

## Approaches, methods and instruments: What and how to evaluate in the framework of the Educational Reform?

### Autonomy and journeys

Juan Carlos  
Romero Hicks  
Carlos Mancera  
Felipe Martínez Rizo

### International experiences

Andreas Schleicher  
(OCDE)  
Nuria Sanz  
(UNESCO)  
José Soares  
y Helber Vieira  
(INEP-Brasil)  
Dagmar Raczynski  
(uc-Chile)

### Local strategies

Baja California  
Estado de México  
Quintana Roo  
Sonora  
Yucatán

### The view from the INEE

Margarita Zorrilla  
Teresa Bracho

### Systematization, self-evaluation and regionalization

Alberto Curi  
Javier Treviño  
José Martín Farías  
Miguel Ángel  
Cuatepotzo

Translation into Sonoran  
and Sinaloan Mayan and  
Chihuahan Tarahumara.







# The logbook

Gazette of the National Educational Evaluation Policy in México

Año I. No. 1 / July-October 2015

## FROM OUR READERS

2

## IN OUR OWN HAND

- 3 Evaluation and the Educational Reform in Mexico  
Members of the Board of the INEE

## FROM THE DESK

- 4 The dialogues and challenges of otherness: three policy-design tools for a National Educational Evaluation Policy

## VOICES FROM THE CONFERENCE

- 5 Autonomous evaluation: the challenges posed by the INEE's autonomy  
Juan Carlos Romero Hicks
- 7 Evaluation in the context of the Educational Reform  
Javier Treviño Cantú
- 9 The role of evaluation in educational improvement  
Alberto Curi Naime

## FACING OUR CHALLENGES

- 11 Comprehensive evaluation in Campeche: beyond implementation and diffusion  
José Martín Farías Maldonado
- 13 Educational evaluation and the regionalization strategy  
Miguel Ángel Cuatepotzo Costeira

## OUR VOICE

- 15 The challenge of using the information stemming from educational evaluations in Mexico  
Teresa Bracho González

- 17 The challenge posed by the Educational Reform: Classroom Observations  
Margarita Zorrilla Fierro

## ROADMAP

- 19 The National Learning-Evaluation Plan (PLANEA): the Involvement of Different Participants in the Education Process in the Evaluation of Results  
Andrés Sánchez Moguel
- 22 Curriculum evaluation  
Rebeca Reynoso Angulo and Raquel Ahuja Sánchez
- 24 A new look at the diffusion and fostering of the evaluation culture  
José Luis Gutiérrez Espíndola and Annette Santos del Real
- 27 The evaluation of educational policy  
Giulianna Mendieta Melgar

## SPEAKING ALOUD

- 29 The challenge of providing high-quality education in Latin America and the implications for the Mexican Evaluation System  
Nuria Sanz Gallego

## IN THE CLASSROOM

- 32 Veracruz: Evaluation faced with the challenges of the Educational Reform  
José A. Ojeda Rodríguez

- 33 Five examples of evaluation in Mexico: local autonomy, experiences and context  
Sonora, Estado de México, Quintana Roo, Baja California y Yucatán

## DOSSIER: COMPREHENSIVE AND CONTEXTUALIZED EVALUATION

### WITH OUT A PASSPORT

- 36 Monitoring compliance with/exercise of the right to education in Brazil  
José Francisco Soares and Helber Ricardo Vieira
- 38 Educational Reform and Public Policy design in the Latin-American countries: What do the Educational Ministers could do with PISA evidences?  
Andreas Schleicher
- 41 The evaluation of educational reforms and policies: reflections on the Chilean experience  
Dagmar Raczynski

### OTHER PERSPECTIVES

- 44 Black boxes and black holes: evaluation and assessment of the Educational Reform in Mexico  
Adrián Acosta Silva
- 46 Towards a possible future: the expectations for, and challenges facing, educational evaluation in Mexico  
Felipe Martínez Rizo
- 47 The school as comprehensive educational-evaluation center  
Carlos Mancera Corcuera

Gazette of National Educational Evaluation Policy in Mexico, Year 1, No. 2, July-October 2015, is a publication of the National Institute for Educational Evaluation by the Educational Policy and Regulations Unit, under the care of the General Direction of Guidelines for the Educational Improvement. Barranca del Muerto No. 341, Col. San José Insurgentes, Del. Benito Juárez, C.P. 03900, México, D. F. Tel. (55) 62.70.66.00.



INEE  
<http://www.inee.edu.mx>



Twitter  
@INEEMX



INEE Facebook  
<http://goo.gl/axitPa>



INEE Youtube  
<http://goo.gl/fHRDvC>

## FROM OUR READERS

In its first month, Number 1 of the *Gazette of National Educational Evaluation Policy in Mexico*, (March-June, 2015): *Training, work and teacher development: What do we know and what do we need to change?*, reached 30 117 readers via email, 33 521 via social networks, and 2 000 in print form. In Facebook, it was consulted by 32 088 surfers, of whom 1 433 entered the microsite to read the full digital version, also being downloaded 3 009 times, in its PDF format, via Twitter.<sup>1</sup>

## Collaborations and suggestions:

Because evaluate is to explore, learn and reach, the *Gazette of National Educational Evaluation Policy in Mexico* is a boarding to share the journey in order that all rightfully receive education with quality and equity. If you want to send comments and suggestions, please get in contact with us:

ARCELIA MARTÍNEZ BORDÓN  
General coordination  
amartinezb@inee.edu.mx

LAURA ATHIÉ  
Responsible editor  
lathie@inee.edu.mx

The *Gazette* has given me ideas and helped me to come with suggestions [...] I'm a teacher-researcher in the undergraduate course in Productive Social Pedagogy at the *Universidad Mayor de San Simón*. We promote strategies for improving educational quality in the rural *Trópico de Cochabamba* region. Evaluation, which is crucial to this endeavor, has only focused on student results and achievement, and we're a long

way from seeing it as a broader, more general, multidimensional process, with several protagonists and factors involved. I'm sure that being in touch with you via the Newsletter would help us to do so.

**DANIEL PEDRO TIRADO RAMÍREZ**

Cochabamba, Bolivia

Congratulations. I've seen the *Gazette* online and it's really wonderful. I'm going to circulate via the network that I run.

**F. JAVIER MURILLO TORRECILLA**

RINACE. Universidad Autónoma de Madrid

The *Gazette* has filled a vacuum in educational evaluation, and I consider it the best publication, at the national level, on that subject [...] It brings together many complimentary writers and opinions, all with the same aim: to improve education.

**HÉCTOR MORALES CORRALES**

General Director of the State of Mexico Educational Evaluation Institute

The most important thing is that it fosters an ongoing debate between and the authorities about education. [...] The *Gazette* is interesting because it publishes the conclusions of people who are doing research in the areas of higher-education and teacher training.

**RAMÓN GUADALUPE LARA CRUZ**

Director of the "Rafael Ramírez Castañeda" Regional Teacher Training College in the state of Sonora

[...] It's well produced and contains important data that give us an idea of the direction in which teaching is moving. I suggest taking teaching contexts more into account, since we need to do this in accordance with the Law Governing the Professional Teaching Service.

**JULIETA ESTER PALACIOS VILLALOBOS**

Teacher with the Federal Distance Learning System. Coatepec, Veracruz

[...] I think it's excellent. It needs to be adapted to satisfy the needs of the different sectors involved. I'd like you to publish something about the new role of managers in education. It would be good if you could publish experiences of evaluation in the rest of country's 32 states, like the accounts you published from Tlaxcala and San Luis Potosí.

**UNIVERSIDAD AUTÓNOMA DE QUERÉTARO**

The *Gazette* is an important source for educators involved in education. [...] It would be a good idea if it became accessible to more educators, in order to broaden viewpoints and promote discussion about education in general.

**JOSÉ A. OJEDA RODRÍGUEZ**

Head of the Unit for Planning, Evaluation and Educational Control of the State of Veracruz

We thank the following people for the very useful comments that they made via the *Gazette's* divers feedback channels:<sup>2</sup>

**DIEGO ALFONSO ITURRALDE G.**, Researcher at the *Centro de Cooperación Regional para la Educación de Adultos en América Latina y el Caribe* (CREFAL-UNESCO) in Michoacán;

**RENÉ HERNÁNDEZ GÓMEZ**, teacher at the "Benito Juárez García" school in Coacoatzintla, Veracruz; **GREGORIO SÁNCHEZ OROPEZA**, researcher, CREFAL-UNESCO Michoacán;

**MARCELA ERÉNDIRA CASTILLO CASTAÑEDA**, State of Mexico supervisor;

**ROBERTO SANTIAGO ARANA**, Director of Educational Planning (SEP-Querétaro); Álvaro López Espinosa, general director, DGESE-SEP;

**LUZ RAQUEL GARCÍA GÓMEZ**, academic director of the *Universidad Autónoma del Estado de México*;

**GENARO RODRÍGUEZ FIERRO**, teacher CEBETIS-DEGETI, Acapulco, Guerrero;

**ROBERTO PEREA AGUADO**, teacher-Researcher, Querétaro, Querétaro; **PEDRO FLORES CRESPO**, teacher-Researcher;

**FERNANDO PAREDES RAMÍREZ**, Director of Planning, Programming and Evaluation (SEP-Querétaro);

**GUILLERMO PADRÉS**, Governor of the State of Sonora, and **HÉCTOR EDUARDO VELA VALENZUELA**, Minister of Education of the State of Durango.

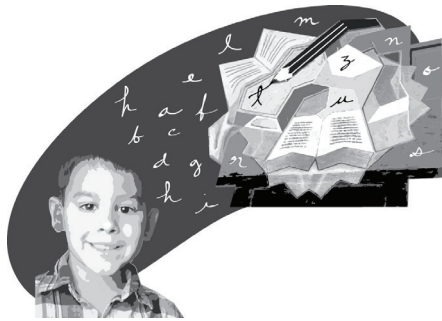
<sup>1</sup> Information from the General Department of Computing and Technical Services, the General Department of Social Communication, and the Guidelines and Educational Policy Unit of the INEE, up to March, 2015.

<sup>2</sup> *Gazette* survey in the INEE's web microsite, telephone survey, personal and mail interview, visits to the INEE's office, mails through gacetapnee@inee.edu.mx and comments on social networks.

## IN OUR OWN HAND

## Evaluation and the Educational Reform in Mexico

MEMBERS OF THE BOARD OF THE INEE



*The greatest good of man is to talk of virtue every day of his life [...], whether examining myself, and examining others, because an unexamined life is not worth living.*  
Socrates

Still beset by the challenges of the current Educational Reform, Mexico is no stranger to dealing with world-class challenges. Fifteen years ago, like many other nations, our country signed the UNESCO's six Education-for-All Goals and the United Nations Millennium Declaration's Development Goals. Subsequently, via the Educational Reform, that was published in 2013 and promoted compulsory education -which has been considered a human right since 1993- stressing high quality and creating a clear legal framework that entitles all Mexican children and young people to high-quality education.

Some basic questions arise regarding the aforesaid entitlement – ones that evaluation should help to answer, since ascertaining what progress has been made in our country's vouchsafing of the said right to high-quality schooling for all is an inescapable task that requires the development of a measurement-and-evaluation system by the National Educational Evaluation System

(Spanish acronym: SNEE) within the framework of the National Educational Evaluation Policy (Spanish acronym: PNEE), for the purpose of giving feedback to the education system about (a) the progress made in providing high quality education to all as a right, (b) the effectiveness of the government actions aimed at solving educational problems, and (c) the gap or shortfall between the current state of education and the high-quality education mandated in the Law, and the extent to which the State has fulfilled its obligations under the latter.

Though the Executive Order requires the creation of the SNEE -along with the formulation, implementation and oversight of the PNEE- we face many challenges in our endeavor to implement and consolidate the said Policy, including, to mention just a few, those of dismantling the current legal framework under which various authorities have different powers and functions, of coordinating and unifying the different entities responsible for the PNEE, of aligning evaluation policy with education policy, and of strengthening the state-level educational evaluation teams.

In order to tackle these tasks, along with those that will face us in the future, it is essential that, within the framework of the SNEE, an agenda be created that includes priorities, structures, procedures, programs, actions and resources; helps us to plot learning-improvement paths based on evaluation; fosters improvement, accountability and the dissemination of results; and helps to build an evaluation culture in all the areas of education and in all the people and organizations involved in different responsibilities from educational evaluation – i.e. teachers, students, parents, organized civil society and ourselves.

Though we still have a long way to go, we are taking important steps in order to comply with the provisions of the Law and provide evaluation that meets the requirements of the education system.

Thus, since 2014, the Dialogues for the Building of the PNEE -aimed at discussing different topics, approaches and organizational proposals related to the evaluation policy- have been held with the federal and local education authorities. Furthermore, the statutory ordinary assemblies of the

SNEE Conference have been held, at which we have reported on the progress made in the INEE's projects and the criteria for the implementation of the different nation-wide evaluations.

This year, almost all the evaluation processes mandated in the General Law Governing the Professional Teaching Service (Spanish acronym: LGSPD) will be carried out in accordance with the guidelines issued by the INEE, and the National Learning Evaluation Plan (Spanish acronym: PLANEA) will be implemented, providing reliable information about the progress achieved in student learning outcomes.

In 2014, the INEE issued its report entitled *The right to a high-quality education*, and, this April, it published another report entitled *Teachers in Mexico*, which describes the challenges inherent in the recruitment, initial training, admission to the public education system, and professional development, of teachers. It also issued a detailed report on the results of the "Prior free and informed consultation with the indigenous peoples and communities about educational evaluation", carried out in collaboration with UNICEF's Mexico office.

The Institute will also be issuing the first guidelines for improving educational policy, based both on the results of the evaluation and also on the retrieved results of various research projects on initial training, professional teacher development and the education of migrant children, as well as the results of the aforesaid "Prior free and informed consultation with the indigenous peoples and communities about educational evaluation".

High-quality education is both a public demand and a national imperative, as well as being a commitment made by our country under an international agreement, but, above all, it is an inalienable universal human right.

The challenge is enormous. It is incumbent upon us, as "obligors" in the context of the Reform, to rethink educational evaluation and convert it into a tool for improvement – i.e. to use every available means, including programs, funding and legal compulsion, to ensure that the over 33 million compulsory-school-age children and young people in our country can have a decent life,<sup>1</sup> a better

one than we have had, by respecting their right to a decent education and giving them schools where they can learn.

To this end, this 2<sup>nd</sup> edition of the INEE's National Educational Evaluation Policy Newsletter (Spanish: *Gaceta de la Política Nacional de Evaluación Educativa en México*) takes an in-depth look at our country's progress in the field of evaluation and endeavors to ascertain the dimensions and scope of the task still facing us. Like those that have appeared in previous editions of this publication, the articles contributed to this edition by our colleagues in Mexico and abroad -this time with an emphasis on

educational evaluation and, above all, on the use of the latter's approaches, instruments and results to achieve the better, fairer high-quality education that we are all seeking-will enable our readers to keep up to date with social developments and findings in the area of education. €

1 The National Population Council (Spanish acronym: CONAPO) estimates that there are 33,524,563 young people under 15 years of age in Mexico. CONAPO, 2015, based on data from the 2010 Census of Population and Housing.

Hence, faced with this *state of otherness*, we acknowledge that, in order to engage in dialogue and thus design a National Educational Evaluation Policy (Spanish acronym: PNEE), it is indispensable that we -the nation as a whole, the federal and state governments, and society in general- jointly map out a route that will ensure that all Mexico's young people have access to high-quality education.

Indeed, the said road began in the first edition of this newsletter (Number 0), where it was asserted that our publication's *raison d'être* was "...to explain the positions of those who are taking part in the Conference regarding the National Educational Evaluation System (Spanish acronym: SNEE)".

Four months later, in response to our question, in the second edition (Number 1) of our Newsletter, about what the reaction to the latter was, we received "a flood of opinions" from our readers. Now, in this edition, we ask ourselves the question, "What system of cogs and wheels does this publication form part of?", to which the response is "a system that brings together initiatives for the building of a shared agenda". Like the SNEE Conference and the Dialogues for the building of the PNEE, the *Gazette of National Educational Evaluation Policy* is a key tool for the formulation, design and Monitoring of the National Educational Evaluation Policy.

Embarked on in 2014, the Dialogues for the building of the PNEE have been forums for one-to-one, high-level discussion between the members of the INEE's Board of Governors and the education authorities about the implementation of the Education Reform and the challenges posed by the Evaluation of the National Education System.

The Undersecretaries for Elementary Education, the Planning and Evaluation of Educational Policy, and Lower Secondary Education, along with the General Directors for the Planning and Evaluation of Policy and Higher Education for Professionals in Education, the heads of various areas of the Ministry of Public Education (Spanish acronym: SEP), and the ministers of education and representatives of all the states except Chihuahua, Baja California Sur and Guerrero took part in the second stage of the said Dialogues, which was held between the 7<sup>th</sup> and the 22<sup>nd</sup> of April of this year.

Via the members of its Board of Governors and its unit heads, the INEE proposed to

## FROM THE DESK

### The dialogues and challenges of otherness: three policy-design tools for a National Educational Evaluation Policy

Untamable but gregarious, we human beings try to live together. Sometimes we are like wildcats, without language or ways to communicate, lacking sufficient creativity to find the right way to exchange ideas with the other – that person whom we do not know.

In his essay, *Encountering the Other*, the Polish journalist and writer, Ryszard Kapuściński, appeals to our instincts when

he says that the three options when we run into the unknown are "to make war, to build a wall around ourselves, or to start a dialogue". How, then, can we react to the other? Words, dialogue, democracy and participation are a good option, especially in the case of a national policy that is attempting to find a solution to a myriad problems that cannot be solved without dialogue, says Kapuściński (2005):

We should seek dialogue and understanding with the new Other. The experience of spending years among remote Others has taught me that kindness toward another being is the only attitude that can strike a chord of humanity in the Other. Who will this new Other be? What will our encounter be like? What will we say? And in what language? Will we be able to listen to each other? To understand each other?





the education authorities that discussions be held about “the National Educational Evaluation Policy and the right to a high-quality education for everybody”, as a basis for developing the PNEE guidelines for the SNEE projects.

Having opened the debate by discussing what was meant by “a high-quality education for everybody” and “evaluating with a stress on “entitlement to a high-quality education”, as well as asking what each member of the SNEE had to do in order to foster, respect, safeguard the said right to a high-quality education, the participants affirmed the need to strengthen the mechanisms of the Dialogues, in view of the latter’s usefulness in promoting the one-to-one exchange of ideas, fostering familiarity with other problems and experiences, and coming up with proposals regarding how to disseminate, raise awareness about, and gradually implement the focus on rights in the building of the PNEE.

Among other things, it was agreed to draw up a governing PNEE document, including in it a rights focus, policy guidelines, proposals as to actions and goals, and a roadmap which, as stated in one of the sessions, would “enable us to write the script for a better future”.

Hence, in confirmation of Helen Hanff’s quotation from John Donne, (2002) in her book, *84, Charing Cross Road*, regarding communication with *the other, the unknown*, we can conclude that the *Dialogues* and the *Newsletter* are tools that help us to map out a route for discussion even where there is disagreement, since: “No man is an island, Entire of itself/Every man is a piece of the continent, A part of the main;/ [...] Because I am involved in mankind”<sup>1</sup> €

1 John Donne, *Devotions upon Emergent Occasions*, quoted, in Spanish translation, in Hanff, H. (2002). *84 Charing Cross Road*. Barcelona: Anagrama.

For more information about Helene Hanff’s book, *84 Charing Cross Road*, see page 64 of the Roadmap section of this *Gazette*.

## References

- Hanff, H. (2002). *84, Charing Cross Road*. Barcelona: Anagrama.
- Kapuściński, R. (2005). *Nobel Laureates Plus and La Nación*, consulted on May 20th, at: <http://goo.gl/pL0UQS>

## VOICES FROM THE CONFERENCE

### Autonomous evaluation: the challenges posed by the INEE’s autonomy

The INEE was afforded autonomy following the Educational Reform, due both to its mission and to structural challenges. Today, it is free to evaluate the education system and its components, and is charged with carrying out this task with maximum efficiency, in collaboration with the other members of the education sector. “The Standing Constituent Committee decided to grant it autonomy because it considered that its activities were a matter of priority for the Mexican State,” says the author of this article.

JUAN CARLOS ROMERO HICKS  
Chairman of the Senate Education Committee  
[romero.hicks@pan.senado.gob.mx](mailto:romero.hicks@pan.senado.gob.mx)

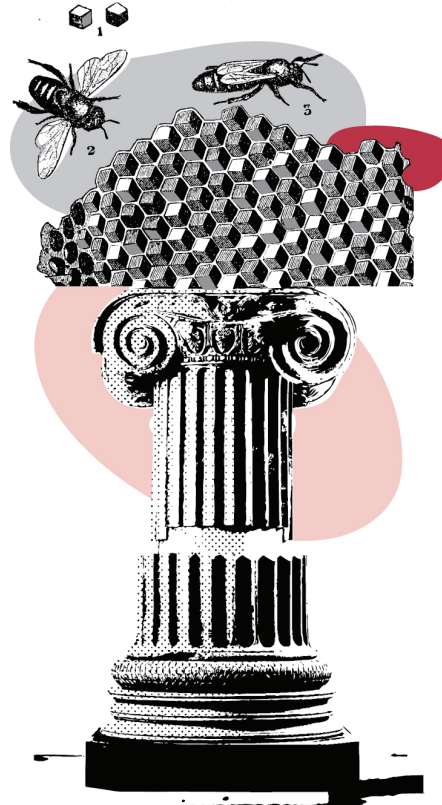
#### The conquest of autonomy

The National Educational Evaluation Institute (Spanish acronym: INEE) is a young institution. Almost 13 years have gone by since it was set up, and during this time it has matured by leaps and bounds due to the importance of its job, the changes in its structure as an institution, and the complexity of the National Education System (Spanish acronym: SEN).

The INEE was set up in 2002, under a decree issued by then president, Vicente Fox Quesada, as a decentralized entity of the Ministry of Public Education (Spanish acronym: SEP), and granted optimal technical autonomy while still being subject to control as a dependency of the latter, having its budget managed by the SEP and being answerable to whichever Minister of Education of Education is in office at a given time. In 2012, during the administration of President Felipe Calderón Hinojosa, the INEE was granted greater independence in view of the results it had produced. By presidential decree, it ceased to fall under the SEP and was given increased political and administrative autonomy, while remaining a Federal Government Agency responsible to the Executive Branch.

Having acquired technical prestige during those first ten years, the Institute produced annual data, through its evaluations, about the state of education in our country, analyzing the said data and developing qualitative evaluation instruments. Its census-based and sample-based tests got better as time went on, furnishing us with serious information about student learning outcomes and the performance of the SEN.

The INEE was charged with administering the international tests pertaining to the Program for International Student Assessment (PISA) of the Organization for Economic Cooperation and Development (OECD), also taking part in regional studies such those carried out by the Latin American Laboratory for Assessment of the Quality of Education (Spanish acronym: LLECE) and in the Second Regional Comparative and Explanatory Study (Spanish acronym: SERCE).



Since evaluating learning outcomes, results and the state of education in general means also evaluating educational policy and, hence, education authorities. Without having had a chance to observe the newly structured INEE in action, the Standing Constituent Committee decided to grant it independence under the Constitution, since it needed to be free of pressure, deadlines and schedules, and unaffected by political vagaries, given that the highly technical work entrusted to it was a priority for the Mexican State.

### Some observations about constitutional autonomy

The word autonomy derives from the Greek word, *αὐτονομία*, which conjoins the prefix, *autos*, meaning ‘self’, the core word, *nomos*, meaning “norm” or “rule”, and the suffix, *ía*, meaning “action” or “quality”. The said word, which passed from Greek to Latin, and from the latter to Spanish, retains its original double meaning - referring to the quality of being independent and also the action of regulating oneself.

Several theorists agree that the concepts underlying modern political thought and underpinning the changes that occurred all over the world in the XVIII and XIX centuries laid the foundations for the system of weights and balances, in a framework of division of powers, that characterizes the different political systems that have been adopted all over the world, establishing what is almost the sole pattern of political organization (Baños, 2006).

