participating countries?
- h. To what extent has the LLECE effectively supported the development of spaces for new ideas, disseminating advances and fostering discussion on quality of education? (effectiveness in advancing the agenda and discussion on quality education)
- i. Has the Laboratorio been effective in its technical role of supporting and coordinating the development of ERCE items including instruments, educational assessment tests and associated factors questionnaires?
- j. Has the Laboratorio been effective in its technical role of supporting and coordinating the gathering of empirical data?
- k. How well has the LLECE encouraged South-South cooperation?

Efficiency

- l. Is the current configuration of the LLECE in terms of financial and human resources adequate for supporting their stated goals?
- m. Is the OREALC CT team based in Santiago receiving adequate support from UNESCO Headquarters?
- n. How does LLECE fit within a greater UNESCO work programme?
- o. How does the LLECE business model and costs compare with that of other large-scale assessment programs?

Sustainability

- p. To what extent are the current financial and organizational configurations of the Laboratorio sustainable for its continuity as a regional tool and its implementation of future versions of the ERCE?
- q. What are the main improvements needed to ensure the Laboratorio's continuity as a regional tool and the accomplishment of future versions of the ERCE?

Methodology

25. The evaluation approach will require a combination of multiple and complementary evaluative methods and strategies collecting both quantitative and qualitative data. The final methodology may include some or all of the methodological elements below. The specific methods will be further refined during the inception phase and in consultation with the evaluation reference group and IOS.
26. Desk review. The evaluation team will need to review a wide array of documents produced throughout the life of the LLECE. These include relevant budget and programme documents, minutes of meetings of the National coordinators meetings; previous reviews and evaluations and applicable Executive Board documents. It will also need to review relevant elements of the PERCE, SERCE and TERCE as well as any documentation produced by the High-level Consultative Technical Council (CTAN). An indicative list of documents is provided at the end of the ToR (Annex 1), however the evaluation team is expected to exercise due diligence in canvassing the relevant literature.
27. Reconstruction / refinement of a Theory of Change of the UNESCO's technical support to LLECE including the results chain and its underlying assumptions.
28. Semi-structured interviews (face-to-face and via Skype) with individuals from the different stakeholder groups. The evaluation team, supported by the evaluation reference group will develop a more precise list of these stakeholders during the inception phase. Some of these groups are:
 - LLECE's technical staff (past and present)
 - education sector staff in UNESCO headquarters and institutes
 - education officials and present and former national coordinators in the participating countries
 - members of the CTAN
 - members of the Strategic Board
 - implementing partners such as the Centro de Medición of Chile's Pontificia Universidad Católica, the Center for Comparative Education Policies of the Universidad Diego Portales³⁵ and the Center of Advance Investigation in ducation (CIAE) of the Universidad de Chile.³⁶

³⁵ The former has been responsible for test development and analysis; the latter for development and analysis of the associated factor questionnaires

³⁶ The former has been responsible for test development and analysis; the latter for development and analysis of the associated factor questionnaires

- funding agencies such as UNICEF, the Inter-American Development Bank (IADB), the Spanish Agency for International Development Cooperation (AECID), the Santillana Foundation.
 - academics and other experts involved in the area of educational assessment or working in assessment institutions, such as IEA, ETS or ACER.
29. Case studies. In consultation with the evaluation team, two case study countries will be selected. The case study will entail an in-depth examination on the modalities of Laboratorio support in the assessment process of that country. It will include a visit to the country and interviews with officials, as well as, if possible, with teachers and school officials from a subset of the schools that were part of an assessment. A possible selection rationale would be to select a country that has taken part in all assessments and a country that has more recently become part of the network.
 30. Benchmarking with other regional large-scale assessments. This will involve comparing selected business processes and performance metrics of LLECE to a selected large scale assessment. The benchmark could be other regional assessments or global exercises such as OECD's PISA, IEA's TIMSS and PIRLS and others³⁷
 31. Questionnaire(s) and/or survey(s) LLECE country representatives (national coordinators).
 32. Participatory workshop to steer the evaluation and to discuss preliminary findings, lessons learned and recommendations.
 33. Data collection, sampling and analysis must incorporate a gender equality perspective, be based on a human rights based approach, and take into consideration the diverse cultural contexts in which the activities are being implemented.

Roles and Responsibilities

34. The evaluation will be managed by UNESCO's Internal Oversight Service (IOS) in close collaboration with OREALC.
35. The evaluator or evaluation team will need to have specific expertise in the substantive areas of evaluation and educational assessment and be based in the Latin American region.
36. OREALC will identify and establish an Evaluation Reference Group. The group the evaluation process and ensure the quality of associated deliverables.

³⁷ PISA is the OECD's Programme for International Student Assessment. Every three years it tests 15-year-old students from all over the world in reading, mathematics and science.

