Executive Summary

Mexico, as the world's 14th largest economy (2009), faces important challenges in education. Despite the significant progress of the past decades in terms of access to education, improvements in completion rates for lower education levels and development of learning assessments, considerable improvement is still needed. Mexico already invests a high percentage of the public budget in education (at nearly 22%, it is the highest among OECD countries). Results from the 2009 round of the Programme for International Student Assessment (PISA) have shown that although improvement is possible in a relatively short period of time, important challenges remain. In addition to improving the quality of educational services, increasing attainment levels and reducing drop-out rates are also priorities. It is equally important, however, for Mexico to ensure that all students, including those from disadvantaged socio-economic backgrounds and indigenous families, have equal educational opportunities.

To address these issues, in 2008, the Mexican government and the OECD established the Co-operation Agreement "Improving Education in Mexican Schools". The purpose of the Agreement was to determine not only *what* policy changes to consider in Mexico, but also *how* to design and implement policy reforms effectively, given local conditions, constraints and opportunities. One of the strands of this Agreement has focused on developing appropriate policies to evaluate the quality of schools and teachers, particularly assessments, and to link learning outcomes to incentives for continuous improvement. This part of the work has been led by a group of international experts forming the OECD Steering Group on Evaluation and Teacher Incentive Policies in Mexico.

This summary report presents the main findings and policy recommendations developed by the Steering Group and the OECD Secretariat over the course of the Co-operation Agreement.[1] It draws on the results of international workshops and technical meetings with stakeholders in the Mexican education system, field visits, thematic reports from invited experts, and the stock of OECD research and knowledge. Since no single model of education reform can serve to guide all of the reform efforts in Mexico, the recommendations draw on experiences from over 20 countries.

OPPORTUNITY FOR EDUCATION REFORM IN MEXICO

The Mexican government established policy priorities for education reforms in its Education Sector Programme 2007-12. To monitor progress towards achieving its objectives, the Mexican Ministry of Education (SEP) established improvement indicators for student achievement as measured by the national ENLACE assessment and PISA. Other key indicators relate to teachers' professional development, school empowerment, equity in educational opportunities, and reforms relating to content and curriculum. To facilitate policy reforms, in 2008, the Mexican government established the Alliance for the Quality of Education with the national teachers' union (SNTE), which helped define the thematic focus of the Co-operation Agreement with the OECD. In this context, the following recommendations and considerations aim to provide SEP and relevant stakeholders with guidance on the policy priorities for a lasting and effective reform process.

1 **The public policy framework for implementing education reforms:** For countries and education systems to adapt and implement policy reforms tailored from best practices and international examples, local conditions, constraints and opportunities must be adequately addressed. When combined with international practices and comparable evidence, local knowledge mobilisation can provide a vital link in adapting best practices for effective education reforms suited to national priorities and contexts. The purpose of the *public policy framework for education reforms* presented in this report is to provide relevant stakeholders in Mexico with guidelines for continued local knowledge mobilisation to inform current and future reform processes.

1.1 In combination with international practices and available research evidence, *country-specific knowledge mobilisation* on particular policy issues is a vital element to effectively design, plan and implement educational reforms that are viable and sustainable given the conditions, constraints and opportunities in Mexico. Experience clearly suggests that reliable and up-to-date knowledge about particular policy topics is crucial in the process of adopting best practices and policy recommendations.

1.2 Current and future education reform efforts in Mexico would benefit from a methodical consideration of each of the following *six dimensions of the public policy framework* for lasting and successful education reforms:

i) **Data, information and indicators:** This implies a consideration of the quality and quantity of relevant data and information available (on students and teachers, schools, performance, and linkages between them), for target-setting and to identify deficient areas to be addressed.

iI) **Social relevance and stakeholder engagement:** This includes considering strategy options for communication, engagement and consultation with primary stakeholders, including the general public, teachers, principals and local educational authorities. It is important to identify how the proposed reform can be translated into a socially relevant and meaningful message for the average family, teachers and principals.