This system consisting of three separate entities, respectively empowered to issue law, enforce them, and settle disputes, has served to protect the citizens of each country from state abuse.

The first steps leading to the creation of autonomous bodies, with the same status as the three branches of government, were taken in post-war Europe, with the setting up of institutions charged with safeguarding human rights, being followed by the establishment of entities responsible for holding and validating democratic elections, and of central banks empowered to issue money (Ugalde, 2010).

In Mexico, the Standing Constituent Committee granted autonomy to the Bank of Mexico in 1993, to the Federal Electoral Institute (Spanish acronym: IFE) in 1996, and to the National Human Rights Commission (Spanish acronym: CNDH) in 1999.

The widespread mistrust in political parties and governments, the need to have specialist in priority areas, and the urgent need to protect these entities and the work they do from political meddling, have led the Standing Constituent Committee to set up new agencies, all with the same level of autonomy.

The entities responsible for antitrust regulation and the oversight of telecommunications, transparency and access to public information, energy and energy resources, tax-collection, public prosecution and law enforcement, as well as the evaluation of public policy (both social and educational), have been deemed by the Union Congress to require the highest possible level of autonomy, and the said tasks have been entrusted, under the Constitution, to autonomous bodies that, at the very least, meet the following criteria:

1. Independence from all the branches of government.
2. The same status as the branches of government and other autonomous entities.
3. Empowerment to issue the regulations that govern them and the area that they regulate.
4. The ability to manage their own resources.
5. Responsibility for specific tasks that are priorities for the Mexican State and so important that the entities in question should be not be interfered with, or subordinated to, other entities with different interests.<sup>1</sup>

### The INEE’s autonomy

Since education is a government responsibility, it is incumbent on all the four powers. Given its importance, it is probably the most crucial and pertinent task not only for the development of Mexico, but also for that of all other countries, since it affects the most sensitive area of all societies – their children and young people. For these reasons, it is vital that we have periodical, objective, technically rigorous evaluations of the SEN as a whole and of its different components, of educational policy, performance and results, and of those involved in education, for the purpose of improvement, correction and planning.

The SEN is one of the most complex systems in the world, involving a total of 50 million people, including teachers, students, management staff, authorities and parents,

whose interests are not always compatible with high educational quality and whose behavior does not always tend to promote it..

Given the magnitude of the challenge, the technical complexity of education and the interests surrounding it -ranging from legitimate concerns, such as the quality of teaching, to economic, political or ideological ones that pervert its aims- the biggest challenge for an entity charged with evaluation is to be objective and assert its independence vis-à-vis a whirlwind of conflicting agendas, interests and very real needs.

This is the main reason why the Standing Constituent Committee, and later the Union Congress, bestowed constitutional autonomy on the INEE, setting up the National Educational Evaluation System (Spanish acronym: SNEE), and assigned it the task of evaluating and overseeing education and issuing guidelines for it.

Like people, institutions gradually build themselves. Autonomy must be exercised and respected, and failure to exercise it and have it respected means breaking the law and flaunting the Constitution. This institutional-development process requires a twofold effort: that of the autonomous institution itself to be its own master, and that of those of us who make up the other branches of government to respect it. Under our Constitution, the INEE has the same status as any other autonomous entity –as the Supreme Court of Justice or the Federal Executive Branch- and its mission is to build an objective, rigorous, reliable SNEE that helps to raise the quality of education.

### Future challenges

The INEE’s biggest challenge, in the view of the Legislative Branch, lies in fulfilling the responsibilities that have been assigned to it under the Constitution and becoming fully effective. A law does not magically transform things as soon as it is passed, but requires hard work, commitment and change in the people responsible for enforcing it. Laws do not enforce themselves, and the conditions needed for their enforcement do not spring up from one day to the next. Rather, we need people and institutions to comply with them and enforce them.

A review of Section IX of Article 3 of the Constitution reveals what the INEE’s mandate is and which challenges this implies:



IX. The SNEE is established so as to ensure that high-quality education services are rendered, and will be coordinated by the INEE), which shall be an autonomous public institution with its own legal status and funding. The INEE shall be responsible for evaluating the quality, performance and results of the SEN, in the areas of Pre-school, Primary, Lower-secondary and Upper-secondary Education, and hence it shall:

- A) Design and carry out evaluations of the System's components, processes and results.
- B) Issue guidelines to be complied with by the federal and local education authorities for the purpose of carrying out the evaluations that are incumbent upon them.
- C) Compile and disseminate information and, based on it, issue guidelines that foster the taking of decisions aimed at raising the quality of education and promoting educational equality as a prerequisite for social equality.

To fulfill the above obligations, the INEE must discover its role in the creation and consolidation of the Comprehensive Educational Reform. The conferral of autonomy on the agency responsible for evaluation, and the creation of a Professional Teaching Service (Spanish acronym: SPD), do not, per se, amount to an Educational Reform. We must build a new educational model by reviewing methods, study plans, curricula, materials, resources, aims, predominant contents, the dissemination of competencies throughout the different states, real school autonomy, initial and ongoing training, and a long list of other pending tasks of which not everybody involved in the process seems to be aware.

The INEE's first challenge lies in creating the SNEE, but one has the impression that some of the people and entities involved in education -especially the local and federal education authorities- think that they are no longer responsible for evaluation. This constitutes a misunderstanding of the regulatory framework, for the main entities on which the legal obligation to evaluate is incumbent are, indeed, the education authorities, with the proviso that they now have to do so in accordance with the guidelines

issued by the INEE and within the context of a national system that will be as strong as its weakest links.

The second challenge is the technical one. Designing and measuring the components, processes and results of a complex system like the Mexican one is a monumental task; one of the reasons why we made the INEE autonomous was so that it could create ever more objective and pertinent evaluation instruments that will enable us to build on whatever works and rectify what doesn't work.

The INEE must learn, very quickly, how to exercise its constitutional powers to regulate and issue guidelines. It will become the key tool for improving and transforming the education system to the extent that its technical and legal teams work together to move ahead with the issuing of guidelines that enable the other members of the System to constantly evaluate themselves.

However, the biggest challenge for the INEE's -and the one that will bear the most fruit if overcome- is that of exercising its autonomy, which is a *sine qua non* for its fulfillment of the noble and difficult task that it faces. Autonomy must be built, exercised, protected and respected. Though inevitably winding, it is the only path. €

1 c.f. Ugalde, J. (2010). Ugalde deems independence and hierarchical level to be the same. However, in the light of what really tends to happen, I believe we need to separate them.

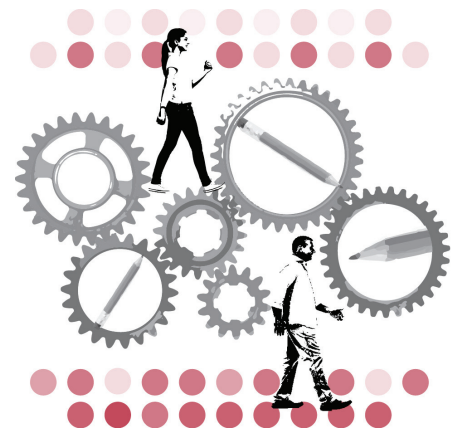
## References

- Baños, J. (2006). *Teorías de la democracia: Debates actuales*, in: *Andamios*. Vol. 2. No. 4, Mexico.
- Ugalde, F. (2010). Órganos constitucionales autónomos, in *Revista del Instituto de la Judicatura Federal*. No. 29, Mexico.
- Senado de la República (2012). *Dictamen de las Comisiones Unidas de Puntos Constitucionales, de Educación, y de Estudios Legislativos, Segunda, con Proyecto de Decreto que reforma y adiciona diversas disposiciones de la CREUM en materia educativa*, in: *Gaceta del Senado*, 20<sup>th</sup> of December. Mexico.
- Senado de la República (2013). *Dictamen de las Comisiones Unidas de Educación, y de Estudios Legislativos, con Proyecto de Decreto por el que se expide la Ley del INEE*, in: *Gaceta del Senado*, 22<sup>nd</sup> of August. Mexico.

## VOICES FROM THE CONFERENCE

### Evaluation in the context of the Educational Reform

Besides placing the government initiatives in context and explaining their importance for the optimal functioning of the education system, the author lists ten challenges that must be faced in the area of evaluation in order to "achieve the major aim of the Constitutional Reform in that area".



JAVIER TREVIÑO CANTÚ  
Undersecretary for the Planning and Evaluation of Educational Policies of the Ministry of Public Education of Mexico  
[javier.trevino@sep.gob.mx](mailto:javier.trevino@sep.gob.mx)

The major aim of the Constitutional Reform pertaining to education, set by President Enrique Peña Nieto and seconded by a large congressional majority, is to make education a force for transformation in Mexico and a basis for peaceful, respectful coexistence in a fairer, more prosperous society. This federal-government priority was attested to by the inclusion of the phrase, "Mexico with High-quality Education", among the five big objectives of the 2013-2018 National Development Plan (Spanish acronym: PND).

The planning process implemented in this context by the Ministry of Public Education was set forth in the 2013-2018 Sectoral Education Program (Spanish acronym: PSE), which includes, among its basic aims, the promotion of educational quality and equity, defines the school as the pivotal concern of the National Education System (Spanish acronym: SEN), and holds the State responsible for creating conditions leading to the strengthening of school-management skills.

The said PSE contains precise stipulations about the types of evaluation to be carried out within the SEN, stating they must be systematic, comprehensive, compulsory and ongoing, taking stock of the demographic and socioeconomic context of those involved, the human, physical and financial resources used, and such other factors as affect the teaching-learning process. Likewise, the Plan stipulates that the results obtained for each educational modality and level must be wide-ranging and employ approaches that can be used in all the different parts of the education sector, so that they may be adopted in a timely manner by the different education authorities.

Thus, evaluation is seen as the instrument, *par excellence*, for achieving improvement in education and fostering accountability, since both the latter are complementary factors in the endeavor to achieve fair, high-quality education.

Hence, the amendment to Articles 3 and 73 of the Mexican Constitution stipulate changes in the structure, planning, management and evaluation of the SEN so as to stress that ongoing raising of the quality of compulsory education is to be supported by periodic evaluation of the latter's components, processes and results.

In this way, the Constitutional Reform with regard to education places emphasis on ensuring both the high-quality of compulsory education and also optimal student learning outcomes based on the best teaching materials and methods, school management and educational infrastructure, as well as on suitable teachers and school managers. Furthermore, a National Educational Evaluation System (Spanish acronym: SNEE) -coordinated by the National Educational Institute (Spanish acronym: INEE) and responsible for designing assessment and measurement tools to evaluate the SEN's components, processes and results- is set up.

It bears pointing out that the General Education Law (Spanish acronym: LGE) defines quality as "congruence between the aims, results y processes of the education system in accordance with the goals of effectiveness, efficiency, relevance and equity" (Article 8, Section IV). For its part, the Law Governing the INEE defines the "quality of an education system" as a blend of "relevance, equity, efficiency, effectiveness, impact and sufficiency" (Article 5, Section III).

### Aspects of educational evaluation in Mexico in the context of the PND and the PSE

Since evaluation is a feature of all the components of the Educational Reform, clearly it cannot be limited to the design and administration of tests and other measuring tools, or to obtaining results relating to the progress made in achieving the aims of the PND and the PSE. In reality, it contemplates a much broader, more deep-reaching and precise endeavor to gauge the relevance, effectiveness, efficiency, impact and sustainability of the programs, projects and educational policies aimed at identifying the weaknesses and strengths of the SEN as part of the decision-making process aimed at promoting constructive all-round changes and developing the competencies of all those involved in education.

Therefore, evaluation processes should focus on all the factors that have to do with educational quality, relevance and equity, such as school organization, the use of time, the distribution of teachers during the school year, the different courses available, the resources and processes, the compulsory-education curriculum and the acknowledgment of, attention paid to and inclusion of, diversity in the SEN.

When evaluating the performance of in-service teachers and school managers, consideration should be given to their working conditions, their class planning, the atmosphere in their classrooms, their teaching practices, the learning outcomes of their pupils, school management, inter-school collaboration and communication with parents.

In this regard, the National School-Technical-Support Service (Spanish acronym: SATE) was set up to help teachers to carry out internal evaluations and to interpret and make use of external evaluations.

Moreover, there are plans to boost Mexico's participation in international evalu-

ations and surveys in order to have access to international yardsticks and be able to develop indices that enable our country to benefit from the experiences of similar nations that face the same challenges with regard to education.

### Challenges for evaluation and how to overcome them

Based on prior experiences of educational evaluation, one can say that the challenges we face, in our efforts to achieve the major aims of the Educational Reform, consist in:

1. Making evaluation a decisive component of the SEN, fully accepted and respected by Mexican society, via better design, reliable administration and the responsible use of results. To this end, we need to engender more openness and willingness in our teachers, and in the other people involved in education, to take part in evaluation processes for the purpose of achieving systematic improvement.
2. Affording the school a pivotal role in the process, with its students being the central focus of teaching-learning, by systematically monitoring the atmosphere in both the classroom and the school in general, as well as inter-student conviviality, since these things affect the quality of education. To do this, we need to provide those involved in education –teachers, management staff, students and parents- with enough clear, timely information, in order to fully convince them that the results will have a real, positive impact on the school environment.
3. Fostering the development of competencies in both teachers and management staff, to enable them to administer evaluations to their students and learn, themselves, from the results of external evaluations.
4. Offering training and specialized programs to those responsible for carrying out evaluations.
5. Boosting cooperative federalism – i.e. improving the relationship between state and municipal governments and the federal one, since education is a shared task. This type of cooperation is conducive to equity, since it ensures that students, parents, teachers and school managers throughout Mexico all enjoy the same opportunities.

## VOICES FROM THE CONFERENCE

6. Designing and administering evaluations for compulsory education that take stock of the specific contexts in which teaching occurs in the different parts of our country, and including features that take into account our population's ethnic, linguistic and cultural diversity, and identify and control features that could result in cultural bias, above all with regard to migrant, poor and handicapped students, among others.
7. Designing and compiling a system of yardsticks that provides enough clear, timely information about educational coverage, quality and equity.
8. Creating a diagnostic survey of all the current evaluation programs so as to detect overlaps and synergies, and also in order to eliminate duplication and deficiencies and foster the capacity to come up with concrete proposals for improvement.
9. To make the transition from evaluation-based decisions to educational-policy ones – i.e. to ensure that evaluation leads to action and really engenders educational improvement, as well as creating efficient, reliable accountability mechanisms that everybody trusts.
10. Building a shared agenda, congruent with the priorities of society, that identifies convergences with, and deviations from, government policy and comprises the features stipulated in the PND and the PSE, and also in the current legal framework.

In order to overcome these challenges, the Ministry of Education is working with the state education authorities, in the framework of the SNEE, to determine which evaluation projects and actions should be implemented in accordance with the National Educational-Evaluation Policy (Spanish acronym: PNEE).

In this regard, the PNEE promotes the evaluation of components, processes and results, not as a type of control or supervision, but rather as the most important instrument for producing information and knowledge aimed at raising the quality of education. €

## The role of evaluation in educational improvement

ALBERTO CURI NAIME  
Undersecretary of Elementary Education,  
Ministry of Public Education of Mexico  
[alberto.curi@nube.sep.gob.mx](mailto:alberto.curi@nube.sep.gob.mx)

*“Evaluation doesn't just serve to measure and classify; it should also be used as a basis for understanding and learning”.*  
(Miguel Ángel Santos Guerra)

All evaluations seek to come up with a data-based assessment, founded on information about an object or process, by means of observation. In education, evaluation has been conceived of as a “systematic process of investigation and understanding of educational realities that endeavors to make a value judgment about the said realities for the purpose of making decisions and improving the intervention” (Ramos *et al.*, 2009:60). In accordance with this interpretation, the decisions stemming from an evaluation have a direct impact on those taking part in the evaluation, since all evaluation is established on actions taken by people who are acknowledged protagonists in the education process, and also on the use of the results to reorient the design and implementation of a public policy.

Hence, in order to foster educational quality, it is important to have a clear understanding of which teaching-learning practices engender significant improvement in processes and results, which is why it is important to know what is happening in the classroom, and, based on that, promote better teaching practices, since, as asserted by Álvarez in his book, *Evaluar para conocer, examinar para excluir* [“Evaluating to Find Out; Examining to Exclude”], evaluation that aspires to be formative must be continually at the service of practice, in order to improve it” (2014:14).

Hence, besides promoting accountability, educational evaluation must be able to give rise to dialogue and consensus that informs

the practices of teachers, school principals, supervisors, local education authorities, and the federal education authorities themselves.

This means that the aims of evaluation coincide with those of education itself and with the pedagogical, ethical and social focus of the National Education System (Spanish acronym: SEN). Thus, in the context of the 2013 Educational Reform, evaluation should basically set out to engender more, better, learning opportunities.

Thereby it no accident that, as a result of the Reform, our country now has various laws and regulations that establish links between evaluation and the State's duty to vouchsafe high-quality education, as mandated in Article 3 of the Mexican Constitution (Spanish acronym: CPEUM), which stresses linkage among teaching methods and materials, school organization, educational infrastructure and the aptitude of teachers and school principals to achieve “maximum learning outcomes in students” (Official Federal-Government Gazette, [Spanish acronym: DOF], 2013a:1), and guide teaching-learning processes – above all the evaluation processes that are carried out, as mandated in the Constitution itself, by the National Educational Evaluation Institute (Spanish acronym: INEE), which is now an autonomous government entity.

The information yielded by evaluation will serve as a basis for decision-making aimed at changing the educational policies that orient the government's actions and also those of the SEN, placing emphasis on helping schools to carry out their mission and more effectively fostering learning in children and young people.

### Evaluation for quality – a crucial part of the Educational Reform

In order to ensure that government teaching services are of a high quality, the Reform mandated the setting up of the National Educational Evaluation System (Spanish acronym: SNEE), coordinated by the INEE, in which the federal and state-level education authorities participate via coordination mechanisms that foster greater local participation, through support that is helping to create ever more comprehensive evaluations, which should, in turn, promote the



use of the evaluation results to improve, and lead to innovations in, the planning and implementation of educational policy.

Since high educational quality also implies equity, it should be stressed that the law governing the INEE empowers the latter to design and carry out evaluations that help to raise the quality of learning, with special emphasis on the different regional populations, on ethnic, cultural and linguistic minorities, and on the handicapped" (DOF, 2013b:6).

High quality is mainly defined in terms of successful learning, and, hence, the way to achieve the latter. Article 3 of the General Education Law (Spanish acronym: LGE) echoes the CPEUM mandate, stipulating that the State "is obliged to provide high-quality education services that ensure maximum learning outcomes in students" (DOF, 2015:19).

Furthermore, the Reform places emphasis on teaching, since schools, and hence the education system, exist to promote teaching-learning. Article 12 of the General Law Governing the Professional Teaching Service (Spanish acronym: LGSPD) cites the CPEUM mandate regarding education, when stipulating "suitable teacher profiles". With regard to the latter, the INEE and the National Department for the Coordination of the Professional Teaching Service are responsible for regulating the entrance of elementary and upper-secondary-level teachers to the Professional Teaching Service, and their promotion, acknowledgment and continuance within it.

These laws and regulations bear witness to the fact that Ministry of Public Education (Spanish acronym: SEP) is concerned about improving student learning outcomes and, hence, teacher performance and also school supplies processes, environments, as well as results.

### **The commitment of the SNEE**

The aim of educational evaluation is to ascertain the extent to which student-achievement expectations are fulfilled, which explains the setting up of systems devoted to "developing suitable measuring and evaluation tools to verify that these expectations are fulfilled" (Ferrer, 2006). Hence, if evaluation is to focus on students and the teaching-learning process, it must furnish teachers with instruments that enable them to assess their students' progress in achieving

pre-established goals and identify the competencies and knowledge that the said students need in order to carry out everyday activities.

The commitment to improve learning is also manifested in the National Development Program (Spanish acronym: PND) and the 2013-2018 Sectoral Education Program (Spanish acronym: PSE), which, in particular, asserts that "a very important change, proposed in the Reform, consists in placing the school at the center of the education system" (DOF, 2013b:8), and proposes strategies that enable schools to "strengthen their commitment to student learning and become promoters of change and social transformation" (DOF, 2013b:8).

Hence, we need to link evaluations of educational quality to decision-making, since the law mandates that the INEE, besides drafting and disseminating an annual report on the state of the components, processes and results of the education sector, must "issue guidelines that help to inform decisions that tend to raise educational quality and equity, as a prerequisite for seeking social equality" (DOF, 2013b:12).

In view of the above, one can see that the SNEE should operate in a context of respect and collaboration among those of us who share the responsibility for promoting an educational policy that fulfills the requirements contained in the Educational Reform itself.

### **Learning evaluations in elementary education**

Indubitably, we need to move ahead with the analysis of the results of the evaluations carried out in the last few years –e.g. the Examinations of Quality and Educational Achievement (Spanish acronym: EXCALE) designed by the INEE, or the National Evaluation of Academic achievement in Schools (Spanish acronym ENLACE) promoted by the SEP, as well as studying the results that spring from the new evaluations produced under the National Evaluation and Learning Plan (Spanish acronym: PLANEA), which was also designed by the Institute and is administered in the third year of pre-school, the sixth year of primary school, and the third year of secondary school, as well as an eminently formative diagnostic test for the third year of primary school, aimed at helping teachers and school principals to plan teaching, not to mention PLANEA will be

jointly implemented by the INEE and the SEP this year, the results of which can be used by the school and zonal technical committees to support school organization, teaching and projects and activities aimed at improving learning starting in the 2015-2016 study cycle. This test is administered in the schools themselves, being implemented by the teachers, who, along with the school principals, will be able to use its results immediately.

### **Some final thoughts**

One can assert that proper evaluation feedback will support the formulation and changing of educational strategy in Mexico. Note should be taken of the clear international recommendations about taking stock of context and the real possibilities of implementing evaluations (OECD, 2010) all the more so when it is hoped to use their results as a basis for formulating strategies aimed at bringing about important changes (Tiana, 2003).

Nowadays access to information is growing exponentially. Comprehension, analytical, summarizing and critical competencies, as well as creativity (García-Valcárcel, 2003), should be taken into account in evaluation, to ensure that it reflects real educational and social progress, as recommended in the "World Declaration on Education for All". It is important that we develop more, better, worldwide, regional, national and subnational evaluation systems that take stock of different components of quality and engender solid tests that serve as a basis for policy formulation and the administration of education systems" (UNESCO, 2015).

Evaluating to raise educational quality and improve learning outcomes is a big challenge that requires group-executed analysis, since, far from placing blame on teachers, "evaluation should engender a feeling of shared responsibility for education as a public asset. [...] We need to promote commitment to education by all those involved in it, each in accordance with her/his position and sphere of action" (Ravela *et al.*, 2008:62).

Elementary education continues to be the main focus of evaluation, not just in order to ensure that more students pass the national examination, but also because of its importance as an area of education that involves the basic learning needed to live a decent life, get a decent job, fully participate one's country's development, improve life

quality, make basic decisions and continue to learn (Jomtien-UNESCO, 1990).