37. The external evaluation consultant (or team) will conduct data collection and fieldwork and be responsible for the analysis of the data and the preparation of an inception report, a draft and a final report (in English and in Spanish). The evaluator will also prepare communication outputs such as infographics, powerpoints, or 2-3 page summaries for different audiences.
38. The evaluators will comply with United Nations Evaluation Group (UNEG) updated 2016 Norms and Standards for Evaluation, UNEG Guidelines for Integrating Human Rights and Gender Equality in Evaluations and UNEG Ethical Guidelines for Evaluation.
39. The evaluation team will be in charge for their own logistics: office space, administrative and secretarial support, telecommunications, printing of documentation, travel, etc.

Qualifications of the Evaluation Consultant

40. The external evaluation team leader is expected to possess the following mandatory qualifications and experience:
 - University degree at Masters level or equivalent in education, social sciences, political sciences, economics, or related field.
 - At least 10 years of project and/or programme evaluation at the international level or in an international setting; particularly of educational projects and programmes.
 - Proven experience, of at least 5 years, with large-scale educational assessments and in three different projects or institutions of evaluation.
 - Senior experience of at least 10 years in knowledge of and experience in applying qualitative and quantitative data collection methods, including case studies.
 - Understanding and knowledge of the UN mandates and its programming in relation to education in the framework of the Sustainable Development Agenda.
 - Bilingual in English and Spanish. Ability to read, interview and write a report in either language to the same level of proficiency.
 - No previous involvement in the implementation of the activities under review.
 - Experience with assignments for the UN.

Deliverables and Schedule

41. The evaluation will commence in September 2019 and conclude in January 2020.
42. The evaluation will consist of four main deliverables: inception report, draft report, final report and summary communication outputs.
43. Initial meeting. The consultants will participate in an initial meeting in Santiago where they will meet with Senior Management from OREALC and further clarify the evaluation purpose, the questions the evaluation seeks to answer and how the clients foresees using the evaluation results. It will also be an opportunity to meet and interview the Technical Coordination Team (CT) and the partners from the Pontificia Universidad de Chile, Universidad Diego Portales and Universidad de Chile.
44. Visit to the Meeting of National Coordinators in Antigua Guatemala. This will be an opportunity for the evaluation team to meet and interview all national coordinators, one of the main data collection elements.
45. Visit to case study locations. The evaluation will integrate two countries as case studies. The evaluation consultant(s) will visit the countries and carry out key informant interviews with relevant officials involved both at the technical and political levels of the assessment. OREALC, IOS and the evaluation reference group will define selection criteria, consult and identify the case study locations.
46. Towards the final phases of the data collection and analysis, OREALC will coordinate a stakeholder workshop in Santiago where the consultant(s) will present initial findings for discussion.
47. Main deliverables:
 1. **Inception report:** After initial discussions with OREALC, IOS and selected members of the reference group, the consultant/ team will prepare an inception report. This consists of a report with a theory of change and any adjustments to the questions or methods presented and timeline presented in the ToR. The inception report shall not exceed 10 pages.

2. **Draft evaluation report:** The evaluation team will prepare a draft evaluation report. It will be circulated for comments among the evaluation reference group. IOS or OREALC will consolidate all comments for the evaluation team. The evaluation report will be drafted in English and translated to Spanish (or vice versa) and follow UNESCO IOS's Evaluation Report Guidelines. IOS will share with the evaluation team the guidelines and a detailed final report template in due course. The main body of the draft report shall not exceed 30 pages, excluding annexes.
3. **Communication output(s):** The evaluation team will prepare a synthesis of the main findings from the evaluation. This might take different formats such as a power point presentation or slides, a 2-3 page brief or an infographic.
4. **Final evaluation report:** The final evaluation report will follow the aforementioned structure. As part of the UNESCO IOS quality assurance processes, all evaluation reports are subject to review by an external expert to ensure compliance with quality standards (if it is corporate). The recommended actions from the quality assurance process will be addressed prior to finalization of the report.

References

- UNESCO Santiago, What affects learning achievements among migrant children? TERCE In Sight, June 2016
- UNESCO, The promise of large-scale learning assessments Acknowledging limits to unlock opportunities, 2019
- Daniel Wagner, Smaller, Quicker, Cheaper Improving Learning Assessments for Developing Countries, IIEP, 2011.
- UNESCO Santiago, Comparison of Results between the Second and the Third Regional Comparative and Explanatory Studies.
- UNESCO Santiago, Third and sixth grade teachers in Latin America and the Caribbean: Characteristics, perceptions and relationship with student learning achievement Resultados Nacionales del Tercer Estudio Regional Comparativo y Explicativo TERCE 2013, Mexico INEE 2016
- UNESCO, Learning assessments at UNESCO, Ensuring effective and relevant learning for all, 2018
- TERCE background, Preguntas frecuentes.
- University of Colorado, UNESCO's Technical Assistance and Capacity Building Efforts in Measuring Learning, 2008
- Universidad de Chile (2009) Evaluación Del Segundo Estudio Regional Comparativo Y Explicativo (Serce) Desarrollado Por El Laboratorio Latinoamericano De Evaluación De La Calidad De La Educación (LLECE) OREALC/UNESCO Santiago.