iii) **Public funding:** It is important to consider amounts and consistency of public funding for development and implementation of the policy reform (*e.g.* whether it is annual or fixed in the budget), including potential cost-benefit analysis, cost projections and economies that can be obtained by re-channelling existing budget items or programmes.

iv) **Institutional arrangement:** This includes a consideration of public institutions (central and state educational authorities), to identify specific bodies that should contribute to developing standards, evaluations, and proposing modifications.

v) **Legal and regulatory framework:** It is important to foresee potential conflicts and possible modifications that may be necessary in education laws and related areas (*e.g.* labour laws) to carry out education reforms.

vi) **Decentralisation and devolution process:** This includes looking at *formal* and *de facto* levels of devolution across the main federal bodies and state educational authorities responsible for providing educational services (including resources, capacity, information management, evaluation and supervision).

2 **Public accountability:** All stakeholders should feel responsible and be held publicly accountable for student learning and overall educational results.

2.1 *Performance, equity and value for investment in education are challenges for Mexico,* as in many other countries undertaking important educational reforms. This is illustrated by Mexico's performance in international comparisons, in the great diversity that exists between and within Mexican regions and states, and in the importance that education spending continues to have in terms of share of the public budget, despite modest per-pupil spending compared to other OECD countries. Holding all actors involved in Mexico's education system accountable for increasing the performance of all students, in all schools, provides a clear message and a way to align efforts and resources.

2.2 *Actors should be held accountable for student learning and growth, and provided with the necessary assistance and capacity building.* A clearly defined accountability system focused on the results of student learning and growth can provide the necessary coherence, given the size, complexity and multiple interests of the participants in Mexico's education sector. The use of student learning as a key criterion against which state education authorities, schools, principals and teachers will be held accountable, reflects a focus on outcomes rather than input-focused policy reforms. International practice regarding performance-based teacher incentives, for example, reflects this change. This does not imply, however, that issues of infrastructure or social inclusion are no longer important for the Mexican educational system. Rather it implies that learning and development for all students – fostered, cultivated, assessed and evaluated through various means – should be the ultimate goal of policy action and reforms. Support to students, schools, principals and teachers, as well as professional development, are vital complements to increased accountability.[2]

2.3 *Accountability focused on student learning and growth implies establishing clear standards.* The development of standards as a key component of the accountability system focused on student learning should address at least three priorities: *i)* appropriate development of standards for content, student performance and teacher performance; *ii)* alignment and coherence between standards, assessment, evaluation and professional development; and *iii)* alignment of standards to international best practice and internationally competitive benchmarks of student knowledge and skills. Within a standards-based accountability framework, actors should have incentives to meet or exceed the expectations that are reflected in standards.

2.4 *Accountability measures should include complementary criteria of effort as well as performance.* A standards-based accountability system for students, schools and teachers in Mexico should consider using measures of student learning and growth (from standardised assessments and other reliable methods, where possible), as well as complementary criteria regarding individual, group and school performance. This is important in Mexico as student and teacher attendance, punctuality and time-on-task remain important issues. An accountability system in Mexico should take into account the fact that some principals and teachers may not be performing to their current capacities. Incentives are needed, therefore, to increase basic effort and performance, as well as supporting capacity building and professional development. Reduction of student drop-out rates, for example, can also be considered as an indicator. Accountability also implies that some teachers who receive adequate technical assistance and opportunities for professional development, but who do not improve performance, would be counselled out of the profession.

2.5 *The focus should be on students, schools and teachers for continuous improvement with the school as a basic unit of accountability.*[3] Although different levels and actors in the education system should be held accountable, the school can serve as the basic unit of accountability, with individual data, information and monitoring for students and teachers. Student and teacher data and information at the school level can be used to support improvement efforts, teacher incentives and stimuli, education interventions for low performers, and the identification of good practices for modelling and to inform the development of teaching standards, for example.

2.6 *It is important to define a gradual process to develop complementary approaches of assessment using multiple sources of evidence.* Developing a robust standards-based accountability system is a gradual process, with clear stages, and with complementary approaches to assessment and evaluation. Both summative and formative assessments of student learning and growth, as well as school and teacher performance, should form part of the accountability system in Mexico. The development of such a system, however, should be delineated in stages with a thorough consideration of current and projected capacities, methods and costs.