For these reasons, as well as indicating the actions that Mexico must take to fulfill its obligation to provide high-quality education to all, evaluation should be a basis for planning the future of education based on better participation mechanisms, strategies and programs, which address the weaknesses of the SEN, and also on public policy that fosters development and brings results. €

## References

- Álvarez, J. M. (2014). *Evaluar para conocer, examinar para excluir*. Madrid: Morata.
- Diario Oficial de la Federación (2015). *Decreto que modifica la Ley General de Educación*. Last reform, April 20th, 2015. Mexico.
- DOF (2013a). *Decreto por el que se reforman los artículos 3º en sus fracciones III, VII y VIII;* y 73, fracción XXV, y se adiciona un párrafo tercero, un inciso d) al párrafo segundo de la fracción II y una fracción IX al artículo 3º de la CPEUM. 26<sup>th</sup> of February, 2013. Mexico.
- DOF (2013b). *Decreto por el que se aprueba el Programa Sectorial de Educación 2013-2018*. 13th of December, 2013. Mexico.
- Ferrer, G. (2006). *Sistemas de evaluación de los aprendizajes en América Latina. Balance y desafíos*. Santiago de Chile: PREAL.
- García-Valcárcel, A. (2003). *Tecnología Educativa. Implicaciones educativas del desarrollo tecnológico*. Madrid: La Muralla.
- OECD (2014). *Creating Effective Teaching and Learning Environments: First Results from TALIS*. 2013. Paris.
- OECD (2010). *Mejorar las escuelas: Estrategias para la acción en México*. Paris.
- Ramos, G., Perales, M., and Pérez, A. (2009). *El concepto de evaluación educativa*. in: Jorner, J. and Leyva, Y. (Coords.) *Conceptos,*

metodología y profesionalización en la evaluación educativa. Mexico: International Technology Research Institute.

- Ravela, P., Arregui, P. et al. (2008). *Las evaluaciones educativas que América Latina necesita*. *Revista Iberoamericana de Evaluación Educativa*, 1(1), pp. 51-63.
- Santos, M. (2007). *La evaluación como aprendizaje. Una flecha en la diana*. Madrid: Boinum Narcea.
- Tiana, A. (2003). “¿Qué pretendemos evaluar, qué evaluamos y qué conclusiones podemos extraer de las evaluaciones?”, en: Iaies, G. y otros. *Evaluar las evaluaciones. Una mirada política acerca de las evaluaciones de la calidad educativa*. Buenos Aires, IIPE-UNESCO. Pp. 111-132.
- UNESCO (2015). *Regional conferences on the evaluation of education after 2015: final statements*. Paris: UNESCO.

## VOICES FROM THE CONFERENCE FACING OUR CHALLENGES

### Comprehensive evaluation in Campeche: beyond implementation and diffusion

Moving full steam ahead with the Educational Reform means developing a comprehensive evaluation culture that goes beyond the use of instruments and the publication of results. One needs to reflect on the results obtained and use them as a basis for actions aimed at achieving high-quality education – i.e. to implement self-evaluation.

JOSÉ MARTÍN FARIÁS MALDONADO  
Former Minister of Education of Campeche  
[seduc.cam@hotmail.com](mailto:seduc.cam@hotmail.com)

### The self-evaluation of those involved in the Educational Reform

The Educational Reform makes it a constitutional obligation to provide high-quality education to all Mexicans – an obligation that implies the commitment of all the three branches of government, as well as students, parents and members of the education system (i.e. teachers, technical-pedagogic advisers, school principals and supervisors), education public servants in the education sector (ranging from analysts and middle managers to ministers of education), the private and public sectors, non-governmental organizations and the general public.

Given the importance of the said Reform, all of us should ask ourselves what role we should be playing in its implementation, whether that role is clearly specified, what attributes we need to play it effectively, which guidelines, criteria and procedures we should follow in order to do so, and what results and products are expected of us.

Based on the legal framework of the General Education Laws (Spanish acronym: LGE), the General Law Governing the Professional Education Service (Spanish acronym: LGSPD), the Law Governing the National Educational Evaluation Institute (Spanish acronym: LINEE), the matching state education laws, and the manuals regarding organization and procedures that



govern education-sector agencies and dependencies, one can barely answer the first two questions. Hence, it would be a good idea to draw up a document that provides precise answers to the said initial questions to both civil servants and the general public – a document that would serve to build consensus, and disseminate information, about the roles that each of us should play. Indeed, though the said task is not an easy one, in Campeche we have, at least, started drafting just such a document.

On the other hand, in order to answer the third question, regarding attributes, we need to draw up a quality profile, establish criteria and procedures, and create instruments to define the said attributes and measure the degree to which we possess them.

To answer the last two questions, we need to draw up a performance profile listing the results and products that are expected, as well as establishing the criteria and procedures, and creating the instruments, needed to measure them. In my opinion, the aforesaid profiles would have to be national yardsticks. Since there would be certain profiles pertaining to contextual characteristics of an economic, cultural and social nature -e.g. the attributes required to be a teacher in the indigenous education system or the procedures, results and products needed for given contexts- these aspects should in no case be seen as contextual or institutional reasons to reduce quality exigencies or disregard parts of the said profiles. One should not forget that the Constitutional amendment concerning education requires us to provide high-quality education services to all Mexicans, using mechanisms such as the evaluation employed to select new elementary- and lower-secondary level teachers.

These problems become more complex and hard to solve in the case of people already involved in education -i.e. in-service teachers, students, parents, education experts and public servants- since the said participants in the education process –including all of us, and especially those who are part of the National Educational Evaluation System (Spanish acronym: SNEE), would have to be given a chance to engage in self-evaluation, reflection and self-directed training, all of this in accordance with *ad hoc* criteria, procedures and instruments that would enable us to see ourselves through “the glass of quality control”, identify our

weaknesses, reflect on them, and, once aware of them, choose the most suitable ways to reduce or eliminate them, via self-directed learning, so as to comply with the quality profile. In view of the above, we suggest that, in the case of people working in the education system, the aforesaid self-evaluation be carried out before the external evaluation to determine their continuance in the said service.

Though the mechanisms currently imposed by the Professional Teaching Service (Spanish acronym: SPD) leave us no time to carry out the aforesaid self-evaluation, it would, nevertheless, be a good idea to make space for it in the next processes governing entry to, and continuance in, the SPD, and, if there is time, to also make a space for those people involved in education who are not teachers, school managers or supervisors to evaluate themselves. Ideally, the said self-evaluation would be subject to a certification procedure whenever the person engaging in it, aware of his/her weaknesses and hoping to have eliminated them via self-directed training, decided to seek such certification.

In the same way, people and entities currently involved in education should be given a chance to engage in an *ad hoc* self-evaluation process that would enable them to evaluate their own performance and improve those aspects of it that they consider to be inadequate.

In Campeche, based on the results of the self-evaluation, of the external evaluation, or of reflection and training, whether self-directed or not, processes for raising quality and improving performance will be implemented and recorded in databases for purposes of follow-up or self-managed control and to support the evaluatee’s direct line manager, subordinates or colleagues.

Since they are the only way to move forward, the said processes will apply to everyone involved in the Educational Reform process, and not only to teachers and school principals, since the latter are not the only people responsible for offering high-quality education services.

### **Self-evaluation of the programs, goods and services supporting education**

Under the Educational Reform, making high-quality available to everybody implies not only the self- or external evaluation of the quality and performance of those

involved in education, but also the carrying out of self-improvement processes.

Likewise, annual programs and their corresponding budgets, as well as compensatory support programs for the education sector, including services, must adhere to quality standards and profiles that ensure that they fulfill expectations and comply with norms - this without forgetting the criteria, procedures and instruments for the self-evaluation and external evaluation of the said programs, budgets and services. At the same time, the people responsible for the creation and running of the said programs should be capable of evaluating their quality using the documents stipulated for that purpose. In this regard, the National Council for the Evaluation of Social Development Policies (Spanish acronym: CONEVAL) possesses validated procedures and instruments that could be used for the previously mentioned purpose.

Additionally, there is a group of goods and services for which quality profiles also need to be drawn up stipulating the standards they need to meet to satisfy the requirements of an educational model based on the criterion of high-quality education for everybody. The most important items among the aforementioned goods and services are those pertaining to educational infrastructure, which leaves much to be desired in most schools and in the local administrative offices pertaining to our country’s decentralized entities. Though we possess descriptions of the current condition of our educational infrastructure, we still have no quality profile -at least in Campeche- that stipulates the criteria it should meet.

Competency-based education, which stresses the learning and intensive use of computers skills, in accordance with the current exigencies of our elementary and upper-secondary system, requires an educational infrastructure that differs from the current one. Nowadays, science laboratories with equipment meant purely for demonstration are no longer useful, having given way to virtual laboratories and multimedia equipment that are more useful and can be used in the classroom. Our drinking-water, drainage, electrical and internet-connection systems also fail to meet current standards and our schools are no longer adequate, in these times of advanced energy-saving technology and sustainable development.

Likewise, we need to draw up quality standards for learning-support materials,



which are now digital or internet-based, and the same applies to stationery and software used for management purposes. We believe that these infrastructure-related items should comply with national standards, with priority attention being given to schools in marginal-

ized or rural areas, and not just to ones in urban, commercial or high-income zones.

Without a doubt, we still have a long way to go in drawing up these updated specifications for physical infrastructure, learning-support materials and management-related

supplies, and also in improving, and also evaluating, our organizational processes. We are far from being able to evaluate our capacities and performance as participants in the education process, but the time has come to start. €

## VOICES FROM THE CONFERENCE FACING OUR CHALLENGES

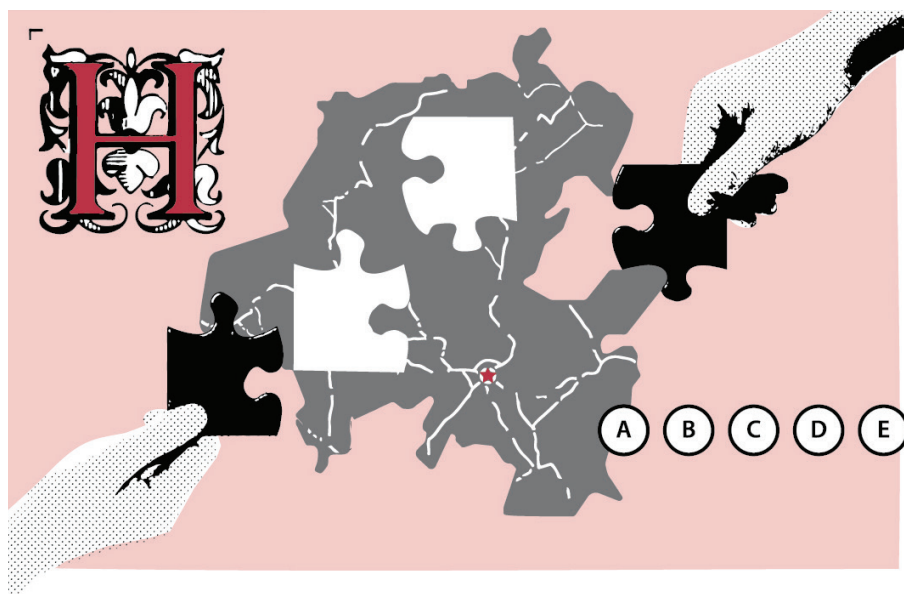
### Educational evaluation and the regionalization strategy

When we hear the word “evaluation”, we automatically associate it with the word “examination”, which in turn we associate with the word “punishment”, or, in the best of cases, with the word “prize”. Given the above, in the context of the Educational Reform we need to instill in people a widespread openness to change, as we in the education system in Hidalgo have managed to do by adopting effective strategies at the local level.

MIGUEL ÁNGEL CUATEPOTZO  
Former Minister of Education  
for the State of Hidalgo  
[miguel.cuatepotzo@seph.gob.mx](mailto:miguel.cuatepotzo@seph.gob.mx)

Since evaluation is equated nowadays with improvement, we need not only to accustom ourselves to positively associating evaluation with improvement or training, but also to design valid, reliable, relevant types of evaluation and evaluation instruments to suit our purposes. Of course, evaluation of the education system, which pivots around high-quality and equity, must set out to improve education.

Hidalgo faces many challenges with regard to education, and the realities in our state demand effective decisions and planning that take into account the latter’s special nature and the idiosyncrasies of its inhabitants, how our students learn and how



our teachers teach, along with the nature and needs of our schools and communities. Undoubtedly, all these challenges require us to adapt to an environment whose diversity makes it both rich and complex.

#### Regionalization

Covering 20,846 km<sup>2</sup>, Hidalgo is characterized by rich diversity and has a unique identity due to its biodiversity and many different cultures.

In this context, in order to achieve true equity, we need an education system that matches its complicated surroundings. Hence, the regionalization of our educational services has been, and remains, a pressing need, and, to satisfy this need, the Ministry of Public Education of the State of Hidalgo (Spanish acronym: SEPH) has proposed the creation of elementary-level educational regions so as to change school management and teacher supervision at the regional level and thus continue to raise the quality of education and make it more equitable.

We are responding to the specific needs of the schools in our state, and supporting them, by implementing different models and levels, bringing management services closer and refocusing and increasing support for teachers, groups, school principals, zonal advisers and supervisors, in a context where federal and local programs are being aligned with each other.

Hence, the project for transforming the management of education has been implemented by setting up 30 regions, bringing together populations that are close to each other and share cultural, socioeconomic and environmental features, so as to group them into clusters that are called “regional territories”.

In this way, we are bringing our educational services nearer to their users so as to respond to the changes that are occurring and satisfy the real needs of our people, making the school the focal point.

A central role in this system is played by the Regional Technical Committee (Spanish

acronym: CTR), the main responsibilities of which are to support the Technical School Committees (Spanish acronym: CTE), monitor progress in the efforts to raise the quality of education, and evaluate results and processes in the regional schools in order to come up with regional strategies. In this way, the CTR is detecting the needs and demands in the regional schools, as expressed in the “Improvement Paths” and the “Intervention Paths”, and formulating strategies aimed at achieving the short-, medium- and long-term aims set forth in the Regional Plan, placing emphasis on linkage between educational levels and institutions, and also between projects and programs that are considered necessary for supporting the said processes.

### The school and social participation

#### Full-time schools

Hidalgo has 588 full-time schools (Spanish acronym: ETC), of which 402 have joined the “Crusade against Hunger” and hence have food services. In the context of the current Educational Reform, we recently organized a series of three one-day workshops for principals of full-time schools, chairmen of Social Participation Committees (Spanish acronym: CPS), Assistant Regional Service directors, regional liaisons, teams in charge of Ministry-of-Education federal and state-level programs, ETC coordinators, supervisors, sector heads and teachers – in all, a total of 1,500 people.

The said workshops were held in order to learn more about how ETCs work, come up with a comprehensive educational proposal, and discuss autonomous management, with a view to promoting harmonious relations in schools, liaison with parents, leadership, and strong relationships among institutions.

The ETCs are the ones that have run pilot projects in regionalization, so as to propose actions that are in line with the idiosyncrasies of each region, and, of course, of each school.

#### Schools devoted to excellence (Spanish: *Escuelas de Excelencia*) in order to reduce educational backwardness

Hidalgo has 535 schools enrolled in the Schools-devoted-to-Excellence Program for Reducing Educational Backwardness (Spanish acronym: PEEARE). Since September of 2013, the General State Education

Department has visited the different areas where these schools operate in order to hold information-gathering meetings, strengthen ties with the Social Participation Committee (CPS), and offer consultancies on the compilation of technical files. In addition, round tables have been held throughout the state, during which the technical team gave advice and checked the schools’ technical files, taking photographs and gathering photographs for the first stage of the PEEARE platform, which is periodically updated.

It should be stressed that the said Program is designed to strengthen the administrative autonomy mandated in the Educational Reform. For the first time, federal resources are reaching the schools and school-supervising entities directly, making it possible to respond to needs and shortages in a realistic way, based on the schools’ priorities.

The main emphasis of this process is on transforming infrastructures, since it involves a high level of public participation and both curricular and extracurricular activities that foster more effective learning, above all in very poor communities. In this way, the CPS’s in our state have become more committed, turning into mainstays for the carrying out of planned projects, and have also ensured that the resources they received were properly used.

#### Ongoing training in the context of the regionalization strategy

We have backed up the federal initiatives for the training of qualified technical-pedagogic advisors as part of the Technical Support Service to Schools (Spanish acronym: SATE) via conferences, study circles and workshops on the design of training courses aimed at honing the skills, and strengthening the competencies, of these advisors, who help to boost educational quality in their zones or regions.

Some regionalized entities have been set up in our state to provide this service; in each of the 30 regions, the CTR has formed supervisory teams that tackle specific problems, involving both teaching and training needs, based on a regional plan that is linked to the CTE Improvement Path. In this regard, work based on self-management, is going ahead in the regions, which have planned and initiated their own training processes.

In another instance, all the higher-education institutions, the agencies of the SEPH itself, and those coordinating federal

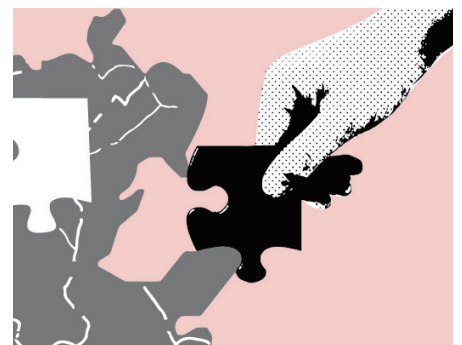
programs, were invited to take part in a public competition, submitting their proposals to be analyzed and approved in accordance with the General Law Governing the Teaching Service (Spanish acronym: LGSPD). To date, we have 104 proposals that will make up the State Teacher Training Catalogue subject to an approval process that will take place soon.

Also, “peaceful school coexistence with a gender perspective” has become very pertinent in our zone. Stemming from a diagnostic study carried out by UNESCO, ten general primary schools in eight municipalities have been chosen to take part in a project called “Citizens for coexistence and peace in schools” run by the Columbian Foundation, “*Convenio Andrés Bello*”. Hidalgo has been chosen to represent Mexico in this event, as part of the pro-coexistence activities. Currently, we are working on a guide to peaceful, democratic coexistence in schools, to be used by 313 schools and 4,000 teachers in our state.

#### Moving from experience to evaluation

In order to achieve comprehensive improvement of the education service in our state, we need to build an adequate infrastructure, make resources available for fostering autonomous management, strengthen entities such as the CTE and the CPS that promote reflection about teaching and reliable diagnoses aimed at coming up with solutions to specific problems, provide relevant teacher-training and improvement courses, and come up with a multidimensional strategy for solving problems in specific contexts.

All of the aforesaid things, along with other actions, will enable us to create and carry out a technically sound, essentially formative evaluation that is widely applied, embraces diversity, and fosters educational improvement. €



## VOICES FROM THE CONFERENCE

## OUR VOICE

## The challenge of using the information stemming from educational evaluations in Mexico

Using information means understanding and finding meaning in it in our particular context. Having said this, the author proposes four possible scenarios for information use as a basis for reflecting on the current situation and discussing how we might use information to help improve the National Education System.

TERESA BRACHO GONZÁLEZ

Member of the Board of Governors of the INEE  
 teresa.bracho@inee.edu.mx

In order to link evaluation and improvement, we need to deliberately and explicitly encourage the people who have to use evaluation results to do so in an appropriate and useful way, using them as input to their activities, grading and other scales, and also to inform their actions and decisions.

### Dilemmas with regard to the use of information

The technical robustness of evaluations depends on how the information is produced, how rigorous the measuring instruments and the analysis and interpretation are, and the extent to which the integrity and security of the data are assured, as well as other factors affecting the latter's validity and pertinence. Concern about the aforesaid issues has given rise to the consideration of other problems which –in a context of authentically *public* policy that involves different protagonists- we need to be aware of if we are to solve them. The following are just some of the said issues.

*The central problem of educational evaluation consists in producing suitable information.* It is wrongly believed that, once this problem has been solved (via solid databases, robust information systems, technically unimpeachable mechanisms for gathering and processing data) potential consumers will automatically use the evaluation results.

*Also it is wrongly assumed that the information will, in most cases, be used appropriately and effectively,* since the real problem consists in developing instruments and processes that ensure scientific rigor – i.e. assure the quality of the information.

*Furthermore, it is wrongly thought that the mere dissemination of the information* (via printed media or Internet) will automatically make it accessible to users, thus ensuring that it is understood and resulting in its being used effectively by various consumers in their everyday work.

In contrast to the above, we need to look at information use from another angle. The main thrust of this short article is that we need to see the use of information stemming from evaluations as a problem per se –just as important as the mode of production of the said information- rather than as a happy consequence that ensues more or less spontaneously. In short, we still have some big problems to solve, having to do with both supply and demand, in this regard, as several specialists in the field agree.

There is considerable leeway with regard to the *supplying* of evaluation in order to engender more effective strategies that generate more, and more reflective, *demand*,

including more consumers –by adapting the contexts, media and messages to satisfy different interest groups- so as to foster better understanding, and positive appropriation, of the said information by all its potential users.

One of the central premises of this article is that the use of evaluation results is the responsibility not only of the users, but also of those who produce the said evaluations.

### The uses of the information yielded by educational evaluation

With regard to the use of evaluations to improve education, one needs to find ways of analyzing the behavior of the protagonists in this hypothetical production scenario de. One way –though not the only one- is by dividing users into two groups: (a) those who produce the information (the *supply*), and (b) those who consume the said information (the potential or actual *demand*).

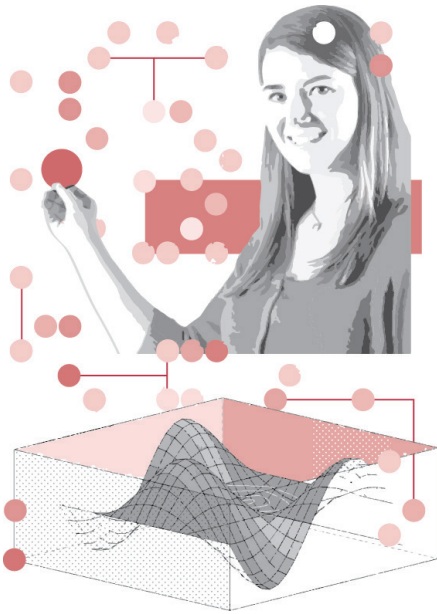
Now, based on the aforesaid definition of information use, one can distribute the behavior of information supply along a continuum one end of which is the creation of homogeneous supply and the other end of which is supply that varies according to its different consumers. Likewise, one could distribute the behavior of demand along a continuum one end of which is indifference to the said information and the other end of which is its actual use in the consumers' daily activities as members of the Education System (Spanish acronym: SEN). By examining the point where these two continuums intersect, we can identify the four possible scenarios shown in the following table:

**Table 1. Scenarios of the types of supply and demand of information stemming from educational evaluation**

		Demand	
		Use with different degrees of appropriation	Undifferentiated
Supply	Heterogeneous	<b>Scenario I</b> Heterogeneous supply Demand that uses the information	<b>Scenario II</b> Heterogeneous supply Indifferent or incipient demand
	Homogeneous	<b>Scenario III</b> Homogeneous supply Demand that uses the information	<b>Scenario IV</b> Homogeneous supply Indifferent or incipient demand

Source: Author, based on the Evaluation Capacity Development: *Framework for Action*, Guerrero (1999).





**Scenario I: Behavior of heterogeneous demand and demand that uses the information (with different degrees of appropriation).** Here we see a supply that recognizes its demand (potential or actual) as a heterogeneous universe composed of different consumer groups with different interests in the educational-evaluation results. In this regard, the behavior of this type of supply is characterized by its efforts to identify as many consumer groups as possible and generate products based on their specific needs, which leads to variation in contents, media and information-dissemination strategies according to group, resulting in specific products and different types of measurement.

The empirical evidence in Latin America shows that heterogeneous demand can arise in the following forms: as general reports on specific learning results, on associated results, or on results per school, and also as publications aimed at teachers or parents (Ravela, 2006).

In this scenario, demand –when interpreted, diagnosed and specifically addressed- would theoretically be more motivated to use the information in its different activities. However, this would continue to depend on motivational factors, interpretative capacity and applicability – the actual uses to which the information is put and whether these are positive or negative. This

scenario pertains to a strong evaluation culture - i.e. a set of shared values and principles that manifest themselves in intensive use and strong benefits.

**Scenario II: Behavior of heterogeneous supply and indifferent or incipient demand.** Here supply continues to behave based on the recognition of a potential heterogeneous demand. What changes is the demand's behavior, since, though different products exist, either there are still factors that cause potential consumers to remain indifferent to the supply per se, or the uses of the said supply remain very incipient, due, among other things, to superficial knowledge or to the fact that the information is considered unreliable or irrelevant by its users.