List of selected LLECE activities, workshops and meetings in 2018

Preliminary versions of the ERCE 2019 instruments were used to test students' learning achievement in mathematics and science.	January 2018	
Meeting of the High Level Technical Advisory Council	February 2018	Washington DC
Workshop on data management and procedures for the ERCE field trial study	February 2018	Dominican Republic
Training delivered to specialists and technicians from the Northern calendar countries		
National Coordinators meeting	March 2018,	Mexico City
Linguistic adaptation process completed	April 2018	
Adaptation of instruments to the national contexts of southern calendar countries initiated.	April 2018	
Participation in the Latin American Congress of Educational Measurement and Evaluation COLMEE	May 2018	Uruguay
Cuba pilots ERCE	May 2018	
Colombia, Guatemala, Panama and Argentina complete the process of quality control of the instruments	May 2018	
Workshop on Data analysis of large-scale evaluations	May 2018	Bolivia
Workshop on data management and procedures for the ERCE field trial study and coding on	June 2018	Ecuador
40th Meeting of National Coordinators	September 2018	Peru
Workshop on the New Vision of Curriculum Analysis in Science and its Challenges	October 2018	Panama
Workshop on Data Analysis	November 2018	Argentina

B. References

Bruns, B.; Akmal, M., y Birsdall, N. (2019). The political economy of testing in Latin American and Sub-Saharan Africa. CGD Working Paper 515. Washington, DC: Center for Global Development.

IEA (2018). Presentación Oliver Neeuschmidt sobre Proyecto Piedra Rosetta.

LLECE/OREALC (2019): Minuta de estimaciones preliminares del presupuesto LLECE 2016-2021

LLECE/OREALC (2019a). Antecedentes históricos del estado financiero del LLECE. Minuta, Santiago, Chile.

LLECE/OREALC (2019). Situación del Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación. Respecto a la Implementación del ERCE en el Período 2016-2019. Minuta interna, OREALC/UNESCO, Santiago, Chile.

LLECE/OREALC (2019). Acta reunión Consejo Asesor Estratégico del LLECE, OREALC/UNESCO. Ciudad de México, septiembre, 2019.

LLECE/OREALC (2019). Acta XLII Reunión de Coordinadores Nacionales del Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación. Antigua, Guatemala, octubre, 2019.

LLECE/OREALC (2019). Situación del Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación. Respecto a la Implementación del ERCE en el Período 2016-2019. Minuta interna, OREALC/UNESCO, Santiago, Chile.

Naciones Unidas (2019). Informe de los Objetivos de Desarrollo Sostenible en 2019. Naciones Unidas, Nueva York, Estados Unidos.

Neeuschmidt, O. (s/f). Presentación Oliver Neeuschmidt y "El proyecto de vinculación Rosetta. Midiendo el progreso global hacia el ODS 4.1".

OECD (2019). Meeting expectations? Reflections on PISA for Development from participating countries. PISA for Development Brief Number 29.

OECD (2018). Peer to peer learning to strengthen dissemination of PISA for Development results. PISA for Development Brief Number 27.

OECD (2017). Who pays for the PISA for development? PISA for Development Brief Number 20.

OECD (2016). How is PISA for Development being implemented? PISA for Development Brief N 4.

Solano-Flores, G. y Bonk, W. (2008). UNESCO's technical assistance and capacity building efforts in measuring learning achievement: An evaluation of LLECE. University of Colorado at Boulder.

UNESCO/OREALC y BID (2017). Consejo Técnico Consultivo de Alto Nivel. LLECE. Roles y Funciones del CTAN.

UNESCO-OREALC (2015). Factores Asociados. Tercer Estudio Regional Comparativo y Explicativo. TERCE. Santiago, Chile.

UNESCO-OREALC (2016). Reporte Técnico. Tercer Estudio Regional Comparativo y Explicativo. TERCE. Santiago, Chile.

UNESCO (2019). Términos de Referencia. Evaluation of UNESCO's Latin American Laboratory for Assessment of the Quality of Education (LLECE)

UNESCO (2019). La promesa de las evaluaciones de aprendizaje a gran escala: Reconocer los límites para generar oportunidades.

UNESCO (2019). El proyecto vinculación Rosetta. Midiendo el progreso global hacia el ODS 4.1. Manuscrito.

UNESCO (2018). El proyecto de vinculación Rosetta. Midiendo el progreso global hacia el ODS 4.1 (Minuta).

UNESCO/OREALC (2016). Reporte Técnico. Tercer Estudio Regional Comparativo y Explicativo. TERCE. Santiago, Chile.

UNESCO/OREALC (2015). Factores Asociados. Tercer Estudio Regional Comparativo y Explicativo. TERCE. Santiago, Chile.

UNESCO/OREALC y BID (2017). Consejo Técnico Consultivo de Alto Nivel. LLECE. Roles y Funciones del CTAN.

Universidad de Chile (2009). Informe evaluación del segundo estudio regional compartido y explicativo (SERCE) desarrollado por LLECE/OREALC, Santiago, Chile.

<https://timss2015.org/timss-2015/about-timss-2015/>

<https://International study fees PIRLS 2021.pdf>

<https://International study fees TMSS 2019.pdf>



**INTERNAL
OVERSIGHT
SERVICE**

Evaluation Office