3	**Importance of student learning outcomes:** Student learning and growth over time should be a key criterion to gauge the performance of schools, teachers, parent participation bodies, state and federal educational institutions, and the system as a whole. Results from standardised assessments are important, but other reliable and valid measures of student learning should be employed for a fuller picture of student achievement.
3.1	*Student learning and growth as the basis of accountability and standards require multiple, cross-referenced, valid and reliable measures.* All of the current measures and instruments of student learning and growth (teacher assessments, portfolios of student work, classroom observation, and standardised tests, among others) present potential sources of error and bias. A complementary approach that uses valid evidence from multiple sources should be gradually developed to take into account current instruments in Mexico, estimate costs, and determine the capacity building and instrument development that are required. With clear content and performance standards of what students are expected to know and know how to do, measures and procedures to assess the learning and improvement expected from students can be further developed.
3.2	*The use of student performance data should be accompanied, when possible, with complementary and reliable measures of student learning, as these are developed, tested and validated.* The relative importance of student data and school-based or teacher assessments can be redefined over time. The state of Victoria in Australia, Hong Kong-China, and Canada are examples of better performing systems that combine standardised assessments with school-based assessments (*e.g.* locally graded but externally moderated), student projects and extended papers.
3.3	*Student performance data, such as those from the annual ENLACE assessment in Mexico, can play an important role in accountability and school improvement efforts.* Current efforts by SEP and state educational authorities regarding the presentation and use of ENLACE demonstrate the high degree of social acceptance and potential of ENLACE. Student performance data aggregated at the group, school, zone or state levels can be employed in static, improvement or growth models.
3.4	*The ENLACE assessment in Mexico has shown to be a valid and reliable measure of student achievement.* This provides Mexico with a valuable opportunity to exploit the potential of the student performance data provided annually by the assessment.
3.5	*A specific development programme should be established for the ENLACE assessment, considering issues of cognitive demand, curricular alignment and coherence.* The best-available evidence on student learning progression and standards should be considered. The development of ENLACE should also set clear stages and goals that address technical (*e.g.* vertical equating), administrative (*e.g.* unique student, teacher and school identifiers and linkages) and logistical (*e.g.* improved test supervision) considerations.[4] With expanded use of the ENLACE assessment in the future, enhanced supervision and security of test administration, for example, should be addressed. The programme should also have a long-term vision that takes internationally benchmarked content and performance standards into account. As content and performance standards are established in Mexico, student performance data can be used, in conjunction with analytical models (*e.g.* growth) for specific policy objectives and programmes. Throughout the process, consideration should be given to the alignment and coherence between standards, assessment and professional development for teachers. A clear vision of the evaluation framework in Mexico should allow for the distinct but complementary purposes of different assessments (*i.e.* ENLACE, EXCALE or possible school-based assessments), and how they should continue to develop in the future within a common national framework.
3.6	*With student performance data and appropriate growth models, low performers, high performers and cases needing follow-up observation can be identified.* As the assessment and evaluation process becomes more established, consequences such as incentives, further observation, and assistance to schools and teachers can be linked to the results. This implies the possibility of having multi-stage consequences and responses to the results. Schools determined to be repeatedly underperforming or performing near the top, for example, could be subject to on-site visits and reviews to identify potential causes and determine appropriate responses relating to improvement, technical support and the channelling of additional resources to under-performers.