**Scenario III: Behavior of homogeneous supply and demand that uses.** The supply behaves differently than in the previous cases, either because the supplier is not fully aware that the demand is diverse -albeit potentially- or because it is aware of the diversity, but, for whatever reason (technical, economic, political or a combination of the three), only general or homogeneous products are created, based on the evaluations and the technical criteria of the evaluator, leading to a homogeneous supply, or even one based on the criteria of the evaluator, who is the only person with enough clout to make decisions about the information that s/he supplies. However, as far as demand is concerned, one can detect that the information is used by different consumer groups, who make requests, ask questions, or make complaints because they need the said information to be adjusted to suit their specific needs.

**Scenario IV: Behavior of homogeneous supply and indifferent or incipient demand.** The supply remains homogeneous as in Scenario 3. Here we see a supply that is indifferently received by the demand, due either to ignorance or lack of understanding, resulting in scant motivation to use the said supply in different contexts. For example, if national reports are issued that could be very appealing to federal and state-level education officials, but not necessarily to people in local departments or schools, because supervisors, school principals, teachers or parents find it hard to understand and apply the said reports in their particular contexts.

However, the indifference that arises also indicates that the consumers who are not taken into account do not complain about this neglect.

What can we infer -or not infer- from the above scenarios? Some points regarding their scope and limits are as follows:

- The model posited requires fine-tuning according to each consumer type. Which would result in the building of specific scenarios for each one of the latter. Here, space does not allow me to elaborate on this point.
- Separate analyses could be developed per consumer type, so as to come up with more than one scenario to describe specific behaviors in the case of a context such as Mexico.

The above is only a preliminary attempt at analysis and does not go into the causes of indifferent behaviors or uses -whether valid or not- in the case of demand.

The scenarios that are roughly outlined above do, however, serve as an initial means of analyzing current cases so as to posit (a) better scenario(s) for achieving the different aims of the Educational Reform.

### Conclusion: the current scenarios and desired scenario

The following questions, to which I have provided answers, might be asked with regard to the model presented above:

- (1) Q. Which scenario does Mexico currently find itself in?
  - A. Somewhere between Scenario III and Scenario IV, depending on which consumers or which protagonists one considers.
- (2) Q. Which scenario would we like to be in and why?
  - A. We'd like to be in Scenario 1.

Briefly, I would support the aforesaid answers by saying that, without a doubt, the ideal scenario for Mexico is one with heterogeneous supply where information is provided to suit the needs of different protagonists, and, above all, the needs of education officials at different levels -ranging from decision-makers to those directly

involved in management in the different zones and schools- but also to suit the needs of school communities (teachers, students and parents). Preferably, one would also supply information suitable for other consumers, such as the executive branch, organized civil society and the media, who are very interested in knowing what is happening in the area of education. The aforesaid adaptations should include help to interpret the situation at any given moment and enable the said consumers to make pertinent comparisons with regard to each particular use to which the information is put.

But such a scenario implies not only an identified heterogeneous demand, but also an *activated* one – i.e. one that uses the information because it knows it exists, and that it is reliable and suits its needs, can interpret it constructively, and has positive motives for using it as evidence aimed at supporting actions in its particular sphere of action.

One is not talking about a willingness to foster the use of information, but, rather, about creating appropriate conditions to foster use -via effective dissemination of the information and encouragement to make sense of it. This scenario would pertain to the presence of a strong evaluation culture in all those involved in the process.

However, one should stress that the creation of this active demand, and the development of suitable uses of information that will enable evaluation to redound in the improvement and transformation of the SEN, is one of the core responsibilities of evaluation itself, and should not depend on the initiative and good will of the latter's users.

## References

- Ravela, Pedro (2006), *Para comprender las evaluaciones educativas*. Fichas didácticas, Grupo de trabajo sobre estándares y evaluación, PREAL. Available at: <http://m.preal.org/detalle.asp?det=729>
- Guerrero, Pablo (1999), *Comparative Insights from Colombia, China and Indonesia*. Evaluación Capacity Development, Working Paper, Series No. 5, World Bank, Operation Evaluation Department, Washington D.C.

## VOICES FROM THE CONFERENCE

### OUR VOICE

## The challenge posed by the Educational Reform: Classroom Observations

**“It is in the classroom that planned, systematic teaching and learning take place; if we want to raise the quality of education, we have to ascertain what we can do, and how we can do it, in order to improve teaching”, says the author of this article that discusses the challenges of teacher observation and proposes a model for implementing it in the Mexican context.**



MARGARITA ZORRILLA FIERRO  
Member of the Board of the National Educational Evaluation Institute (Spanish acronym: INEE)  
[margarita.zorrilla@inee.edu.mx](mailto:margarita.zorrilla@inee.edu.mx)

**“It is not because things are difficult that we do not dare; it is because we do not dare that they are difficult.”**  
(Seneca)

### 1. Introducción

A lot of things need to be done to make the Educational Reform a reality. The National Educational Evaluation Institute (Spanish acronym: INEE) is responsible for doing the

technical work and helping the authorities to take the decisions that are needed so that the Reform is carried out in a way that produces the best possible results for everybody involved throughout the National Education System (Spanish acronym: SEN).

One of the mainstays of the Reform is the creation of a professional Teaching Service (Spanish acronym: SPD) that aspirants join based on merit, and within which they are promoted, acknowledged and also remain, based on merit, all four of the aforesaid processes being subject to evaluations that are applied to aspirants for, and occupants of, teaching, managerial or supervisory positions in the compulsory-education system in order to determine their capacities.

The yardsticks, aims, methods and instruments of the evaluation process need to be defined, and, to a considerable extent, the quality of the evaluations depends on these aspects, which should be divulged in a timely manner to all potential candidates.

The profiles, parameters and yardsticks chosen by the federal and state-level education authorities are validated by the INEE, and, based on them, the stages, methods and instruments that will be used in each evaluation process are designed.

This article will discuss the challenges posed by the classroom observation of teachers in an endeavor to help readers gain a better understanding of the said process.

### 2. The evaluation of teacher performance

In its definition of the SPD, Section xxxii of Article 4 of the General Law Governing the National Teaching Service (Spanish acronym: LGSPD) refers to a “set of activities and mechanisms for Admission to, Promotion and Acknowledgment within, and Continuance in, the government education service, and also to the promotion of ongoing training, so as to ensure that teachers and management and supervisory staff in government elementary and upper-secondary schools run by the State and its decentralized agencies possess the requisite knowledge and competencies”.

The SPD's processes are linked to the evaluation ones, and hence aspirants for entry to, or promotion or continuance in, managerial or supervisory posts, or positions

as technical teaching advisors, must take competitive examinations, with the evaluation of performance being the basis for decisions regarding certain types of promotion, acknowledgment or continuance.

The LGSPD stipulates the features that must be taken into account when designing and carrying out evaluations of teacher performance.

**Article 14.** To achieve the aims of the SPD, profiles, parameters and indices must be developed that serve as yardsticks for good professional practice, to which end it is necessary that the said profiles, parameters and indices, at the very least, make it possible to:

- I. Have a General Framework regarding High-quality Education and minimum standards for teaching and schools, with which education authorities, decentralized agencies and members of the SPD must comply.
- II. Define the main tasks carried out respectively by teachers, managers and supervisors, including, in the case of teachers, the planning and mastery of contents and the classroom atmosphere, teaching practices, evaluation, the achievement of student learning, in-school cooperation and communication with parents or guardians.
- III. Identify basic criteria for the performance of members of the Professional Teaching Service in diverse social and cultural environments, in order to achieve adequate learning outcomes and develop all their students within a context of inclusion.
- IV. Take into account compliance with calendars and the proper use of teaching time.
- V. Establish competency levels for each of the categories applying to teachers, managers and supervisors, so that the said staff, schools, school zones, and, in general, the different people responsible for providing education in the compulsory system may have access to yardsticks pertaining to ongoing improvement and compliance with the requisite profiles, parameters and yardsticks, which shall be reviewed from time to time.

As the reader can see, Section II of the article of the LGSPD cited above stipulates a set of typical features of teaching, making particular reference to classroom atmo-

sphere, teaching practices and student evaluation. All these things relate to teaching in the classroom, and the following section of this article will talk about how to observe the aforesaid activity.

### 3. Observing teaching in the classroom

Since educational quality mainly depends on what happens in the classroom, which is where planned, systematic teaching and learning take place, in order to raise the quality of education we must find ways to improve the teaching in all our classrooms.

Classroom-observation techniques and methods have been being developed for several decades now, focusing, among other things, on features such as classroom interaction, the use of teaching materials, time-planning, and the use of different methodologies depending on the subject taught.

With regard to the performance evaluation stipulated in the LGSPD, various experts have agreed that one needs to observe teachers *in situ*, and some people even affirm that this is the only way to evaluate teacher performance.

However, it is one thing to insist that the observation of teacher performance in the classroom is crucial, and another to actually carry out *in situ* observations, using standardized mechanisms, protocols and instruments and qualified observers, of the almost 1.5 million teachers working in compulsory education in Mexico's elementary and upper-secondary schools. Indeed, the task of observation is even more complex, since, in addition to considering all of the aforesaid aspects, we need to decide which aspects of classroom teaching to observe, and why, how, for how long and how often.

The different teachers, school principals and supervisors that I have consulted have concurred that classroom observation should focus on what is essential – i.e. on student-student interaction, student-teacher interaction, knowledge, resources and teaching materials.

Expressed in this way, observing and evaluating teachers in the classroom would seem to be a simple task. However, nothing could be further from the truth, since classroom observation is a complex task – an art that is based on knowledge of different subjects and on intuition, as well as on certain competencies, attitudes and values that I will discuss below.

### 4. Methodological approaches to classroom observation

To give the reader some idea of how one might build a model for classroom observation, I will mention two approaches to the said task that are currently being used. I wish to stress that these are not the only possible approaches, though I believe that they are useful examples of how one might proceed.

The first of these approaches is called the Stallings Observation System (SOS for short), after its inventor, Jane Stallings, who teaches at Vanderbilt University in the USA, and, developed in the 1960's, has mainly been used to evaluate the teacher's use of time during secondary- and primary-level classes, as well as the amount of resources used by teachers, students or both in such classes. This approach has recently become popular again as a result of research done by education experts from the World Bank.

The Department Responsible for Federal Education Services in Mexico City (Spanish acronym: AFSEDF) has adopted the Stallings system and is applying it via school supervisors who use it during their visits to classrooms. Information about the said system and its use by supervisors in AFSEDF classrooms can be found in: *Class observation based on the Stallings system. Manual and User's Guide* (2011): <http://goo.gl/CZ6syF>.

The aforesaid Stallings system, which focuses on the teacher's use of time, is useful for some things, but not for others. Undoubtedly, any teacher will find it useful to know how s/he apportions time during his/her class, what things s/he emphasizes, what s/he needs to pay more attention to, and so on. Depending on the role assigned to the observations within their system, the managers of schools can establish policies that lead to better use of teaching time.

More recently, another classroom-observation system called the Classroom Assessment Scoring System (CLASS, for short) began to be used to evaluate the quality of classroom interactions for the purpose of improving teaching. The said system comprises the three aspects of a) emotional support, b) classroom organization, and a) pedagogic support, which are respectively defined as follows:

- a) emotional support: classroom atmosphere (inter-personal relations), teacher sensitivity and ability to see things from the students' point of view.



- b) classroom organization: class control, productivity (use of teaching time) and teaching strategies.
- c) support of learning: development of concepts, quality of feedback and effective use of language (communication, in-depth explanation, advanced language).

There are scales for each of the previously mentioned aspects, which can be found in the Guide to this system. The observer, who should be trained and certified, scores the teacher observed in the aforesaid 3 areas, on a scale of 1 (minimum) to 7 (maximum), so that, when the observation -which is carried out in accordance with set rules- has finished, the observer has a comprehensive picture of the lesson observed.

As the reader will appreciate, both the SOS and the CLASS system have specific aims, as well as including a manual, scales and protocols for the purpose of carrying

out the observation and organizing the information obtained.

The above simple description, along with the comments in the paragraphs preceding it, should serve to give the reader an idea of what is meant by classroom observation.

### 5. Coming up with an approach for classroom observation

In Mexico, as in many other countries, we aspire to increase teachers' efficiency by improving their pedagogical practices. Since we must make our aspirations reality, we are developing an approach to classroom observation that takes stock of the following factors:

- a) Classroom observation is strictly formative and should only be used to improve teaching.
- b) Feedback must be given to teachers immediately after the observation.
- c) We need to develop a Mexican model

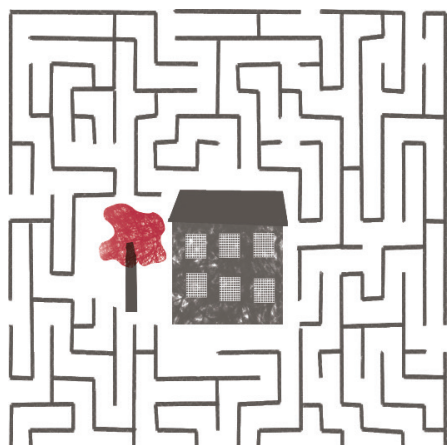
that suits our cultural makeup and education system.

- d) The model needs to be versatile so that it can be used with pre-school, primary, lower-secondary and upper-secondary teachers.
- e) The model needs to be adaptable for use by the tutors of teachers who have just joined the SPD, by technical teaching advisors, and by school principals and supervisors.

My colleagues and I are enthusiastically working to create an observation system that meets the above criteria and is of optimal help to all those involved in the Mexican education system. We have a long way to go to convert the aims sketched out above into a useful instrument, and great deal of patience, knowledge and staying power will be needed to make that journey. €

## ROADMAP

## The National Learning-Evaluation Plan (PLANEA): the Involvement of Different Participants in the Education Process in the Evaluation of Results



After the Education Reform, the evaluation of student achievement in Mexican elementary and upper-secondary schools took a new turn. The National Evaluation of Academic Achievement in Schools and the Examinations of Educational Quality and Achievement were replaced by PLANEA, which contemplates the involvement of different participants in the education process in the analysis of results. Below, we present the modes for the implementation of the said plan and reflect on the challenges that they pose.

ANDRÉS SÁNCHEZ MOGUEL<sup>1</sup>  
 General Director for the Evaluation of Educational Results of the Unit for the Evaluation of the National Education System of the INEE  
[andres.sanchez@inee.edu.mx](mailto:andres.sanchez@inee.edu.mx)

In early 2015, the National Institute for Educational Evaluation (Spanish acronym: INEE) unveiled the National

Learning-Evaluation Plan (Spanish acronym: PLANEA), which includes a set of tests for students in elementary and upper-secondary schools, for the purposes of obtaining results to be used to improve the education system.

### Types of evaluation activity

It is a good idea to begin by recalling the main features associated with evaluation. Preskill and Russ-Eft (2004) comment that, according to various definitions, evaluation is a systematic, planned process, with specific aims, that involves data gathering, increases knowledge and leads to better decision-making based on value judgements about what is evaluated, and is felt to be implicitly or explicitly useful for understanding what we do and how it effects reality.<sup>2</sup>

When we flesh out these ideas in the context of the PLANEA, it is clear that implementing tests (i.e. gathering data) and analyzing and publishing their results do not constitute a complete evaluation process; rather, the core activity consists in making value judgements based on the data, and it does not become meaningful unless it increases our understanding of

what education is and leads to improvement. The aforesaid value judgements do not have to be centralized, and the PLANEA requires the participation of different people involved in education, as stated in one of the principles that governs the INEE's activities: "Evaluation implies the participation of society and, above all, of those involved in education (i.e. authorities, students and teachers)".<sup>3</sup>

### The aims of the PLANEA

The overall aim of the PLANEA is to ascertain the extent to which students acquire a set of key learnings at the different levels of compulsory education. Its specific aims are to:

- a) inform society about the state of education in terms of student learning outcomes.
- b) provide relevant, useful information for the monitoring, planning, programming and running of the education system and its schools.
- c) provide contextualized information<sup>4</sup> for improving teaching in schools.
- d) help to develop guidelines for improving education via relevant information about educational achievements and their contexts.<sup>5</sup>

### Modes of Implementation

a) *Evaluation referred to the National Education System:*<sup>6</sup> These tests are implemented every four years with representative samples of students in the final year of elementary school, making it possible to evaluate a wide range of key curricular learnings.<sup>7</sup> This interval suffices for monitoring changes in the education system, which do not tend to occur over short periods, and the tests provide the education authorities with pertinent information for monitoring, planning, programming and running the education system, and allow society to verify the state of education. The INEE will furnish the education authorities and the general public with reports of nation-wide and state-level results, presented in context.

b) *Evaluation referred to the schools:* These tests evaluate a limited range of key curricular learnings and will be implemented annually by the Ministry of Education, in collaboration with the state education authorities, in every school in the country, in the final years of primary, lower-secondary and upper-secondary school. Besides providing

the education authorities with information, just like the multiple-choice tests, they furnish information for improving educational processes in schools, and each of the latter will receive a report on the results, accompanied by information that will enable them to suitably contextualize them.

c) *Diagnostic evaluation census:* These tests will be given to fourth-grade primary-school teachers at the start of their courses, being implemented and analyzed by the said teachers, for the purpose of class planning, in accordance with a guide. They provide information for the purpose of improving teaching in schools and, unlike the other types of test, are administered by the teachers, who are responsible for gathering and analyzing the data, making contextualized value judgments about them, and using them to plan their teaching.

### People involved in reflecting on PLANEA results and using them

People from different areas are expected to play a part in analyzing and reflecting on PLANEA results and making value judgments about them. Placing these activities within the context of the Knowledge Utilization Models,<sup>8</sup> we consider that it is necessary to turn the resultant information into knowledge, sharing the latter with whoever can use it to make concrete improvements in education – i.e. not merely transferring information about the results, but, rather, promoting collective analysis of them.

The PLANEA was designed within the context of the new legal framework that governs the INEE and other educational institutions, and establishes lines of direct, ongoing communication among authorities at the federal and state levels, other participants in the educational process from the public and private sectors, and society in general. The intention is to set up mechanisms for disseminating, and reflecting on the results.

In Table 1 below, following Paul Hood, we ask a series of questions the answers to which will help us to identify the disagreements about -and challenges inherent in- the dissemination, understanding and use of the results of the different PLANEA variants:

### Final observations

When we talk about disseminating the examination results, we are assuming that the system for administering the PLANEA

includes technical features that ensure the high quality of the information. This plan includes a series of processes aimed at preventing the results from being distorted, and, above all, from being inflated, including emphatic warnings that the PLANEA must not be used to evaluate either teachers or schools, or to apportion blame, but, rather, in order to provide information as a basis for serious, informed reflection for the purpose of making changes, as well as specific examination protocols, supervision of implementation, statistical verification of result consistency, and the use of external evaluators – i.e. ones from outside the school.

One of the temptations with evaluations such as the PLANEA is to develop school rankings. We consider that these do little to foster improvement and can overshadow the more productive kinds of analyses that school communities can achieve by discussing and interpreting their results and coming up with clear, contextually relevant proposals for supporting the learning of their students, which is, after all, the main purpose of the whole exercise. €

- 1 The author wishes to thank Mariana Zúñiga García and Carolina Cárdenas Camacho, both from the UESEN's Department for the Evaluation of Educational Results, for checking this text.
- 2 Preskill, H., Russ-Eft, D. (2004). Overview of Evaluation. In: *Building Evaluation Capacity*. Thousand Oaks, California: Sage Publications.
- 3 Schmelkes, S. (2014). *Reunión Anual de Evaluación 2014. Cinco principios del INEE*.
- 4 In the case of the PLANEA, *contextualizing* means identifying the circumstances that surround learning in order to better understand the results.
- 5 The perceptive reader will notice that no further consideration is given, in this article, to this particular aim of the PLANEA. The aim of helping to develop guidelines for improving education differs from the previous ones at least in the following ways: (a) the users of the information belong to the INEE itself, exercising the powers vouchsafed to them under law in order to issue guidelines that foster the making of decisions aimed at improving the quality of education and making it fairer (Article 47 of the Law Governing the INEE), and (b) the information about educational results that can be used for this purpose

Table 1. PLANEA variants

	Evaluation referred to the National Education System	Evaluation referred to the schools	Diagnostic evaluation census
¿What kind of information do we want to use?	Updated every four years. Applied at the national and state levels and over a large range of class levels.	Updated annually. Analyzed at the school level.	Updated annually. Aggregated at the group level or disaggregated per individual student.
Who are the main users of this information?	National and state-level education authorities.	The school community, school supervisors and municipal authorities.	Class teachers and the school community.
What do we know or assume about these users?	Decisions to be taken at the macro educational-policy level and subpopulations with low scores identified.	The information should supplement what is already known about the school and the results should be reflected on in context. Skills are required in order to analyze this information.	The information produced by the school itself should be taken, along with that obtained from this evaluation, as input for planning classroom teaching. Skills are required in order to analyze this information.
For what purpose is it used, and what results are sought.	Evidence-based decision-making.	School community encouraged to reflect on, and discuss, its circumstances and how they relate to the learnings, in order to develop contextualized improvement strategies.	Provision of a panorama of learnings that entry-point students should have acquired and identify those that they lack in order to come up with remedial strategies.
What must we know about the context of this information?	The contextual variables that are most closely associated with the results.	The explicit information that is provided should be taken, along with the tacit information possessed by the school community, in order to extract meaning and facilitate its use. <sup>9</sup>	The teacher should notice when the group in general, or individual students, need to acquire knowledge that is already supposed to have been acquired.
What are the implications of the answers to the above questions for achieving better use of the results?	Besides furnishing timely information to decision-makers, the INEE should develop a systematic dialogue with them in order to reach shared interpretations that foster the development of coherent joint actions.	The SEP must provide the school community with support and high-quality means of dissemination in order to enable it to analyze the results. The INEE should follow up on the use of the information furnished and recommend the changes that need to be made to the reports so as to adjust the information that is given to the schools.	The teachers should receive clear, detailed information from the SEP about the diagnostic instruments and their use, the analysis of the results, and the procurement of useful information. The INEE should monitor this process and suggest adjustments. When the activity is successful, it could be extended to other school levels or subjects. This variant could serve as a foundation for enhancing teachers' ability to design and implement formative-evaluation strategies, since there is no substitute for teachers' evaluations of their students.

Source: Author, DGERE-UESEN.

pertains to each of the disaggregation levels mentioned and needs to be combined with information from the Institute's other sources, with local evaluations, and with studies of the components and results of the National Education System, a task so complicated that there is not enough space in this article to discuss it.

6 "Since the sample-based evaluations evaluate a broad range of curriculum contents, they consist of a series of *multiple-choice* sections, of which each student answers only one, revealing his/her knowledge about this broad content set". (Andrade, E,

*et al.* (2015). Sample PLANEA design. Internal INEE document).

7 Incorporating degrees for this type of evaluation will be staggered: in 2015, the first application will be for sixth grade and ninth grade is made and the second will be in 2019. In 2017, the first application for pre-school third and last grade of high school education is made, and the second will be in 2021. For this type terminals degrees of each level are evaluated, because they provide enough information for evaluating the entire education system.

8 The Knowledge Utilization Models are theoretical structures whose purpose is to better understand the relationship between knowledge production, knowledge transfer and knowledge use. cf. Hood, P. (2002). *Perspectives on Knowledge Utilization in Education*. WestEd.

9 I refer to the "knowledge spiral" and the concepts of explicit and tacit knowledge posited by Nonaka (1994) in his article, A dynamic theory of *organizational knowledge creation*, published in *Organization Science*, Vol. 5, No. 1, 14-37.



## Curriculum evaluation

Educational reform inevitably implies thinking about the educational syllabuses and programs pertaining to the levels affected by it. How, therefore, can we tackle the task of evaluating our country's elementary- and lower-secondary-level curriculums? Below, the authors propose some guidelines for moving ahead with the linking of two educational reforms -that of institutions and that of curricula- which do not always go hand-in-hand.

REBECA REYNOSO ANGULO  
Director for the Evaluation  
of the Educational Prospectus  
[rreynoso@inee.edu.mx](mailto:rreynoso@inee.edu.mx)

RAQUEL AHUJA SÁNCHEZ  
General Director for the Evaluation  
of the Educational Prospectus  
[rahuja@inee.edu.mx](mailto:rahuja@inee.edu.mx)

Unit for the Evaluation  
of the National Education System

While learning outcomes tend to be at the center of all discussions about what is currently happening in education, nevertheless curricular issues are a decisive factor in the provision of high-quality classes. Since the curriculum is one of the main frameworks for underpinning and defining working conditions both in our schools and in our education system (Gimeno, 2007), and also a feature that tends to be targeted for reform when efforts are made to improve the range of courses available, it is pertinent that it be evaluated so as to identify those aspects that help or hinder our endeavors to ensure that everybody has access to high-quality education.

One of the challenges faced by the National Education System (Spanish acronym: SEN) concerns the making of decisions about the learnings that we want to foster

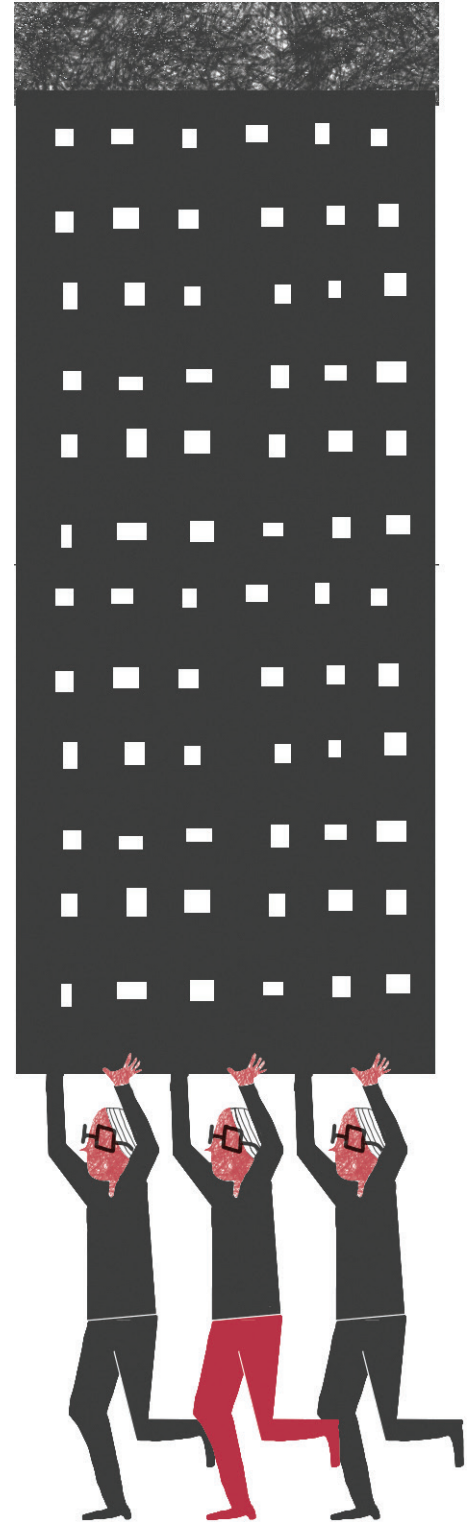
in our schools so that students can develop personally and participate in the society to which they belong. These learnings, along with the guiding principles of the educational process, are formally enunciated in the curriculum (Coll & Martín, 2006).

Our country's experience of curriculum evaluation is basically in the area of higher education and has involved the accreditation of curriculums. Recently, institutional evaluation has been promoted, also for purposes of accreditation, in the upper-secondary-level National Baccalaureate System, with one of the criteria being the adoption of the Common Curricular Framework that has been in place since 2008. While there have been some scattered efforts in the area of elementary education, there is no institutional policy. Although the Ministry of Education has carried out internal evaluations in order to monitor the progress achieved with regard to the curricular reforms, prescribed textbooks so as to assure that they are in line with the curricula, and asked both Mexican and foreign external bodies to carry out evaluations of the curriculum, these efforts have not been systematic.

Among other things, the National Institute for Educational Evaluation (Spanish acronym: INEE) has studied learning outcomes, determining that the curriculum is effective, but, until now, no evaluations had been developed that focused on the latter. Since 2014, these evaluations have formed a part of the work that the INEE does via the Department for the Evaluation of Educational Contents and Methods of the General Department for the Evaluation of the Prospectus (Spanish acronym: DGEOE). These instruments focus on the pertinence of course and curriculum contents, on the programs used in classrooms, on the educational methodology used during compulsory education, and on the opinions that the different people involved in the process have about the latter.

### Why evaluate the curriculum?

Article 3 of the Mexican Constitution stipulates that student learning outcomes are the central concern of education and the only *raison d'être* of schooling, and asserts that, in order to ensure the high quality of compulsory education, teaching materials



and methods, among other things, must be conducive to maximum student achievement, having to do not only with learning outcomes, but also with the overall school environment. For its part, Article 10 of the General Education Law (Spanish acronym: LGE) mandates that teaching curriculums and materials must be key components of the National Education System.

Furthermore, the INEE's evaluations are centered on vouchsafing the right to education that is enshrined in the Mexican Constitution, and on high educational quality.

### **A model for evaluating curriculum quality with a right-to-education focus**

The right-to-education model posited by Tomasevski (2004) consists of the four components of *availability*, *accessibility*, *adaptability* and *acceptability*. In order to vouchsafe the right to education, a curriculum must be *adaptable*, with its contents and formats being open to modification by teachers and schools in accordance with their students' makeup and environment – i.e. what it offers must be meaningful and relevant to the school's pupils. This means also taking stock of the curriculum's *acceptability* – i.e. the extent to which it allows students to feel that they play a central role in their own learning, that they are a part of their school, and that the said curriculum belongs to them, furthering their interests and being of use in their present and future lives.

Article 5 of the Law Governing the INEE defines educational quality as “the quality of an education system, consisting of relevance, pertinence, equity, efficiency, effectiveness, impact and sufficiency”. In this regard, evaluating a curriculum's quality starting at the design stage means assessing: its *pertinence*, in terms of whether its methodological base, contents, and approach to teaching-learning and evaluation take the students' make-up and needs into account; its *relevance* in terms of whether its aims and contents satisfy the needs of society; its *sufficiency* in terms of whether it suffices to achieve the desired educational goals; its capacity to foster *equity* by taking stock of, and seeking to rectify, inequalities stemming from the students' social background or personal situation; its *efficiency* in terms of whether the proposed resources, the tasks set and times assigned for completing them, and the methods of evaluation, are appropriate for fulfilling its aims; its *efficacy*

in terms of whether, as a whole, it succeeds in fulfilling learning aims, and; its *impact* in terms of whether it fosters long-term learning that enables people to develop both individually and in their family, working, social and civic lives.

Curriculum evaluations also set out to take stock of *consistency* and *congruence* – i.e. to ascertain whether the different parts of the curriculum are coherent and lasting, and relate to the core items (e.g. the conception of learning, the epistemological unfolding of the discipline), and also whether they are conducive to the fulfillment of the learning aims in each part of the curriculum and in line with those of the other curriculum subjects), since the aforesaid things are crucial to proper implementation.

If the curriculum is to serve as a pivotal tool for educational practice and to be a key factor in the making of decisions about school staff and resources (teachers, time allotment, teaching resources and infrastructure, to name just a few) – since it is via these things, and especially the classroom, that teaching-learning takes place (Díaz-Barriga, 2003) – then it is essential that it be evaluated in order to find out what is happening in the school and propose improvements. By observing what is taught in the classroom, identifying the endeavors to promote learning and seeing how curricular stipulations are put into practice, we can obtain key information about how pertinent and relevant education is. The opinions of students, teachers, administrative staff and parents about what is meant to be learned in schools and what is actually learned there enable us to ascertain whether the education offered in the said schools is fulfilling the needs, and furthering the interests, of those who attend them.

However, in order to design a curriculum one needs to be clear about educational aims, specifying the basic target learnings and describing how to achieve them. Consideration of this relationship between content and how to teach it inevitably has an impact on other aspects of teaching and leads to decisions as to which knowledge enables information to be assimilated or is essential for making students think. It also implies endeavoring to find out which types of teaching foster which kinds of learning. Selecting given contents and ways of approaching them implies conceptions about the subjects taught and how they affect the process, assessment during the teaching

process, and the kinds of materials and other resources needed (Díaz-Barriga, 2013). These ingredients of the curriculum are prescribed not only by the education system, but also in the course of the interaction that takes place in the different places where education occurs – above all in the classroom.

### **What focus should curriculum evaluation focus on?**

The curriculum evaluation that is being promoted by the INEE contemplates a flexible model that takes stock of the three curricular dimensions of design, implantation and impact. This implies designing ad hoc models in answer to questions about evaluation and requirements for specific information, and the touchstone for the building of this model is entitlement to high-quality education, the vouchsafing of which all evaluation should help to ensure.

Curriculum design is seen as a process whereby ideal criteria for education and its foundations, aims, organization, contents, approaches and evaluation are laid down. Curriculum evaluation implies taking stock not only of the study plan and the programs pertaining to all the school levels, including the areas of the curriculum pertaining to each of the said levels, but also of the curricular frameworks for each type of education and for specific types of population, as well as of the materials that stem from the aforesaid considerations and make it possible to implement the curriculum in different contexts. The INEE has embarked on a two-fold line of curriculum evaluation, covering (a) all the fields of study pertaining to compulsory education, with the Natural-Science and Mathematics curriculums currently being studied and compared with those used in other countries, and (b) general studies into teaching principles and trends, with one such study having begun recently.

The second aspect to be considered is the implementation of the formal curriculum, known as the “operative” or “implemented curriculum” (Posner, 2005). It is interesting to witness those teacher-student interactions that enable us to ascertain what is taught in the classroom and how, in order to identify the relationship with the formal curriculum, ascertain whether the learning experiences are high-quality ones and hence promote the said relationship, and find out which of the factors or conditions that converge in the teaching-learning process

determine learning outcomes. As part of these actions, the INEE is carrying out both qualitative and quantitative studies of curriculum implementation in schools, as well as studies of the institutional-management aspects of implementation.

As far as the fourth aspect is concerned, we will endeavor to evaluate the impact of the curriculum in terms of real or perceived positive or negative effects on the current or future lives of people who have completed a learning path, in order to help generate data about the individual, job-related and social effects of education and the benefits that educators and other people perceive to have resulted from a given study plan.

Work has also been done on the conceptual and methodological aspects of curriculum evaluation, with a first step being the formation of a Network of Curriculum Evaluators in which we are endeavoring to include all the academics and researchers who are interested in the subject so as to generate synergies that will result in more vigorous and systematic development of the area.

### Challenges at the classroom level

In order to come up with information about the extent to which the right to high-quality education is being vouchsafed, we must take stock of the different areas or levels of

realization where educational intentions become reality and the curriculum takes shape. At the *social* level, a society's aspirations, outlined and expressed in the aims of educational institutions, are realized. At the *institutional* level, the said aims take the shape of a specific curriculum project in accordance with the type and form of educational service in question, with the principles, aims and tendencies that will guide educational praxis being formally expressed in official documents and curricula. In the case of elementary education, while there is a nation-wide curriculum, there are different curricular frameworks for specific populations, which must be taken into account in order to reach conclusions about this level of realization, while the challenge in the area of lower-secondary education is bigger due to the large number of different study plans and programs pertaining to the different subsystems that this level comprises. At the *school* level, educators form part of the institutional project, including both its internal and external context. Here the challenge for evaluation consists in taking stock of contextual differences in order to come up with information about the extent to which the right to high-quality education is vouchsafed. Finally, at the *classroom* level, the institutional project is realized via inter-

pretation and reformulation by both teachers and students. Evaluation at this level means coming up with information about what happens in the classroom, which is an enormous challenge. €

### References

- Coll, C. and Martín, E. (2006). *Vigencia del debate curricular. Aprendizajes básicos, competencias y estándares*. Mexico. SEP.
- Díaz Barriga, Á. (2003). El currículo. Tensiones conceptuales y prácticas. *Revista Electrónica de Investigación Educativa*. Vol 5. No. 2, Mexico. Consulted on April 20<sup>th</sup>, 2015, at <http://bit.ly/1fUcVMk>
- Díaz-Barriga, Á. (Coord.). (2013). *La investigación curricular en México 2002-2011*. Mexico: ANUIES-COMIE.
- Gimeno, J. (2007). "Aproximación al concepto de currículum". In: *El currículum: una reflexión sobre la práctica*. Spain. Morata.
- Posner, G. (2005). *Análisis del currículo*. Mexico. McGraw Hill.
- Tomasevski, K. (2004). *El asalto a la educación*. Buenos Aires. University of Buenos Aires.

### ROADMAP

## A new look at the diffusion and fostering of the evaluation culture

For the INEE, fostering an educational-evaluation culture implies creating conditions for influencing those who make decisions about the system, providing them with pertinent, useful information that enables them to identify challenges, and disseminating information about the progress achieved in fulfilling the legal mandate to provide universal, high-quality, equitable education.

JOSÉ LUIS GUTIÉRREZ ESPÍNDOLA  
Director of Evaluation-Usage and Promotion  
[jgutierrez@inee.edu.mx](mailto:jgutierrez@inee.edu.mx)

ANNETTE SANTOS DEL REAL  
General Director for the Diffusion  
and Fostering of the Evaluation Culture  
[asantos@inee.edu.mx](mailto:asantos@inee.edu.mx)

Unit for the Divuligation and Fostering  
of the Evaluation Culture

Educational evaluation is a way of improving the education system, and its ultimate purpose is to help comply with the legal obligation to provide everybody with high-quality, equitable education.

In this regard it is important that the evaluations comply with the highest technical standards so as to ensure that their results

are valid and trustworthy, and also that the said results be widely disseminated, so that, in the first place, those involved in education, and in the second place, the general public, are acquainted with them, assume them as part of their way of looking at the world, and are able to use them appropriately and constructively in their respective fields of activity. Looking at the whole picture, fostering an evaluation culture, disseminating the results, and encouraging people to use them, should be just as important as producing the information itself.

As a coordinating entity of the National Educational Evaluation System (Spanish acronym: SNEE) and an authority on the subject, the National Institute for Educational Evaluation (Spanish acronym: INEE) is responsible for making people aware of what evaluation is and what it is for. It is called



upon to help disseminate information that will encourage those involved in teaching to commit to the effort to provide everybody with high-quality education as mandated by law, and hence willingly agree to having their performance, and, indeed, all the components of the education system, evaluated, to learning about evaluation and being willing to use the information that it yields to achieve significant improvements in education, and to updating their own classroom-evaluation practices.

For evaluation to work and fulfill its ultimate purpose, we need to foster a new view of it, rid it of its punitive connotations, and provide solid evidence about its ongoing role, in compliance with the legal obligation to provide high-quality education to everyone, in creating a better education system, helping to professionally develop those working in the latter, and improving the teaching-learning process.

To do this, we need to lay the foundations for rethinking and reorganizing evaluation in a way that enables us to actively promote patterns of behavior that foster accountability, educational improvement, and, in general, effective monitoring of respect for the right to education. In a way, this new evaluation paradigm would appear to have been anticipated in several sets of norms developed by the different associations specializing in evaluation.<sup>1</sup> The three norms described below provide a different view of evaluation and help us to rethink the role of educators in the said process, how to disseminate the resultant information, and how to use it to influence education policy.

### Involving educators in evaluation

All the aforesaid norms acknowledge the importance of involving educators in the evaluation process, since this has many advantages that are decisive for the success of the endeavor, including the following:

**Policy:** making those indirectly or directly involved participants and sharers in the process helps to confer legitimacy on both the evaluating body and on evaluation per se, as well as on the results eventually obtained.

**Ethics:** The evaluation of a given educational manifestation not only provides a basis for subsequent interventions that modify that manifestation, but also is, per se, a kind of intervention. Though it is feasible to carry out an evaluation without involving the evaluatee(s), from an ethical standpoint, the latter should not be converted into mere subjects of the process. In order to prevent the evaluation from becoming an imposition, one must let the evaluatees know that they can make a relevant contribution both to the evaluation and also the process of change and innovation of which the evaluation is a part.

**Practices:** Functionally speaking, an evaluation process in which the evaluatees actively participate will naturally be more fluid, as well as being less subject to the hiding of information and administrative obstacles.

Taking the interests and points of view of those involved in evaluation in no way implies subordinating the evaluation to private conflicting interests that could bias it. The entity carrying out the evaluation must not lose sight of its basic commitment to the public good, above all if the said entity is a publicly funded government agency.

### Disseminating results

The standards also stress the dissemination of results is pivotal and should not be limited to mere diffusion, but also involve the offering of information, guidance and suggested actions, so that the potential users of the said results, in their respective spheres of action, can make constructive use of them and convert them into new ways of thinking about, and carrying out, education, to the benefit of society.

In this regard, the standards indicate that complete, relevant information should be directly offered to potential users in a timely manner, via appropriate channels and using understandable language, with a view to fostering: (1) proper understanding of the results and their scope and limitations, (2) linkage of the said results with practical aspects of the evaluatee's working or professional life, and (3) the use of the information to foster educational improvement.

In general, the extent to which evaluation leads to educational improvement depends on whether the necessary steps are taken to help the different people involved in the process to fully understand and appropriately use the results.

### The impact of evaluation on decision-making

The most important of the possible impacts, and the one that evaluation most explicitly seeks to have, is that of affecting decision-making at different levels – and not only at the macro level.

In this regard the standards encourage the evaluating body to communicate the results to decision makers so as to increase the possibility of their having an impact,



be it conceptual (i.e. changing the way that problems are seen) or practical (i.e. leading to change and innovation in institutions, processes, programs or teaching materials).

Beyond the recommendations contained in the standards, there would appear to be general consensus that evaluation should have some type of impact on decision making, though there is no consensus at all about the kind of impact that should be sought or how to achieve it. For example, some evaluators try to influence the way in which authorities and educators see the problems in the education sector, rather than endeavoring to have an impact on specific decisions. Others endeavor to offer specific advice to decision makers, believing that the latter should be a veritable roadmap for change.

The former reproach the latter for the instrumental focus of their recommendations, and, above all, for their ingenuous view of decision-making, which they believe to be a purely rational process, stressing that the formulation of government policy means reconciling conflicting interests, leading to solutions that are not the most sensible ones from a strictly technical point of view.

An alternative way of proceeding is to invert the terms of the equation, and, rather than trying to convince politicians and other people involved in education to use the information yielded by the evaluation, attempt to supply information that may be relevant for all the parties involved and in line with the type of decisions that they are able to take within the limits of their competencies and in accordance with their position in the education system.

One way in which evaluators can achieve this is by talking to high-level decision-makers and other people involved in the education system throughout all the stages of the evaluation, until policies and programs have been drawn up and put into practice.

This is a route that can effectively relink educational evaluation with decision-making in the context of the new policy-formulation paradigm, which is characterized by democratic participation. Here lies what may nowadays be the biggest challenge for educational evaluation.

Given the above, one of the strategic components of the new paradigm is the evaluation culture, understood as the conceptions, judgments, values and attitudes of a society, or certain segments thereof, regarding educational evaluation in general - the organizations that that design and

implement the latter, the processes that govern the education system and its different components, the results produced, and the ways that different people use it, the impacts, real or imaginary, that they believe it to have, etc.

The evaluation culture is composed of perceptions, beliefs and convictions that are shared by society and articulated via a series of mechanisms that serve as grading codes whereby society as a whole, or certain groups or areas of it, adopt a given attitude towards educational evaluation.

While educational evaluation is a single entity at the conceptual level, at the level of social reality educators and other members of society develop different evaluation cultures, depending on their positions, interests and expectations. Each of these cultures is built around more or less stable around more or less stable social representations that are crystallized into patterns of interpretation, and hence not easily changed – at least not in the short term. These educational evaluation cultures are sets of assumptions that *condition the educational-evaluation practices and behavior of different protagonists and social groups*.

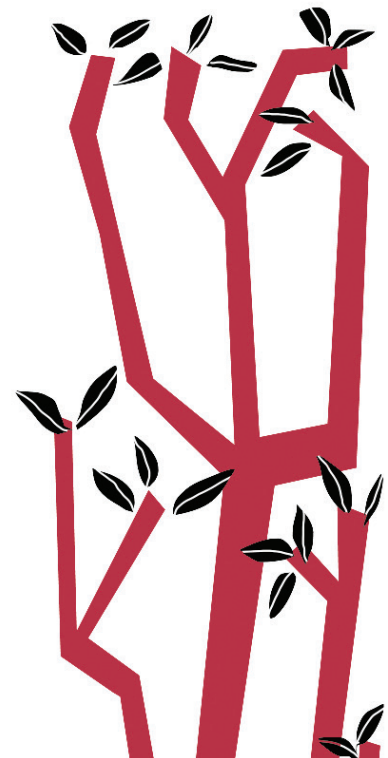
Though, in its most general sense, *fostering an evaluation culture* refers to the creation of a widespread *social culture*, clearly different groups are bound to have different demands. For society as a whole, it suffices to disseminate and affirm a general awareness of the existing evaluating bodies and their motives, and also of the general nature of evaluation and its usefulness for improving the education sector. However, in the case of people working in the education system, politicians and other groups directly involved in education sector, one obviously has to do more than this, and hence it is necessary to establish priorities. While one has to help explain what educational evaluation, who carries it out and what it is for, to society in general, one needs to focus one's efforts on authorities, people and entities involved in education, and other persons and groups involved in evaluation, such as parents, civic organizations, researchers specializing in the area of evaluation, intellectuals and journalist.

Hence, our institute's mission with regard to the educational-evaluation culture, the dissemination of evaluation results, and the fostering of constructive use of the latter, is to create conditions whereby we can have an impact on decision-making, offering useful,

pertinent information to authorities and to all those involved in education, and furnishing teachers, school principals and supervisors with diagnoses that will help them to identify problems in the areas of teaching-learning and management, and find innovative solutions to them; also our task is to inform society in general about general developments in the SEN and the extent to which the obligation to provide everybody with high-quality, equitable education is being fulfilled, so that our citizens may be vigilant and hold the authorities responsible for this task accountable. €

## References

- American Educational Research Association (2011). *Ethical Standards*, AERA. Available at <http://goo.gl/EIGffX>. Consulted in May, 2014.
- American Evaluation Association (2004). *Guiding Principles for Evaluators*, AEA. Available at [www.eval.org/p/cm/ld/fid=51](http://www.eval.org/p/cm/ld/fid=51). Consulted in May, 2014.
- DeGEval-Gesellschaft für Evaluation (2001). *Evaluation Standards*. DeGEval. Available at <http://goo.gl/1CBVAi>. Consulted in May, 2014.
- Sanders, J. (1998). *Estándares para la evaluación de programas*. Comité Conjunto de Estándares para la Evaluación Educativa. Bilbao: Ediciones El Mensajero.
- Scottish Educational Research Association (2005). *Ethical Guidelines for Educational Research*, SERA. Available at <http://goo.gl/tG8EXV>. Consulted in May, 2014.



## ROADMAP

## The evaluation of educational policy

One of the tasks that the INEE is now empowered to carry out is that of evaluating educational policy. Regarding how the said evaluation is being designed, Mendieta Melgar, the author of this article comments: “The challenge is to produce information that is of relevance for national policy, helps to bring about positive changes in the fundamental political spheres, and fosters compliance with the right to education of all Mexico’s children and young people”.

GIULIANNA MENDIETA MELGAR

Director for the Evaluation of Educational Policies and Programs General Direction of Guidelines for the Education Improvement of the INEE’s  
mmendieta@inee.edu.mx

The system for the evaluation of educational policies and programs by the National Educational Evaluation Institute (Spanish acronym: INEE) is being designed to provide a comprehensive approach to the challenges faced by the National Education System and has the clear aim of creating input, information and evidence to support and guide educational improvement. How, then, does it differ from the other evaluation systems developed in our country to date?