4	**Fair assessment of the value-added of schools:** All students, regardless of socio-economic, ethnic or linguistic background, should have the same opportunities to learn and achieve at higher levels. Although student performance has been shown to be highly correlated with family background, results from assessments and evaluations should reflect the true contribution to students' learning and not the socio-economic context of the school or its students.
4.1	*Given the large diversity of educational contexts in Mexico, value-added models can offer a fair and more accurate measure of student growth and school performance.*[5] Current efforts by SEP and state educational authorities regarding the presentation and use of ENLACE results are a good starting point and could be built upon with value-added results for schools. The challenges involved in designing, planning and implementing an assessment system for accountability and school improvement that uses value-added modelling should be addressed rigorously throughout all stages, including the initial knowledge mobilisation, analysis and application phase of education reforms in Mexico.
4.2	*Value-added models can offer a better option than raw test-scores to accurately and fairly identify the contribution of schools to student learning, by taking into account the context and background of the students.* The technical challenges involved with developing an assessment and evaluation framework based on value-added modelling should be considered and addressed from the initial phases of design and planning. The robustness and frequency of the ENLACE assessment in Mexico, however, provides an invaluable opportunity.
4.3	*Given the current conditions of the educational system in Mexico, value-added models can be based primarily on the school as the unit of accountability,* although school zones,[6] student groups, municipalities and states can all be used for analysis and action. Vertical equating should be among the first of the technical issues to be reviewed in the further development of ENLACE. The quality and availability of information that could be used for contextualised value-added models should also be assessed.
4.4	*The first phases of the development of value-added modelling in Mexico should use actual student data* to identify the weaknesses and strengths of different value-added models. Even before applying value-added methods to student performance data, however, schools could be grouped based on socio-economic contexts, and contextualised attainment models could be used as possible precursors of full-fledged value-added analysis. Therefore, the process of establishing value-added modelling can have different phases:
	i) Stratification of similar schools (based on type and socio-economic or other relevant information) for within-group comparisons of average results of raw scores. Issues regarding quality and completeness of test data and contextual information should also be identified and addressed.
	ii) Internal value-added modelling exercises conducted by education authorities to select models and address technical issues with data. A three-year-moving average is suggested for the modelling. In addition, education authorities could use VAM analysis to monitor and conduct evaluation trials of specific policies, programmes and jurisdictions, such as *Programa Escuelas de Calidad,* for example, with particular emphasis on differences within and between municipalities, school zones, states and ethnic groups, among others.
	iii) Public information, awareness and engagement with stakeholders on the merits, challenges and opportunities of value-added modelling, which could be linked to a re-launching of the ENLACE assessment, for example, with a clear plan for its further development.
	iv) Attributing consequences (low-stakes at first for under-performing schools (further exploration, observation and assistance), as well as for high performers. The same value-added analyses could be used by SEP and state education authorities to identify schools that may have teachers and practices worthy of replication and modelling. Logistical issues relating to test administration should also be addressed.

5 | **Evaluation of teachers for accountability and improvement:** Teachers are vital to student learning. It is difficult to improve, however, what is not measured. A fair and reliable in-service teacher evaluation process should provide incentives for teachers at all levels of the performance spectrum to improve, to be recognised and to contribute to overall educational results.

5.1 | *Teacher standards should be developed to provide teachers with clear guidance as to what is considered good teaching practice.* Teaching standards could also be used in designing opportunities for professional development and improvement (training, modelling, observation, technical support, etc.). International examples and models of standards provide a useful starting point to further adapt and develop appropriate teaching standards in Mexico.[7] The standards developed by Mexico should meet the following criteria:

i) cover all of the defined teaching domains;

ii) establish different levels of competency for each specific aspect that defines the domains of teacher and school work;

iii) reflect a core group of performance traits that should be observable in all teachers and all schools;

iv) define and operationalise intended goals and outcomes of good teaching; and

v) be dynamic to ensure proper scaling.

The standards should also cover at least the following domains: use of instructional time (attendance, punctuality, time on task), planning and preparation (the design of instructional activities and evaluation procedures for all students), classroom environment (making the classroom a safe place for risk taking), instruction (adapted to different students, engaging and challenging), and professional responsibilities. Special care should be placed on the ability of the teacher to strive for equity: to attend to the needs of the diversity of students in order to achieve learning outcomes for all.

5.2 | *Establishing consensus among stakeholders on the importance of developing a comprehensive, transparent and fair in-service teacher evaluation framework is vital.* Given the importance of teachers to student learning, an in-service teacher evaluation framework should be designed and planned for the short, medium and long term. The gradual building of capacities at different levels to ensure fairness and objectivity of the evaluation framework, including a cadre of well-trained external evaluators, should be considered. To facilitate acceptance and sustained reform, it is important to set up *a communication, engagement and consultation strategy* with primary stakeholders (including the general public, opinion leaders and teachers). Other important policy dimensions such as legal, regulatory and financial considerations[8] can ensure that solid foundations for reform are established in a transparent and participatory manner, even if the complete teacher evaluation system is not implemented during the mandate of a single government administration.

5.3 | *In the context of increased accountability and along with opportunities for capacity building and professional development for teachers, it is important to ensure that all teachers meet minimum levels of professional performance and results.* Growth in student learning should be at the heart of the evaluation process. In addition, however, issues relating to basic teacher effort, such as attendance, punctuality and time-on-task, can be included in the earlier stages of the teacher evaluation framework as a way of getting all teachers to perform at capacity. Basic criteria such as these can produce considerable and timely gains for the teacher evaluation system in a cost-efficient manner (*i.e.* ensuring that all of the "low-lying fruit" is collected first).

6 | **Incentives and stimuli for in-service teachers:** Although performance rewards have been used effectively in other fields of employment, their recent use in the education sector, particularly for teachers, is still being explored, monitored and evaluated. Thus, SEP, state educational authorities and stakeholders will need to determine the specific combination of monetary and non-monetary incentives and stimuli that will be most effective in Mexico. Regardless of the rewards or consequences that are linked to results, however, for teachers to be considered effective, their students should demonstrate satisfactory levels of achievement growth, while no teacher should be rated as ineffective if students show satisfactory levels of achievement growth.

6.1 | *For an effective and sustained in-service teacher incentives policy, the following five principles should guide its development:*

i) Incentives should reflect the quality of teaching. The criterion for success of the incentive programme should not simply be better pay for better performing teachers, but the contribution of teachers to improved student learning outcomes.

ii) The incentive system should recognise and support the individual teacher, the team of teachers at the school and the profession as a whole. Incentives should be embedded in a system that supports the continuous improvements of students, teachers, schools and the education system. In the longer-term, incentives based on tests should be complemented by a sound human resource management capacity in schools and at local levels that can accurately assess the quality of work, with robust external validation and corroboration methods jointly owned by government and the teaching profession.

iii) The incentive system should build on a sound understanding of what motivates teachers and should embrace multiple dimensions of motivation, with the aim to foster an attractive work environment, create and facilitate advancement along a career path, provide access to professional development, and identify and promote effective teaching practices. Incentives and stimuli should therefore consider financial and non-pecuniary incentives, such as working conditions, material inputs for schools and classrooms, social recognition, enhanced training and professional development opportunities, or a combination thereof.

iv) The incentive system should provide good feedback mechanisms and access to professional development, to ensure that teachers who do not receive the incentive understand what they can do to improve performance and have incentives to change behaviour. It should foster a culture based on evidence and data.

v) The incentive system should reward both good performance and relative improvement, and consider the value added by teachers and schools, net of socio-economic factors. While value-added analytical models are being developed, however, simpler methods can be employed to ensure that students, schools and teachers are compared with those in similar contexts (*e.g.* socio-economic stratification and/or contextualised attainment models).

6.2 | *It is important to clearly distinguish an in-service teacher-incentive policy from other teacher-related programmes that may appear to be similar, but that do not fundamentally provide incentives to teachers to improve performance.* Incentive policies should be communicated clearly to teaching professionals in advance of the assessments and measures that will be used for the awards. In addition, each eligible teacher should have a probability of being rewarded for outstanding performance that is greater than zero, which is currently not the case for similar programmes in Mexico. Of particular importance will be finding a balance between national guidelines for the incentives (and at least partial funding), and state-level flexibility and co-participation in resources (financial or otherwise) for incentives and stimuli to teachers. Finally, a pilot of possible incentive programmes is highly recommended to ensure viability and cost-effectiveness of policy design. Pilot exercises should be rigorously monitored and evaluated in order to be most useful and worthwhile, with a base line and as much control as possible (*e.g.* randomised or quasi-experimental trials, if conditions allow).