The evaluation of federal education programs, and, to a lesser extent, that of educational policy, has been the responsibility of the National Policy Evaluation Council (Spanish acronym: CONEVAL), an autonomous entity charged with measuring policy levels and evaluating programs, aims, goals and actions pertaining to social-development policy, which includes the overcoming of poverty via education. Just over ten

years after CONEVAL was set up, and ten years after the implementation of the General Guidelines for the Evaluation of Federal Public-administration Programs,<sup>1</sup> and one can affirm that one of the main virtues of its model is the latter’s institutionalization of the development of periodic evaluations of the performance of social programs.

The said evaluation’s distinctive feature is that it explicitly focuses on the various phases (diagnosis, design, processes, results and impact) of the program’s cycles. However, with regard to the usefulness of the results of these evaluations for guiding policy decisions in order to foster educational improvement, it is clear that there are still windows of opportunity due to two factors:

- 1) The evaluation of public policy has been little exploited for use in decision-making.
- 2) Evaluation results are still not used enough as a basis for the making of policy decisions, program improvement and budgeting.

Hence, the challenge facing the INEE and other specialized evaluation agencies consists in producing information that is relevant for the formulation of national educational policy, helping to bring about basic changes in the said policy, and helping to ensure that the right of all children and young people to education is respected.

### The proposed model

Since the main aim of the evaluations developed by the Institute is to improve education, the evaluation-model developed by its Direction for the Evaluation of Educational Policies and Programs sets out to make the transition to a model for evaluating *strategic educational policies* that tackle problems that are major public and social priorities having to do with equity, rather than evaluating individual programs which, though useful, are disjointed and fragmented.

The evaluation that the INEE is developing stems from the strong evidence generated by the aforesaid evaluation model, and hence it is a matter of assessing the coherence, coordination, articulation, pertinence and effectiveness of an educational policy and ascertaining the extent to which it helps produce the expected results and create guidelines.

With regard to the brief description of the model that is given in this article, it should be stressed that, by *educational policy*, we mean the coherent, articulated structured, stable, deliberate, reasoned, systematic set of *decisions* or *failures to act*, via which the State tackles concrete educational problems. Any educational problem can become a *public problem* to the extent that it becomes part of an agenda – i.e. when it is taken up, made evident and formalized by those in government or in the social sector, who propose actions to solve it.

Educational policy unfolds in a given institutional and organizational context, with laws, norms, administrative processes and both formal and informal rules of the game, as well as both explicit and implicit resources, assigned for its development, and areas and paths that should be interconnected. In this context, a set of protagonists participate, with different interests and motivations, whose postures derive from a set of technical positions, policies and political ethics (Aguilar and Bustelo, 2010; Cardozo, 2012).

In this regard, the educational-policy that the INEE is proposing sets out to assess the successes and failures of government action in order to move towards respecting, safeguarding and promoting the right to a high-quality education for all. Hence, it proposes to evaluate the coherence, coordination, articulation, pertinence and effectiveness of those components that constitute a policy –i.e. the programs, services, actions, resources, capacities, restrictions and motivations of the different protagonists- in order to solve the educational problem.

The aim is both to better understand the public problem that gave rise to a given policy proposal, assess the change theory underpinning it, and also ascertain how complex the implementation of that policy will be, via an approach that will make it possible to unravel the set of restrictions that impedes the said policy from being implemented, and guide the results of the said policy so that they may be used to achieve improvement by decision-makers and education-sector protagonists.

The main goal of this type of evaluation is that of providing evidence that will enable progress to be made towards building coherent *policy* and *inter-policy* systems



that generate synergies in order to achieve the strategic objectives that have been established in line with the right to a high-quality education for all, with special emphasis on underprivileged groups - e.g. handicapped children and young people, migrants, refugees and indigenous people).

### The INEE's evaluation of educational policy: Roadmap

Traditionally, and for purposes of analysis, the policy cycle has been seen as comprising seven processes that feed back into each other and overlap in time: (1) The inclusion of a problem on the public agenda; (2) The structuring of the said problem; (3) The set of possible solutions to the said problem; (4) The policy-analysis of the aforesaid potential solutions; (5) The reaching of decisions as

According to Cejudo and Michel (2014), "The specificity that is a distinctive feature of policy focus deems policy itself to be the sole focus of analysis and can lead to short-sightedness, since we concentrate entirely on each public policy without paying attention to the entire set of public policies that are, automatically harmoniously interrelated. But a set of well-designed, well-implemented public policies is not the same as an assortment of policies that all complement or reinforce each other in order to solve complex problems." (p.1). Thus, one can identify the following three levels: (1) *Policy* (i.e. internal) *coherence*, which is directly linked to the causative theory that structures a policy and ascertains whether there is logical linkage and a relationship of cause-and-effect among the definition of the problem, the policy instruments and the expected solution; (2) *Coherence among policies*, i.e. the presence of various public interventions that interact, reinforcing, rather than blocking, each other. When policies are not coherent, they may hamper each other, overlap, or fail to pay attention to certain aims or people; and (3) *Coherence among policy spheres* - i.e. the relationship among policies in one sphere and those in other spheres, and how the said relationships achieve major goals.

to the most technically sound, and politically supported, option; (6) The implementation of the public policy; and (7) The possible evaluation at each stage of the cycle.

According to Cronbach, evaluation doesn't just mean choosing a method to answer specific questions that are especially important to a given audience, but also implies focusing on understanding both social problems, and also the ways that programs or policies can be evaluated so as to solve them (Greene, 2013). Hence, the evaluations that the Institute is developing set out to become instruments that foster improvement of government action, in order to gradually promote and vouchsafe respect for the right to high-quality education for all.

The improvement of a given aspect of education is not achieved exclusively via the implementation of *formative* evaluations, but also by means of *summative* ones, which is why the evaluation model proposed by the INEE is generally oriented towards developing the former, based on a copious analysis of limits and opportunities in the context of government action. However, this position does not preclude the development of summative evaluations - e.g. of results and impact.

The INEE's educational-policy model seeks to foster a comprehensive understanding of the latter, with the evaluation path comprising the following three phases:

**Phase one. Identification of the series of educational rights and obligations that the State must safeguard/fulfill**, along with the latter's interdependence with other rights. It should be stressed that, though it is not intended to carry out an in-depth analysis regarding compliance with the other rights involved, it is intended to link the right to high-quality education to other related rights, rather than separate it from them, and also to link compliance with the said right to other strategic sectors (Serrano y Vázquez, 2013). Thus, this first phase includes a rough overview of what is desirable and of the minimum progress that it is hoped to achieve in the endeavor to comply with the previously mentioned right to high-quality education.

**Phase two. The fostering of a better understanding of the complexity of the public problem that the policy addresses** regarding effective compliance with the

right to high-quality education, and also the assessment of what the State is doing to safeguard the said right. The public policy in question is analyzed in terms of actual compliance with the right to high-quality education, both at the formal level -i.e. the level explicitly addressed in the State's regulatory and guidance documents- and also at the operational level, in terms of the policy actions that the State has embarked on (or failed to embark on) as part of its daily political agenda.

This phase implies: (1) *Building a change theory*<sup>2</sup> pertaining to the aforesaid actions; (2) *Mapping the key protagonists* or interest groups that participate in the public policy in question, the obligors and the obligations; and (3) *Analyzing the funding* required to foster compliance with the right. These steps will make it possible to compare the formal strategic educational-policy proposal with the public policies already in place, and also to compare the map of the State's decisions, and failures to act, with the map pertaining to gradual progress towards complying with the right to high-quality education posited in Phase one.

Additionally, Phase two includes *the carrying out of an analysis of the current state of education, taking into account the state of the different components of the National Education System* - i.e. in accordance with the General Education Law (Spanish acronym: LGE): I. Students and parents; II. Teachers; III. Education authorities; IV. The Professional Teaching Service; V. Study plans, curricula, methods and materials; VI. State schools and decentralized organizations; VII. Government-authorized private schools ones whose programs are officially validated; VIII. Educational evaluation; IX. The information and school-management system; and X. The educational infrastructure (Article 10, LGE).

**Phase Three. Assessment of the said policy actions based on a series of basic criteria**, which are, or should be, components of all public policies - i.e. pertinence, sufficiency, equity, efficiency, effectiveness, impact, efficacy, coherence, coordination and articulation. Based on the said assessment, it would be possible to make a value judgment regarding the distance between

(a) the frame of reference stemming from the map of the State's rights and obligations to safeguard the right to high-quality education, and (b) public actions currently underway - e.g. the formulation of policy recommendations and the corresponding guidelines.

### Some final considerations

The module for the evaluation of policies and programs that is being set up in the INEE is intended to evaluate compliance with the right to high-quality education in accordance with the lge based on the dynamic interaction between the rights of children and young people and the obligations undertaken by the State and other obligees to safeguard the said right.

When taking technical and methodological decisions, we will, in a complementary way, favor using a *participative methodology* during all the stages of the evaluation, endeavoring to assign a central role to children and young people as obligors, and also to take into account the obligees – i.e. the State, parents, international organizations, schools, the private sector and civil society.

The evaluation is not just a technical exercise, but also has ethical and political dimensions. Hence, the INEE intends to

asses strategic educational policies based on series of basic criteria, in the context of the endeavor to gradually comply with the right to high-quality education.

Therefore, the said evaluation explicitly sets out to foster the use of its results to improve education, via the drawing up and issuing of guidelines as tools for guiding -or reorienting- the strategic educational-policy decisions taken by the education authorities. €

- 1 Secretaría de Hacienda y Crédito Público, Secretaría de la Función Pública and CONEVAL (dof. 30<sup>th</sup> of March, 2007): <http://goo.gl/FxoOQw>.
- 2 According to Carol Weiss (1997), policy or program theory has two components: (1) Implementation theory – i.e. the definition or description of the activities and of some of the intermediate intervention results; and (2) programmatic theory, which specifies the mechanisms used to bring about the desired changes. Weiss asserts that “change mechanisms are not program activities, but, rather, the effects produced by the activities” (p. 46).

### References

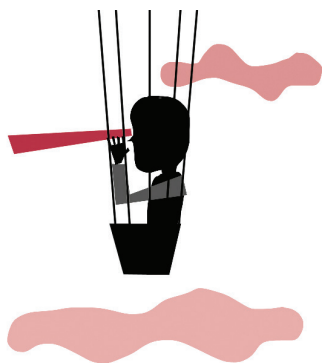
- Aguilar, L. and Bustelo, M. (2010). Gobernanza y evaluación: una relación potencialmente fructífera. *Revista Gestión y Análisis de las Políticas públicas*, 4.
- Cardozo, M. (2012). *Evaluación y metaevaluación en las políticas y programas públicos*. Estado

*del arte*. Universidad Autónoma Metropolitana. Mexico. Unidad Xochimilco. División de Ciencias Sociales y Humanidades.

- Cejudo G. and Michel, C. (2014). Coherencia y políticas públicas. Metas, instrumentos y poblaciones objetivo. *CIDE working document*, 284.
- CONEVAL (2013). *Construyendo un sistema de evaluación del desempeño para el desarrollo social*. Mexico: CONEVAL.
- Diario Oficial de la Federación (2007, 30<sup>th</sup> of March). *Lineamientos Generales para la Evaluación de los Programas Federales de la Administración Pública Federal*.
- DOF (2013, 11th of September). *Ley del Instituto Nacional para la Evaluación de la Educación*.
- DOF (1993, 13th of July. Última Reforma: 2015, 20<sup>th</sup> of April). *Ley General de Educación*.
- Green, J. (2013). An interpretation of Lee J. Cronbach's Vision of Evaluation. in M. Alkin, *Evaluation Roots. A wider perspective of theorists' view and influences* (p.99). California, USA: SAGE.
- Serrano, S. and Vázquez, D. (2013). *Los derechos humanos en acción: operacionalización de los estándares internacionales de los derechos humanos*. Flacso-Mexico.
- Weiss, C. (1997). Theory-Based evaluation: past, present and future. *New directions for evaluation*, (76), 41-55.

## SPEAKING ALOUD

## The challenge of providing high-quality education in Latin America and the implications for the Mexican Evaluation System

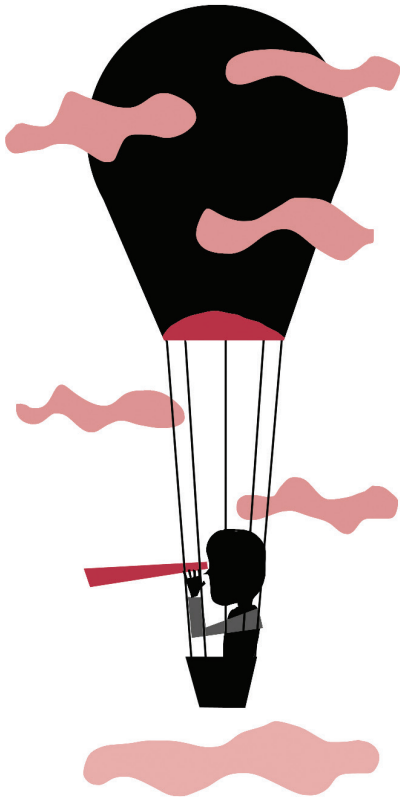


There is international consensus that raising the quality of education and improving learning outcomes will be crucial to worldwide development after 2015. Our country is facing educational challenges in the context of a new educational reform. What role will evaluation play as the year 2030 grows closer? “Mexico has a great opportunity to build an inclusive evaluation system,” says the author of this article.

NURIA SANZ GALLEGU

Director, and representative in Mexico, of the United Nations Organization for Education, Science and Culture (UNESCO)  
[n.sanz@unesco.org](mailto:n.sanz@unesco.org)

Almost 25 years ago, the Education for All (EFA) movement was born in Jomtien, Thailand. On that occasion, delegates from all over the world signed the *World Declaration on Education for All*, a historic undertaking to “satisfy the basic learning needs of all” by providing universal primary education and dramatically reducing illiteracy levels. At the World Education Forum, held in Dakar in the year 2000, six EFA goals were officially set, covering all the different aspects of elementary education, from early-childhood learning to adult literacy and educational quality, with the year 2015 being established as the deadline for achieving the said goals. In the worldwide discussions that were held in order to set goals for the year 2030, entitlement to education has been reemphasized and acknowledged as an ‘enabling’ right whose safeguarding which underpins all other human rights.



Latin America committed itself, from the outset, to pursuing the worldwide education goals, and hence, since the 1980's, its nations have, with varying aims and different degrees of impact, implemented a series of educational reforms. The first of these, which took place in the 1980's, set out to extend educational coverage. Subsequent reforms, in the 1990's, focused on raising quality and making education more equitable. Currently, new reforms are being implemented, this time placing even more stress on school autonomy and so called 'educational decentralization' – i.e. the shifting of decision-making power from the government to schools, with emphasis being placed on the teaching-learning process.

One of the main components of the Educational Reform that was carried out in Mexico in the year 2012 is quality, which the said Reform vouchsafes to all children and young people in compulsory education. Likewise, the 2013-2018 Sectoral Education Program (Spanish acronym: PSE) assigns priority to high-quality education as a means of achieving political, social,

economic and cultural development in Mexico, and states that such education is a prerequisite for reducing social inequality in that country. The PSE sets forth aims, goals and strategies for complying with the mandates of the Reform, stressing high-quality education and continuance in the different educational levels.

Mexico's efforts are focused on meeting the challenges identified in the Education for All monitoring reports issued each year by UNESCO. The raising of quality and the improvement of education will play a decisive role in post-2015 world development. The 2014 EPT Monitoring report<sup>1</sup> states that there are 250 million children who are unable to read and write or have not developed basic arithmetical competencies, of whom 130 million are in school, asserting that, in order to remedy this situation, we need to reinforce national evaluation systems and use evaluation results to formulate educational policy. It also places special emphasis on the fact that the learning crisis affects the underprivileged more, and stresses that the worldwide disparities conceal great inequalities within countries.

The education ministers who attended the 'Education for All in Latin America and the Caribbean' Meeting of Ministers of Education that was held in Peru in 2014 agreed that the said region has special features that will be the basis for the post-2105 agenda, which is as follows:

- Education is a basic human right and underpins other rights.
- The region's biggest challenge is inequality and its negative effects on countries' progress. The high levels of inequality in the region's countries are due, above all, to social-class differences, ethnic origins and geographical location.
- High-quality education is essential for reducing inequality and poverty and creating more inclusive societies.
- The concept of sustainable development should be at the core of all efforts to foster social development that is both inclusive and equitable.
- Education should encourage people to be active, effective citizens at both the national and global levels, and help to strengthen democracy, so as to promote dialogue, settle conflicts and reduce all forms of violence.

- Education systems should offer young people who are unemployed and/or not in school alternative social and job opportunities.

The main actions proposed in the post-2105 agenda focus on reducing inequality by providing high-quality, lifelong education. At the aforesaid meeting of ministers, the Latin American countries also acknowledged the rich cultural diversity in the region, stressing that interculturality is a prerequisite for high-quality education, which is defined in terms of *equity, relevance, pertinence, efficiency and effectiveness*. The said agenda stresses the importance of learning processes and outcomes, as well as that of producing data and research. The countries undertook to develop evaluation systems that take quality into account and yield data that support policy making and education-system management.

The evaluation of quality is a complicated activity, since quality is subjective and the word has different meanings. While most of the reforms that are being implemented associate it with results and teaching, the most important thing is that students learn and develop life skills.

Mexico has taken part in regional discussions about the kind of educational evaluations that are needed to measure educational quality, with one of the topics at the said discussions being the need to build a learning-centered evaluation culture to support all the changes in the various levels and areas of the education system.<sup>2</sup>

The countries in the region have good educational-statistics systems, which provide enough data to provide an overview of the processes and results of their education systems, as well as access to them, and the said statistics constitute a good basis for creating a mechanism aimed at holding the various countries' systems accountable for quality shortfalls.

However, despite all this progress in evaluation and information systems, we still have no data about the effects/use of the information obtained on/in decision making or on/in teaching processes in schools. The 2015 EFA report stresses the need to spur international reflection about bringing the different learning-evaluation systems into line with each other, and the countries have asked UNESCO's Statistics Institute in Montreal to hold consultations aimed at making



it possible to obtain results that can be compared with each other, since this is especially necessary now that the post-2105 Development Agenda requires a system for regional implementation and follow-up.

At the same time as designing their own educational norms and standardized national tests to measure student-learning outcomes, the countries are also taking part in international studies that measure the acquisition of skills and competencies, which will make it possible to compare the progress made in countries with different socio-cultural characteristics.

For example, the studies promoted by the Latin American Laboratory for Assessment of the Quality of Education (Spanish acronym: LLECE) have made it possible to measure improvements in, and imbalances among, the education results obtained in the different countries. The LLECE is a forum for debate about, and reflection on, the development of quality-evaluation systems that are suitable for the region. One of these projects, carried out in 2013, was the Third Regional Comparative and Explanatory Study (Spanish acronym: TERCE), whose relevance resides in its focus on student learning outcomes and its emphasis on building high-quality education systems, in all their dimensions. It is important to evaluate reading, writing, mathematical and scientific competencies in order to find out which factors are responsible, respectively, for the countries' successes and failures in their efforts to provide universal high-quality education.

By comparing the results of these tests, we are able to see the impressive progress achieved in some countries, and, for example, note that the gap between reading-skill levels in urban and rural areas in most of the Latin American countries has been getting smaller, which attests to increases in both educational quality and educational equity.

Hence, the information yielded by these international tests has helped us to get an overview of individual countries, or compare different countries within a group, though it should be stressed that not enough information has been produced to enable us to come up with solutions, or set up projects at the intermediate or local levels, as required by the states, municipalities and schools. To do that, we need to have information about the factors associated with successful learning, along with an analytical model in line with

local realities that enables us to establish priorities, and come up with options, for bringing about changes in school management and teaching.<sup>3</sup>

Mexico has a great opportunity to create an inclusive evaluation system that reconciles the heterogeneous characteristics of its different regions. Such a system could generate new, harmonious types of comprehensive evaluation that conjoin interculturality and an education that is suited both to each age group and also to each of the many different ethnic groups and cultures that make the country so diverse and rich. The said system should be participative and embrace all those involved in teaching children, young people and adults. The results it produces should provide meaningful information that enables education authorities to identify the factors that hamper student learning and institutional development, and, furthermore, they should serve as a basis for formulating educational policies aimed at achieving both national aims and the global goals that were committed at the World Education Forum that was held in South Korea in May of 2015 to discuss how to achieve high-quality, inclusive, equitable, lifelong education for all by the year 2030.

The new Education Agenda is inspired by a humanistic view of education and development based on human rights, dignity, social justice, protection of the vulnerable, cultural diversity, shared responsibility, and accountability. The seven global goals established during the Korea forum are:

1. By the year 2030, at least x% of all children should be ready to attend primary school.
2. By 2030, all children should complete an elementary-education program that lasts at least 9 years.
3. By 2030, all children, and at least x% of all adults should possess enough reading, writing and arithmetic skills to be able to fully participate in society.
4. By 2030, at least x% of all young people and y% of all adults should possess enough knowledge and competencies to find a decent job and live a decent life as a result of technical and professional training provided during upper-secondary and higher education.
5. By 2030, all students should acquire the knowledge, competencies, values and attitudes -via education for world citizenship and sustainable development

that are needed to build a peaceful, sustainable society.

6. By 2030, all governments should ensure that all the students in their schools are taught by qualified, professionally trained teachers who are motivated and properly supported.
7. By 2030, all countries should devote between 4% and 6% of their gross domestic product, or between 15% and 20% of their public spending to education, favoring the most needy.

Mexico is capable of building an evaluation system that enables it to monitor its progress towards achieving these goals for the year 2030. €

1 UNESCO (2014). EFA Global Monitoring Report: Teaching and learning: achieving quality for all.

2 UNESCO-LLECE (2008). Reflections on the evaluation of educational quality in Latin America and the Caribbean.

3 *Idem.*

## References

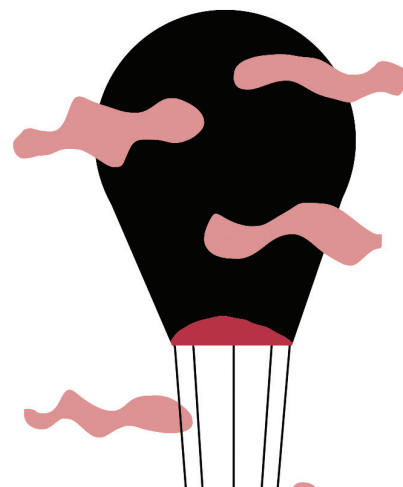
Diario Oficial de la Federación (2013). Decreto de Reforma de los artículos 3° y 37° de la Constitución Política de los Estados Unidos Mexicanos. Mexico.

Secretaría de Educación Pública (2013). Programa Sectorial de Educación 2013-2018. Mexico.

UNESCO (2014). EFA Global Monitoring Report: Teaching and learning: achieving quality for all.

UNESCO (2015). EFA Global Monitoring Report: Education for All 2000-2015. Achievements and challenges.

UNESCO-LLECE (2008). Reflections on the evaluation of educational quality in Latin America and the Caribbean.



## Veracruz: Evaluation faced with the challenges of the Educational Reform

“The big challenge is to ensure that the results of the evaluations for entry to the teaching profession, and for promotion, acknowledgment and continuance within it, are taken into account when designing strategies for improving initial and ongoing teacher training,” asserts Dr. José Ojeda, who looks at developments in the State of Veracruz in the context of the new national evaluation system.

JOSÉ A. OJEDA RODRÍGUEZ  
Director of the Educational Planning,  
Evaluation and Control Unit  
of the State of Veracruz  
[upeace2015@gmail.com](mailto:upeace2015@gmail.com)

Accelerated population growth and economic, social and cultural change having called traditional educational solutions into question, countries are pinning their hopes for a better future on making the inalienable universal right to education a reality, an aim that is only achievable if we focus on the main protagonist in the process – i.e. the teacher.