6.3	*In-service teacher incentives in Mexico should motivate individual teachers to improve performance, but use the school as the basic unit of accountability,* given the current state and prospects of data systems and the quality of information available. With the data and systems currently available, local education authorities can develop measures to confirm and validate the eligibility of teachers for incentive awards (*e.g.* with on-site inspections and data validation of student, school and teacher information). As a robust and credible individual teacher evaluation system is developed, incentive policies could be modified to ensure that teachers are able to receive incentives individually in the future.[9] For school incentives, schools should be made publicly accountable for the additional resources received. If schools have discretion over the allocation of resources provided by the incentives, mechanisms should ensure transparency and the progressive involvement of relevant stakeholders, including parents and local school councils.
6.4	*Financial and non-financial incentives and stimuli to teachers should be based on a fair and adequate assessment and evaluation process.* Given the diversity of the Mexican educational system, a valid and reliable assessment process to identify eligible teachers for incentives needs to be developed. The success of incentives is directly linked to the credibility and fairness of the assessment and evaluation process upon which they are based. Models that take into consideration the socio-economic diversity of Mexican students, as well as other factors that can largely influence student performance, such as Spanish as a second language and ethnicity, for example, should be used when making comparisons among schools and their teachers. Special-education schools and programmes, as well as pre-primary schools, could be evaluated on the basis of appropriate measures of teacher performance and student learning, where possible. Given the diversity between and within states, the incentives policy should also consider a relative premium for disadvantaged rural schools, as opposed to non-disadvantaged urban or rural schools. Incentives should also support continued improvement of schools and teachers across the entire performance spectrum.

NOTES

1. The other publications are *Improving Schools: Strategies for Action in Mexico*, stemming from the work of the OECD Steering Group on School Management and Teacher Policy in Mexico, and *Evaluating and Rewarding the Quality of Teachers: International Practices*, edited by Susan Sclafani for the OECD. In addition, a Spanish edition of the OECD report on evaluating school contributions to student learning *La medición del aprendizaje de los alumnos: Mejores prácticas para evaluar el valor agregado de las escuelas* has been updated and produced through the Co-operation Agreement (OECD, 2010). There are also numerous working papers from invited experts and OECD staff that have contributed to the work of both Steering Groups.

2. Recommendations regarding support, capacity building and professional development for teachers, for example, are provided in the sister OECD publication *Improving Schools: Strategies for Action in Mexico*.

3. The unit of accountability refers to the level at which the effort, capacities and performance of students, teachers and principals are monitored, assessed and evaluated. Although students are assessed individually, for example, and teachers should be motivated at the individual level to improve their performance, results are grouped so that individuals are held collectively accountable at the school level.

4. The specific technical, administrative and logistical recommendations on further development of the ENLACE assessment are presented in Chapter 5.

5. A detailed discussion of the benefits, characteristics and design issues of value-added modelling is presented in the updated OECD 2010 publication available in Spanish: *La medición del aprendizaje de los alumnos: Mejores prácticas para evaluar el valor agregado de las escuelas*.

6. In Mexico, the school zone is an administrative designation of a group of schools for the purposes of supervision and administrative monitoring. Similarly, municipalities are one of the three basic jurisdictional units of government and can contribute significantly to infrastructure and material conditions of schools.

7. For example, C. Danielson's *Framework for Teaching*, Perrenoud (2004); Rewards and Incentives Group (2009); Ontario Ministry of Education (2009); Khim Ong (2008); and Singapore Ministry of Education (2006). Current development efforts in Mexico regarding standards should be considered and evaluated based on recommended criteria.

8. As suggested by the public policy framework presented in Chapter 2.

9. Regarding the appropriate amounts for incentives, a review of international programmes shows that individual teacher incentives can range from less than 1% to more than 360% of monthly salary (OECD, 2009), although experts suggest that between 4% and 8% of annual salary can be adequate for incentives to be meaningful but not cause unwanted behaviour.