To bring about these changes, educational policies in Latin America, which, over recent decades, have focused on promoting change aimed at raising the quality of education, and, in Mexico, pertinent input materials are being created to support the current Educational Reform, a process of radical renewal that is transforming our education system from the ground up by making “quality” the touchstone for “ensuring maximum learner achievement” stemming from a synthesis of teaching methods and materials, school organization, adequate physical conditions and suitable teachers and management staff, in accordance with the stipulations of Articles 3 and 4 of the Mexican Constitution. To com-



ply with the aforesaid mandate, we require clear evaluation mechanisms in the form of tools for improving the education system and its processes and results. “Reforming” means changing structures, taking decisions and implementing policies to satisfy the needs that are detected.

### Progress and challenges

The 2011-2016 Veracruz Development Plan and the 2011-2016 Veracruz Education Program bear witness to the current administration’s commitment to face the challenges posed by the Education Reform and commit to a forward-looking education that supports fair, equitable development, at both the individual and collective levels, as a motor for promoting development in every corner of our state, a highly complex territory covering 71,826 km<sup>2</sup> with 7,643,194 inhabitants, 212 municipalities and 27,125 towns (National Institute for Statistics and Geography [Spanish acronym: INEGI], 2015), 92.3% of which have 500 or less inhabitants.

The Veracruz education system caters to 2,389,237 students<sup>1</sup> -i.e. 7.1% of Mexico’s total student population- and 78.6% coverage, as against a nation-wide coverage of 72%.

Since 2006, our state has been steadily increasing its coverage of the National Evaluation of Academic Achievement in Educational Institutions (Spanish acronym: ENLACE), which, initially being applied to 92% of all students enrolled in elementary education, covered 95.1% of the said population -i.e. 0.7%

above the national average- in 2013, also being initially applied to 93.58% of all students enrolled in lower-secondary education in 2008 and covering 95.70% -i.e. 5.1% above the national average – of the said population in 2014.

Results-wise, the number of primary-school pupils getting “good” and “excellent” scores in Spanish has increased from 17.9% to 37.9%, and in mathematics there has been an increase from 17.9% to 37.9%, while the number of secondary-school pupils getting “good” and “excellent” scores in Spanish has increased from 15.2% to 21.8%, and the number of pupils at that level getting “good” or “excellent” scores in mathematics has increased from 3.6% to 25%.

Furthermore, the results obtained by Veracruz in the Program for International Student Assessment (PISA) mirror those of the ENLACE, with an increase of 23 points in the sciences, one of 45 points in reading, and one of 45 points in mathematics, between 2013 and 2012.

Around 7,800 of the 109,053 elementary-school teachers in the state have taken part in the evaluation for choosing teachers for new-entry positions, a process that began in Veracruz in 2006 and has now been rendered compulsory by the Educational Reform.

### New challenges: using the results

According to the evaluation calendar of the Professional Teaching Service (Spanish acronym: SPD) and the National Institute for

## IN THE CLASSROOM

## Five examples of evaluation in Mexico: local autonomy, experiences and context

In the Dialogues for the Creation of a National Educational Evaluation Policy, various key players in the National Education System have talked about the challenges and options offered, by the Educational Reform. In the interviews transcribed below, the heads of the institutions responsible for evaluation in the states of Sonora, Mexico, Quintana Roo, Baja California and Yucatán describe how they tackle their jobs in their different contexts.

### Interpretations of comprehensive evaluation

In order for an evaluation to be deemed to be comprehensive, one has to take stock of all its components, ranging from the context in which it is carried out, to those who take part in it, to those who stand to benefit from it (e.g. parents, students, teachers, school principals and supervisors). Furthermore, its results, both positive and negative, should be set forth in a balanced manner that enables us to make clear, responsible decisions aimed at achieving improvement. *Guadalupe Leticia Quetzal Hoil, Director of the Center for Educational Improvement of the State of Yucatán (Spanish acronym: CEEEY).*

Educational evaluation should be conceived of as a tool for improving educational practices, processes, programs, involve all the components of the educational process, and explain its results while acknowledging the important role that context plays in them. *Joaquín Caso Niebla, Director of the Institute for Educational Research and Development (Spanish acronym: IIIDE) of the Autonomous University of Baja California (Spanish acronym: UABC),*

Educational Evaluation (Spanish acronym: INEE), 2015 will mark the beginning of a new phase involving the evaluation of all the components of the education system so as to provide benchmarks for the making of decisions – above all about teacher updating.

We in Veracruz are convinced that comprehensive, contextualized, longitudinal analysis of the results of these evaluations will provide us with important tools for the phase in question, making it necessary to take stock of the relationship among the three types of results -i.e. student results (learner evaluations), teacher results (SPD evaluations) and evaluations of school conditions (contextual evaluations), taking into account the socio-cultural environment and also to create mechanisms for facilitating ongoing communication among the protagonists in the education system.

To achieve the above, schools need to commit to an evaluation and self-management culture. The teachers and managers in the elementary and lower-secondary schools in Veracruz have access to the training possibilities –i.e. the online and classroom-taught courses, workshops and online courses- that are offered by the Ministry of Education, in response to the needs detected during the evaluations or expressed by the interested parties, for the purpose of developing professional competencies aimed at improving their performance and getting better student results.

Alongside this, over the last few years strategies for improving institutions, based on information and communications technology, have simplified and speeded up administrative work so as to enable teaching and administrative staff to devote more time to high-value activities.

For their part, education-sector computer systems are being restructured in order to satisfy the requirements set forth in the federal regulations stemming from the Reform. This is the case with academic-control information and the updating and validation of staff, the use of the information garnered via the Census of Elementary-level and Special Schools (Spanish acronym: CEMABE), the work-center catalogues, the systematic recording of teachers' training paths, and, recently, the State Evaluation System, consisting of databases containing the results of the standardized and internal evaluations of institutions' with regard to school organization, performance, drop-out,

lag-behind, failure and terminal-efficiency indices, location, infrastructure and equipment, and initial teaching- and management-staff training aimed at helping those involved in education to plan and come up with operational strategies.

This collection of well-ordered, clear data, which forms part of the National Educational Information and Management System, will provide us with yardsticks for understanding the results obtained in Veracruz and rectifying flaws both in the system and in our education policies, enabling us to determine which actions to take in the furtherance of equity and educational quality.

One of the big challenges consists in ensuring that the results of the evaluations that will be carried out regarding entry to, and continuation in, the teaching service, and promotion and acknowledgment, are analyzed and taken into account when designing specific, detailed strategies for improving the initial and ongoing training of teachers and educational administrators, and also for designing performance- and results-based incentives. Given the importance of achieving the previously mentioned aims, both teachers and society in general should have access to open forums where they can express their opinions and exchange ideas in a way that lends certainty and credibility to the processes in question.

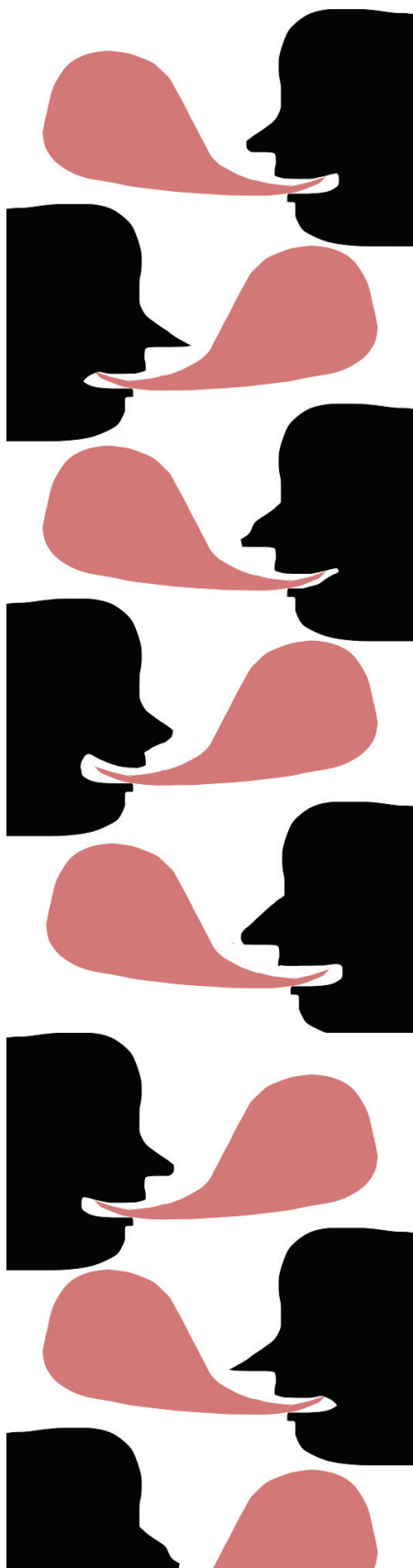
If the Reform, as has been asserted time and again, deems teachers to be crucial to achieving the high quality with equity that we long for in our education system, then our other challenge is to ensure that the said teachers have opportunities for training, professional development and acknowledgment that will enable them to achieve a decent standard of living. €

1 I am referring to people between the ages of 4 and 24.

### References

Diario Oficial de la Federación (2013). Articles 3 and 73 of the Mexican Constitution. Mexico. National Institute for Statistics and Geography (March 2015). Sole Catalogue. Mexico, INEGI.





*an institution that works hand-in-hand with the state's Ministry of Education and Social Welfare.*

Comprehensive evaluation must involve all those who play a part in the education system – i.e. teachers, who are the main protagonists, and parents, whose participation is growing day by day. Of course, evaluation results must be transparent, socially relevant, made public, and used to guide educational policy. *Edgar Hiram Sallard, General Director of the Institute for Educational Innovation and Evaluation of the State of Sonora (Spanish acronym: IIEEES).*

### **Evaluation and context: experiences at the state level**

#### **Teacher evaluation**

A public-consultation project, involving teachers and other people working in elementary education, was carried out in the state of Yucatán in January, 2013, for the purposes of updating the profiles for professionals working in elementary schools. Groups of academics chose and validated the contents of the said on-line consultation, which was drawn up in order to achieve greater coverage and facilitate participant response. In a subsequent second stage with focus groups affording direct access to teaching professionals, representative sections of the law, covering things such as entry to the profession, performance, promotion and acknowledgment, were consulted, and all this information enabled us to obtain useful results. *Guadalupe Leticia Quetzal Hoil (CEEY).*

Since 2008 we have had in place a program entitled “Ser Maestro” (Being a Teacher) stemming from an agreement signed with the Union of Teachers at the Service of the Mexican State (Spanish acronym: SMSEM). The teachers belonging to the state subsystem (around 60% of all the teachers in our state) join this program voluntarily and take part in an evaluation-and-acknowledgment program leading to a cash stimulus. The program is currently under review and we are engaged in discussions with the National Institute for Educational Evaluation (Spanish acronym: INEE) and the National Department for the Coordination of the Professional Teaching Service aimed at deciding which route to take. *Héctor Morales Corrales, General Director of the Institute for Educational Evaluation of the State of Mexico (Spanish acronym: IEEM).*

Also we carry out field studies and analyze the results published by the INEE and the Ministry of Public Education (Spanish acronym: SEP), in addition to which the Institute works with teachers in a voluntary basis to improve training. Over the last six months, we've with two thousand lower-secondary-level teachers. While the latter understand what “competencies” are, the problem is that it's not clear how they work to develop competencies in line with the 2011 reform. *Roberto Wilbert Castillo Tamayo, General Director of the Institute for Educational Research and Development of the State of Quintana Roo (Spanish acronym: IINDEQ).*

#### **Educational achievement**

Last year, for the first time in nine years without interruption, we stopped carrying out the State Evaluation of Student Performance in primary, lower-secondary and upper-secondary institutions due to the Educational Reform, which stipulates that the only institution empowered to design [evaluation] instruments. We still haven't had any technical guidance on authorizing any evaluations that occur here. *Edgar Hiram Sallard (IIEEES).*

#### **Test design**

We've designed exams that meet the requirements of the State Education System aimed at analyzing our students' contextual variables. Before the National Evaluation of Academic Achievement in Schools (Spanish acronym: ENLACE) program was put in place, we gave questionnaires to students, teachers and administrative staff in order to put our findings and results into context. This is an example of the use of input from the National Education System in a state-wide evaluation, and the results of our research were passed on not only to the different people involved at the state level, but also to civil society in Baja California, and, above all, to parents. *Joaquín Caso Niebla (IIDE).*

#### **What still needs to be done**

We need to train professional evaluators. Here we've learned by trial and error, and we need funding in order to train people and help them to grow and become specialists in educational evaluation. *Héctor Morales Corrales (IEEM).*

We don't talk very much about funding in order to train our human resources. This aspect shouldn't be forgotten, because, though there's a desire to improve and some

strengthening in the area of evaluation, the fact is that, since there's no funding to enable us to solve our problems and get what we need, we can't do very much. We lack knowledge about the implications of the evaluation process, ranging from the design of the instrument to its use. *Guadalupe Leticia Quetzal Hoil (CEEY)*.

Financing is our main, ongoing, challenge, while the second one is the lack of clearer guidelines and state-linked regulations enabling the INEE to provide the legally mandated technical guidance. We need to have a shared vision on the part of those in charge at the national level and the state education authorities, to understand the purposes of evaluation and be aware of their scope, and the INEE and the Ministry of Education of Mexico (Spanish acronym: SEP) itself need

to work in harmony to ensure that the state-level evaluations are really comprehensive and enable us to carry out a fairer diagnosis. *Edgar Hiram Sallard (IEEES)*.

One model that can be explored in other states is the one that links the INEE, the State Education System and the university, since it's the latter that possesses the technical know-how needed to carry out the tasks that the Institute has assigned to the states. We'll get better results all round if we reconcile the state models, the local contributions and the states' expectations with federal requirements. *Joaquín Caso Niebla (IIDE)*.

Our main challenges in Baja California concern diversity. It's a very complicated state, with 80% of its population concentrated along the border in addition to which indigenous people from Oaxaca, Chiapas and

Guerrero migrate there, settling temporarily. The teachers in our state sometimes have to handle groups in which as many as five native languages are spoken. *Juan Gálvez Lugo, Director of Educational Evaluation (DEE) of the Ministry of Education and Welfare of the state of Baja California*.

Given the current laws, and since we need to adapt the materials that we have, we need to choose teachers who can provide support. One way to ensure the success of the [educational] reform as a whole is to adapt the syllabi and the courses currently available, but doing both of these things thoroughly. *Roberto Wilbert Castillo Tamayo (INDEQ)*. €

Interviews by Laura Irene González Mendoza and Magdalena Alpizar

### Autonomy, structure and *raison d'être*

Entity	Legal status	Structure
<b>IEEEM</b> <b>State of Mexico</b> Set up in 2007	Decentralized organization, established in accordance with the Decree Governing Decentralization. Its budget for 2015 is Mxp 23 million.	Three departments: Management, Data Collection and Model Analysis; each with two sub-departments. Team of 60 professionals.
<b>CEEY</b> <b>Yucatán</b> Set up in 2006	Decentralized organization answerable to the state's Ministry of Education and without any funding of its own.	Four departments -Evaluation-process-and-Instrument Design; Data Evaluation, Analysis and Dissemination; Operations; and Study Syllabi and Programs; plus a management area. Team of 33 professionals.
<b>IEEES</b> <b>Sonora</b> Set up in 2004	Decentralized, technically autonomous government organization with federal and state funding, coming under the control of the Ministry of Education and Culture of the State of Sonora.	Five departments: General; Innovation; Instrument Design; Evaluation and Logistics; and Planning and Management. Team of 12 professionals.
<b>DEE</b> <b>Baja California</b> Governed by the 2008 regulations	Belongs to the Undersecretariat of Lower Secondary Education, Upper Secondary Education, Higher Education, Teacher Training and Evaluation of the Ministry of Education and Social Welfare (Spanish acronym: SEBS), with federal and state funding.	Two sub-departments: an Operations department in charge of evaluations and a Technical department in charge of analysis and outreach workbooks.
<b>IIDE</b> <b>Baja California</b> Set up in 1990	Academic unit of the Autonomous University of Baja California (Spanish acronym: UABC), which provides research support.	Organization for the creation, application, divulgation and transfer of scientific and technical knowledge about education.
<b>INDEQ</b> <b>Quintana Roo</b> Set up in 2011	Private institution devoted to the evaluation of the education and business sectors.	Two areas: an academic one that processes diagnoses and turns them into input for diploma programs, and general and specialized courses, and an administrative one. It has a team of specialists in curriculum design and another area in charge of field studies. Team of 18 professionals.

## Monitoring compliance with/exercise of the right to education in Brazil

The authors offer an overview of current developments in the monitoring of the legally mandated right to education in Brazil, and of the main characteristics and challenges in the biggest educational system in the Latin American region.

JOSÉ FRANCISCO SOARES

Chairman of the National Anísio Teixeira Institute of Educational Studies and Research (Portuguese acronym: INEP)

HELBER RICARDO VIEIRA

Adviser to the Chairman of the INEP Brazilian Ministry of Education



Article 205 of the Brazilian Constitution states that “Education, to which everybody has a right, and which is an obligation of the State and of the family, shall be promoted and fostered with the cooperation of society, its purpose being to fully develop the individual, to prepare him/her to exercise citizenship, and to provide him with the qualifications to get a job”.

The Constitution clearly stipulates that this right is only [fully] respected when every citizen acquires enough knowledge to achieve the three aims of education, in accordance with the right to learn. Other articles state that the said learning must be acquired via teaching, offered in schools forming part of education systems run by the state or municipality. The Law Governing the Rights and Obligations of Education acknowledges that there are other entities that offer education, but stresses the central role of the school, and hence of the education systems.

Besides being legally mandated, all social rights must be safeguarded by institutions set up for that purpose, and respect for the said rights must be monitored. In the case of the right to education, the following three aspects should be taken into account:

**First aspect:** Access to, continuance and promotion in, and conclusion of, the different education cycles. Quality in this aspect is defined in terms of regularity – i.e. the different school cycles should be completed at the expected age, so that an irregular trajectory that is longer than necessary is tantamount to not exercising the right to education. Not entering school is the most serious failure to exercise the said right.

**Second aspect.** Since the text of the Constitution is so vague that it does not make clear what the rights to education are, it is necessary for people to reach an agreement about the competencies needed to vouchsafe a decent life to every citizen. The discussion of this issue, which should begin with a definition of the basic competencies, has still not been held in Brazil. However, in view of the conclusions already reached, and what has occurred in other societies,

the acquisition of the said competencies should start with the development of cognitive competencies consisting of clearly defined knowledge and skills, non-cognitive competencies, the values required for individual development, the exercising of citizenship, and preparation for the workplace. Quality, in this aspect, consists in acquiring the aforesaid competencies to a suitable level.

Since it is impossible to vouchsafe to all our citizens the same amount of learning –given that there are free, personal choices that prevent it- we need to define the difference between the education paths and the amount of competencies acquired, which should be compatible with a law-abiding society.

Since the learnings needed in order to exercise the right to education are acquired in schools, it is there that compliance with the said right should be monitored, and hence the information yielded by that monitoring can, and should, be used in the evaluation of schools and education systems, though it is just one of the activities pertaining to the education process. The measurement of –and yardsticks applied to- these results, are simplifications of the complex realities of education, and, while they help to bring about improvement, more needs to be done.

### 1. Data for the monitoring of compliance with a focus in the right to education

**2.1 Learning path.** In Brazil, both the population census and the National Survey of Households (Portuguese acronym: PNAD), yield information about entry to education, which is the first stage of the learning path. The other aspects of the said path –i.e. continuance, promotion, and completion- are monitored by the two-stage Educational Census of Elementary Schools<sup>1</sup> that is carried out by the The National Anísio Teixeira Institute of Educational Studies and Research (Portuguese acronym: INEP), in collaboration with the Ministries of Education of Brazil’s 27 states. Various legal statutes govern both public and private elementary schools and stipulate the last Wednesday of May as



the date for the first (i.e. entry stage) and the gathering of data, since this is when the Brazilian education system is most stable, so that the data in question can serve as a yardstick for the rest of the year. In the second stage, which occurs a year after the first one, schools issue a final progress report on each student, indicating whether s/he has passed, failed or dropped out.

Being used as a basis for formulating and monitoring different government policies, the school census is an essential tool that enables educators and society in general to have access to data about learning-path performance and statistics about students in each state and municipality. The INEP uses a broad range of yardsticks to gather data about school and student performance, including pass rates, academic progress, and mismatches between age and level. The information about teachers makes it possible to compile statistics about the training, and continuance in employment, of each school's teaching staff.

**2.2 Learnings.** The term *competency* is often used to denote the three aims of education under our Constitution. Discussion about this topic starts with the definition of basic competencies. Some countries stipulate that their students should acquire cognitive and socio-emotional competencies, as well as values, during elementary schooling. The cognitive competencies include mother-tongue reading, writing and speaking skills, plus mathematical and scientific skills.

**2.2.1 Empirical learning evidence.** The learnings can only be monitored via tests. Since the said monitoring involves millions of students, we need to build a big logistical system whose processes are rigorously standardized, and which limits the number of learnings to be monitored. The monitoring activities have to do with school evaluation, which complicates the task, given the difficulties and risks that it entails.

The main yardstick is the *Prova Brasil* (Brazil Test), which records each student's attainment in reading and mathematics in a national database that was created in 1995. All students registered in the final years of government primary school and government secondary school respectively take the said test each two years. As well as providing basic information for monitoring compliance with/exercise of the right to education,

the results are used for several other purposes, including evaluating schools, measuring academic irregularities, assessing the effectiveness of government policies and study programs, and holding school systems accountable.

However, as often occurs in democracies, different groups use the data for different, very specific, purposes. For example, there is a lot of disagreement about the use of the *Prova Brasil* data to create school incentive systems. Many of those involved in education consider that the said data are not suitably used, and also affirm that eliminating the gathering of learning data would make it impossible to monitor compliance with/exercise of the right to education.

### 2.3 Contextualization and the linkage of learning yardsticks to the curriculum.

In order to properly use the data that are gathered, the schools that supplied the latter must have access to, and understand, them, for which purpose we need to place the said data in their social context when disseminating them and explicitly indicate their relevance for the curriculum.

The purpose of the said contextualization is to provide details both about the students whose learning outcomes were measured, and about their schools, which is necessary so that society can understand that the results reflect the living conditions in the communities in question, since it is unreasonable to expect students in very different geographical, social and emotional contexts to achieve the same learning outcomes. The results of schooling and learning that are disseminated should be accompanied by socio-demographic profiles of the students and the resources available to each school. Nevertheless, this contextualization should not be used to justify backwardness, but, rather, to stress that it is more difficult and costly to achieve success in certain environments and groups. The measurement of progress in this regard is now becoming a regular part of monitoring and evaluation in Brazil, and its results can be consulted on the INEP's website.

Linkage between the tests and the curriculum is achieved by summarizing the results in terms that educators can understand, as usually occurs in evaluation reports, so that they can be presented to students, teachers and parents, giving clear information about the know-how of students in given school levels.

In this regard, the INEP is working on a project called "Teachers' Views about the Educational Evaluations", which will begin in the second half of the year 2015. The said project consists of an Internet platform that will contain the items used in the *Prova Brasil* and descriptions of what each of them tests, its statistical parameters, and, finally, teacher comments. In this way, the competency-level data will make it possible to really assess student learning outcomes and issue reports as a basis for recommending curriculum changes.

### Evaluating Schools

Attendance at elementary school is crucial to compliance with/exercise of the right to education. The said type of schooling differs from other types in terms of its objectives and the basic organization of the program aimed at reaching them. Such schooling can be defined as the voluntary coming together of a group of professional teachers and a group of children and/or young people still involved in training, for purposes of personal-development, the former being charged with the task of instruction and education and the latter being afforded a chance to learn and become educated. The interaction between the two takes the form of a dialogue, though their roles are clearly different.

Since the way that we conceive of school evaluation strongly influences the way we gather data for monitoring compliance with/exercise of the right to education, one needs to reflect about this issue. Each and every organizational process, including those pertaining to schools, should be continuously evaluated. However, since the Brazilian government is legally bound to evaluate and certify the country's schools, we need to create yardsticks so as to comply with the said obligation.

The purpose of this type of evaluation is to identify changes in policies, practices and resources, in order to ensure that its yardsticks are suitable and hence foster improvement in all our schools. Though the general statutes and regulations governing improvement play an important role, they are no substitute for reflection by the staff in each school – i.e. evaluation is an internal activity and better defined as an organizational strategic-planning activity.

While there is still no consensus in Brazil as to which set of yardsticks should normally be used, based on the conceptual

model adopted in this article and the literature cited herein, we suggest that a school can be evaluated in terms of the following categories: *a)* Students: the number of these and their sociodemographic characteristics (socioeconomic and cultural levels, commitment and motivation); *b)* Resources: school infrastructure, teaching turns and loads, payroll; *c)* Teachers: teacher training, experience, commitment; *d)* Curriculum: study program (learning aims, relevance and pertinence, teaching methods, use of technology); *e)* Culture: learning emphasis, discipline, community links, shared decision-making; *f)* Management: leadership, organized processes (assignment of teachers, effective use of teaching time, accountability); *g)* Results: student learning outcomes, parent, teacher and student satisfaction; and *h)* Costs. The only ones among the above categories for which there are clear yardsticks and measurement processes are student results and costs.

With the introduction in 1995 of the Elementary Education Evaluation System (Portuguese acronym: SAEB), the use of scales to measure student competencies in reading and mathematics -two important teaching-learning areas whose results are very socially relevant- became widespread in Brazil. These scales were recently included in the elementary-education-development scale (Portuguese acronym: IDEB), which was

created in the year 2007 and can be defined as a performance yardstick based on the average student-competency level achieved in the *Prova Brasil*, and also on a performance yardstick that measures the average approval rate pertaining to the school or system, which is measured in the school census. The value of the IDEB increases or decreases in accordance with increases or decreases in student-learning-outcome levels. One of the reasons why it is widely respected is that it combines both an achievement and a performance yardstick -the two areas that determine compliance with/exercise of the right to education- in a single evaluation, while also giving a very good idea of students' socioeconomic level, and thus underlining the need to contextualize results.

Each year, the INEP disseminates information about how much each student in elementary education costs, obtaining this information via an annual survey that is administered jointly with the state-level and municipal ministries of education. The data show that, in 2013, Brazil invested 6.6% of its gross domestic product in education, and that it currently invests twice as much in the said activity as it did in the year 2000, which attests to the fact that the debate about compliance with/exercise of the right to education is leading the country to take concrete steps towards providing better education services to its citizens.

## Conclusions

In the last few years, Brazil has made enormous progress in reflecting on how to create an elementary-education monitoring-and-evaluation system. Various efforts in this direction have now been made, in the context of the SAEB itself, to broaden the range and dimensions pertaining to this process and, hence, those of its results. Based on this large body of experiences, efforts and reflections, we are striving to propose a model which -based on a clear vision that is shared by everybody, takes stock of everybody's needs, and fosters cooperation among the different bodies charged with gathering, analyzing and disseminating evaluation and monitoring results- will enable all those of us who are involved in this endeavor to work hand-in-hand towards achieving our shared aim.

To this end, we will strive to promote debate on this topic, in order to put together the pieces of this great jigsaw whose assembly requires all of our efforts. €

- 1 This consists in the answering of questionnaires about institutions, teachers, classes, and students, for each elementary school level and type. Each student is assigned a unique code, and socio-demographic data are (date of birth, race/color and current school level) are gathered. Each school is also assigned a unique code. Classes are assigned the names used in the school, Teachers are identified using their taxpayer-registration code.

## DOSSIER: COMPREHENSIVE AND CONTEXTUALIZED EVALUATION WITHOUT A PASSPORT

### Educational Reform and Public Policy design in the Latin-American countries: What do the Educational Ministers could do with PISA evidences?

The policy implications of PISA for Mexico and other Latin American countries are many. **Andreas Schleicher**, Director for Education and Skills, and Special Advisor on Education Policy to the Secretary-General at the Organisation for Economic Cooperation and Development (OECD) in Paris, knows the dynamics of a National Educational System in many countries with opposite and similar conditions. Students in all the world learn in different ways, but nowadays, all of them need to achieve better skills and knowledges. What are education leaders doing with the PISA evidences? How can we design public policies with its results?

### Since de beginnings of PISA: Approaches & methodologies have changed

The Programme for International Students Assessment (PISA) is an example of relentless innovation. Since the first round back in 2000, there have been new domains (i.e. problem-solving and collaborative problem-solving, digital literacy or financial literacy), different background questionnaires (i.e. cross-curricular competencies, educational career, reading for school, teacher questionnaire), innovative approaches to disseminate the results and new PISA test questions are constantly developed. A significant methodological change is the transition from a paper-based to an interactive digital assessment that has been implemented this year in most of the participating countries and economies.

However, PISA has kept its mission and focus to ensure comparability across countries, cultures and over time. This means:

**PISA is policy-oriented.** It focuses on providing data and analysis that can help guide decisions on education policy. By linking data on students' learning outcomes with data on key factors that shape learning in and out of school, PISA highlights differences in performance patterns and identifies features common to high-performing students, schools and education systems.

**PISA is carried out every three years** to enable countries to *monitor their progress* in meeting key learning objectives. The basic survey design has remained constant to allow for comparability from one PISA assessment to the next and thus to allow countries to relate policy changes to improvements in education outcomes.

**PISA assesses both subject matter content knowledge**, on the one hand, and the capacity of individuals to *apply that knowledge creatively*, including in unfamiliar contexts, on the other.

**PISA is designed to provide comparable data** across a wide range of countries. Considerable efforts are devoted to achieving cultural and linguistic breadth and balance in assessment materials. Stringent quality-assurance mechanisms are applied in the test design, translation, sampling and data collection. An age-based rather than a grade-based target population is used to ensure valid international comparisons of educational

performance. The PISA for Development project, that provides support to the least developed countries, is an effort aimed at enhancing the PISA instruments so as to make them available and more relevant for countries that have thus far excluded from global educational comparisons. The PISA-based test for schools allows individual schools to assess where they stand among the world's most successful schools.

**PISA is a collaborative effort.** Decisions about the scope and nature of the PISA assessments and the background information collected are undertaken by leading experts in participating countries. Governments oversee these decisions based on shared, policy-driven interests

#### **OCDE countries & PISA: uses with the results**

OECD countries and partner countries and economies participate in PISA because it provides reliable evidence about their students, schools and education-systems. This information is collected through a comprehensive student assessment and questionnaires to students, school principals, - parents and, starting in 2015, teachers. This evidence is used by countries to design and implement better educational policies. PISA allows countries to compare themselves with other countries and to learn from the world's top performing education systems.

Countries in PISA also gain insight by analysing changes in their own performance and equity levels over time across PISA waves. To provide an example: In Germany, the education policy debate and changes in light of PISA 2000 were intense (e.g. Ertl, 2006, Grek, 2009). Confronted with lower-than-expected results in student performance, PISA triggered a sustained public debate about education policy and reform that came to be known as 'PISA shock'. For example, equity in learning opportunities across schools had often been taken for granted, as significant efforts were devoted to ensuring that schools are adequately and equitably resourced.

The PISA 2000 results, however, revealed large socio-economic disparities in educational outcomes between schools. Further analyses that separated equity-related issues between those that relate to the socio-economic heterogeneity within schools and

those that relate to socio-economic segregation through the school system, suggested that German students from more privileged social backgrounds are directed into the more prestigious academic schools, which yield superior educational outcomes, while students from less privileged social backgrounds are directed into less prestigious vocational schools, which yield poorer educational outcomes, even where their performance on the PISA assessment was similar.

This raised the spectre that the education system was reinforcing, rather than moderating socio-economic background factors. These results, and the ensuing public debate, inspired a wide range of equity-related reform efforts in Germany, some of which have been transformational in nature: This includes giving early childhood education, that had hitherto been considered largely an aspect of social welfare, an educational orientation; establishing national educational standards in a country where regional and local autonomy had long been the overriding paradigm; or enhancing the support for disadvantaged students, such as students with a migration background.

For many educators and experts in Germany, the socio-economic disparities that PISA had revealed had not necessarily been surprising. However, it was often taken for granted and outside the scope of public policy that disadvantaged children would fare less well in school. The fact that PISA revealed that the impact which socio-economic background has on students and school performance varied so considerably across countries, and that other countries appeared to moderate socio-economic disparities so much more effectively, showed that improvement was possible and provided the momentum for policy change.

#### **Korea, Hungary, Portugal and Finland: success, a conscious result**

More generally, showing that strong educational performance, and indeed improvement, is possible seems to be one of the most important merits of international comparisons. Whether in Asia (like Japan, Korea, Singapore or Shanghai-China), in Europe (like Finland) or in North America (like Canada), many countries displayed strong overall performance in PISA and, equally important, showed that poor performance in school does not automatically follow from a disadvantaged socio-economic





background. Some countries showed that success can become a consistent and predictable educational outcome: In Finland, the country with the strongest overall results in PISA, the performance variation between schools amounted in 2009 to only 5% of students' overall performance variation, such that parents can rely on high and consistent performance standards in whatever school they choose to enrol their children. Last but not least, some countries have shown that significant educational improvement can be achieved within a limited time span. As noted before, Germany saw significant improvements both in quality and equity of its school systems between 2000 and 2009. Korea's average performance was already high in 2000, yet the Koreans were concerned that only a small elite achieved levels of excellence in the PISA reading assessment back then. Within less than a decade, Korea was able to double the share of students demonstrating excellence. A major overhaul of Poland's school system helped to dramatically reduce performance variability among schools, turn around the lowest performing schools and raise overall performance by more than half a school year. Portugal was able to consolidate its fragmented school system and improve both overall performance and equity, and so did Hungary. Even those who claim that the relative standing of countries in PISA mainly reflects social and cultural factors had to concede that educational improvement is possible.

### **Recommendations to Mexico and Latin American countries to design educational policies on the results of PISA**

The policy implications of PISA for Mexico and other Latin American countries are many. For example, investing resources where they can make the most difference is important. Among countries whose cumulative expenditure per student is below US\$ 50,000, higher expenditure on education is predictive of higher PISA scores. Latin American countries invest significantly less per student, and distribute these resources much more unequally across schools, than the average across OECD countries. Mexico has increased its educational expenditure as percentage of GDP since 2000, but the absolute annual spending per student is still much lower than the OECD average. Directing the resources to students and schools

who need them most can be an effective strategy. Shanghai-China provides a telling example in this respect, because it has very small variability in school performance despite of considerable social and economic inequalities in their population. This has not come about by chance, but rather as a result of considerable efforts to improve the school system by converting so-called weaker schools to stronger schools. In a large and complex educational system such as Mexico's, equity in the distribution of resources between schools is important. Some of Mexico's lowest performing schools in poor rural and indigenous regions tend to have fewer resources than urban schools. Funding mechanisms that allocate resources to the schools where they are most needed, within a model of clear goal setting, assessment and accountability, can increase the equity of resource distribution even among the most remote and marginalised schools.

### **Context in evaluation**

Schleicher says that life is unfair, so it is necessary to use the PISA test results, in a country like México, with so enormous differences between local regions, how do contextualized test?

PISA looks at students' ability to apply knowledge and skills under unknown situation. When students become adults and get involved in work and civil life, they are expected to mobilise skills and solve problems that they have never encountered. PISA provides an opportunity for education systems to better understand the skills of young people to live and thrive in this challenging life context.

One of the main strengths of PISA is that the results can be compared with other education systems around the world, and especially with other countries sharing similar socioeconomic characteristics. This is because PISA test questions have been designed so that they measure the same competencies across all cultures. However, even if test questions should be the same across participating countries, and regions within countries, the background questionnaires can be used to better understand the influence of specific local conditions, for instance by opting in or out of specific questionnaires or questions, and the interpretation of the findings should be adapted to the local context.

Moreover, some countries, like Mexico, have previously gathered data at the

subnational level. This opens the possibility of comparing educational performance and its relationship with student, school and system characteristics between the 31 Mexican States and the Federal District. In a country with sharp regional differences, like Mexico, these regional comparisons can help contextualise the results and provide with relevant policy implications. The results and policy recommendations for Aguascalientes, where the academic performance is almost as high as in Greece and above Chile, might be very different than those for Guerrero, at the bottom of the PISA rankings.

### **Mexican Educational Reform: policy recommendations to strengthen its implementation**

Two important goals of Mexico's education reform are the strengthening of the teaching profession (through the Servicio Profesional Docente) and of the system of evaluation of quality of education (through the Sistema Nacional de Evaluación Educativa). These efforts are well-aligned with the evidence about what the best educational system in the world do to foster student achievement. The quality of an education system cannot exceed the quality of its teachers and principals. Successful school systems have made teaching an attractive career and are able to bring and retain the best candidates in schools, where they are most needed. Successful systems also make intensive use of evidence to design their policies, provide feedback to teachers and students and to inform parents about their children's schools. To successfully implement these reforms, Mexican authorities should communicate to the public a clear vision about the aims and means of the reform, and ensure the coherence of all the new initiatives. Aligning coherent policies and practices over sustained periods of time is an important factor for the implementation of ambitious education reform.

### **Improve PISA tests in the current global context: what and why?**

In PISA we strongly believe that there is always scope for improvement and that is why we are continuously revising, fine-tuning and updating the methodology and the way findings are disseminated. Vital improvements are already being implemented, such as the transitioning to computer-based assessment and the teacher questionnaire, are currently being developed, such as the measurement of

global competences, or are waiting for their opportunity, such as the measurement of linguistic skills. These innovations are the result of the ongoing collaboration between the OECD, national governments and educational experts from all over the world.

Particularly, developing global competence is increasingly important in the current context where learning, working and living environments become ever more global, interconnected and interdependent. PISA 2018 will assess global competence as an innovative domain. This will be a significant innovation in assessing the knowledge, skills and attitudes of students that value di-

versity, openness, flexibility, and resilience, which are crucial for young people to be able to interact, communicate, study, work and live effectively in a global world.

A solid foundation of knowledge in key disciplines is an essential prerequisite for students to be successful in life. But it is also important that students develop creative, critical thinking and collaborative skills, and that they build character attributes, such as mindfulness, curiosity, courage and resilience. Together, these will enable to lead successful lives and contribute to the well-being of societies. €

Interview by Gazette's writing desk



## DOSSIER: COMPREHENSIVE AND CONTEXTUALIZED EVALUATION WITHOUT A PASSPORT

### The evaluation of educational reforms and policies: reflections on the Chilean experience<sup>1</sup>

Having “given a rough outline”, in this article, of the progress and limits of evaluation in Chilean education, and talked in more detail about the benefits of process evaluation and its nature and methodological challenges, the author ends by identifying some of the educational-policy challenges posed by the findings of the research and evaluations that have been carried out.

DAGMAR RACZYNSKI  
Senior Research fellow in Development Consultancy  
Member of the Faculty of the Pontificia Universidad Católica de Chile  
*dagmar.raczynski@asesoriasparaeldesarrollo.cl*

The proliferation of studies of, and research into, education in Chile, including evaluations of policy programs and initiatives aimed at raising the quality of teaching-learning processes, has led to the identification of problems in the sector, especially with regard to the “fit” between reforms, policies and programs on the one hand and schools on the other (Raczynski and Muñoz, 2008). These and other results have posed new thematic and methodological challenges both for evaluation and also for the formulation and implementation of policies.

#### Progress in, and limits to, the evaluation of educational policies and programs

Like other disciplinary areas, the evaluation of educational policies and programs has expanded and flourished in Chile over the last 10 to 15 years. Many different factors, having to do with government requirements, have driven this development.<sup>2</sup> Hence, there is plethora of information and knowledge at the macro, meso and micro-social levels about the realities of education. Until recently, most of the studies of education in our country focused on the learning outcomes obtained by students and schools,

the aggregate deviation trends (decline, stagnation, improvement nation-wide, etc.) and associated factors (students' socioeconomic level, participation of the school in a specific educational-improvement program, support for students at risk, etc.). Several of the said studies use the latest econometric techniques, and some use quasi-experimental designs and make the international rounds in conferences, seminars and well known scholarly publications. One of the factors that have facilitated these studies is the availability, since the mid 1990's, of the results of standardized learning tests –i.e. ones pertaining to the Chilean Educational Quality Measurement System (Spanish acronym: SIMCE)- administered to the universe of fourth-, eighth- and tenth-grade students in the areas of Languages, Mathematics, Social Studies and Natural-environment Studies. Chile's participation in international and Latin American tests has also resulted in important contributions to knowledge about trends and learning gaps, both in the said country and between it and other parts of Latin American and the world.

Overall, these sources have identified the social sectors or geographical regions where student learning outcomes are very inadequate, ascertained that there are very marked academic and social disparities in the school system, and debunked the widespread myth that private schools are better at managing education than government ones. However, the information produced for the purpose of designing social policies and programs is scant and only indirectly



useful and accessible, helping the authorities to establish priorities, but not to formulate concrete policies. Some research projects evaluate the impact of specific, focalized programs, taking, as a dependent variable, the SIMCE results, which show a low or zero impact and, in cases where duration has been measurable, only a fleeting effect. These results come as no surprise, when one considers that a program's impact on student learning outcomes is never immediate, since many extraneous factors are involved, and should be measured three, four or more years after implementation, rather than the same or following year, as often occurs with research.

An evaluation that only measures impact does not yield data that enable us to ascertain the program's effect on the internal dynamics of the school, its administration, or its teaching practices. Nor does it show us which aspects of the program worked and which didn't, or why. To obtain this information, we need evaluations of processes and studies of mid-term results, the first of which should ascertain how the program was divulged to the schools, how it was perceived by the school community, and to what extent it was understood and taken on board and internalized by the latter, fitted in with its program, conditions, needs and existing competencies, and was coherent with the previous courses run there.

It bears mentioning that, when talking about mid-term results, we are not referring, as is often assumed, to the activities emanating from the program -the "products" as they are often called in goal-oriented or logical-framework-approach language- but, rather, to the imprint that the said activities will leave on the school and its pupils and management, the systems that will be left up and running.

Increasingly, in Chile, the terms of reference of the studies and evaluations for which the Chilean Ministry of Education (Spanish acronym: MINEDUC) invites companies to tender require the latter to use qualitative methodologies -case studies, open interviews, focus and discussion groups, *in situ* observations- with broad, statistically representative, samples of students, teachers and schools, in order to complement information gathered from official statistical records and structured surveys.<sup>3</sup> This is a big step forward, since it places the emphasis on the individual school as a working organization,

and on the relations between the members of its community. It's a matter of finding out, in detail, what problems impede implementation and which situations were not foreseen at the design stage, and basically of making sense of micro-level educational-policy developments.

### **Some methodological challenges in evaluation process**

The qualitative methodologies are less codified; the field work takes longer, the samples are smaller, the statistical-representation criteria are weak, and one has to make substantiated choices about case studies at the school and classroom level or situations that merit observing and decide which situations are worth observing, and, very importantly, make it clear what role the observer is going to play in the functioning of the school.

As far as the evaluation is concerned, it's a matter of making a comprehensive, systemic diagnosis of the realities of the school and its surroundings that includes suitable input from insiders. Also, it's essential to take stock of the school's development and track record and its successes and failures, placing the program to be evaluated in this context. The different teams and individuals who make up the school community will give varying accounts of key processes. For example, one teacher will assert that the backwardness of a given student might be due to social and family reasons, while a school official might attribute them to the teacher's low expectations for the student, a member of the administrative staff to too much government red tape and ministry-of-education requirements that make it hard for him to do his job, and an academic coordinator to incompetent teachers. In this context, the research team is hard set to reconcile these viewpoints and, based on them, reach conclusions about processes, crucial issues and alternative improvement paths.

In Chile, members of school communities -especially school principals and teachers- tend to mount long-winded defenses regarding their students' bad results, blaming other participants in the education process, talking about "what ought to be" and what is "socially desirable". Hence, in our evaluations, we've learned that: (a) questions about how often certain practices are carried out tend to elicit more informative answers, teachers talk more honestly about what the school principle does than the

principal himself, and students reveal more about what their teachers do than the teachers themselves; and (b) it's important to supplement interviews and questionnaires with direct observations of different situations in the school. In general, we've chosen to do case studies of schools, the results of which are combined and triangulated with quantitative and document-based information. The cases to be studied are chosen carefully, based on explicit criteria and depending on the school and social environments in which they are carried out.

### **Challenges for the educational reforms, policies and programs**

There are challenges for educational policy at least in three interrelated levels.

#### **a) The work that policy practitioners do with schools**

The literature about school-improvement processes and their effectiveness emphatically stresses that such improvement only occurs if it is implemented from inside the school. Schools do indeed need external support in order to improve, but such support won't achieve lasting results if it doesn't involve the school community, which must believe that change is possible, and have the will to achieve it and the initiative to lead it. Current educational policy in Chile requires each school receiving the Preferential School Subsidy (MINEDUC, 2012) to carry out a self-diagnosis and put together a four-year educational-improvement plan (initially these plans covered one year). The said self-diagnosis is the first step, requiring outside support in order to detect blind spots and establish priorities for action based on the school's strengths, weaknesses, opportunities and threats.

The support subsidies granted to schools in Chile are small and vary a lot, being split among three different recipients - supervisors working in provincial education departments (regional MINEDUC branches); private and municipal school board; and the market or system of private External Technical Support entities (Spanish acronym: ATE). On the one hand, it is essential to raise the standards, and increase the competencies, of this support, and, on the other hand, the support provided should focus on promoting profound reflection about teaching in school communities. Unless efforts at improvement have support both within, and



from outside, school communities, many of the policy initiatives will have only a short-term positive impact at best.

### b) Structured but flexible programs: first- and second-order designs

Almost without exception, ministry-of-education programs in Chile are standardized ones, in the sense that their design pre-defines the actions to be taken in the school, the times and sequence assigned to them, the *modus operandi*, etc. This format leaves little leeway for adapting to each school's specific situation, making it hard to achieve a proper "fit" between the exigencies of the program and the needs, strengths and problems of the school in question. In this regard, it serves to distinguish between two phases in the design of a policy or program – a first-order phase that concerns macro-social and political decisions (the aims pursued –e.g. comprehensiveness and participation- and the strategic orientations and principles that it is hoped will guide the actions) and a second-order phase pertaining to the design that the program's "local operators" create based on the first phase (Concha *et al.*, 2001). The second-order design is the task of those responsible for implementing the policy, generally seen as "executors" who have nothing to do with the design. One of the big challenges for social and educational policies is that of reclaiming these people's role in, and responsibility for, the design, and preparing them for this task.

### c) Coherence in the policy decisions and the regulatory framework

In Chile, and perhaps in other countries in the region, we urgently need to curb the proliferation of programs and actions that are implemented in -or affect- schools at different times. In our country, we have teacher-development programs alongside pedagogical-improvement and -innovation programs involving training and betterment, along with a textbook-distribution program and various other programs involving textbooks and other teaching materials, to mention just a few examples. This *modus operandi* of the Ministry has prevented us from getting a comprehensive overview of the school and its real, specific needs. Policy should be slanted so as to foster an approach that aligns the different supports offered by the Ministry with the needs and possibilities of the school and its managerial staff, teachers

and pupils, thus achieving a better "fit" between the two. It's a matter of achieving coherence not only among the programs, but also with regard to the legal and regulatory framework and ground rules that underpin the system, simplifying them, reducing ambiguities, and eliminating contradictions.

### Achieving better linkage between evaluation results and policy-making

The main challenge for evaluation consists in achieving methodological rigor along with an ability to provide valid information and conclusions. Nevertheless, in our opinion, the evaluator should also assume other responsibilities, including the paramount one of doing everything possible to ensure that the evaluation's findings really inform the decision-making processes. To achieve this, we need to: (a) spend time understanding the needs of the entity that seeks the evaluation, its reasons and purposes and its concerns about the program or policy, as well as its expectations regarding the results and processes, (b) understand the frame of reference and language of those responsible for the program or policy; (c) ensure regular contact between the two sides, involving feedback about the expected results and the strengths and weaknesses detected, so as to ensure that the information produced is adequately taken on board by those responsible for implementing the programs; (d) prepare the results and insist on the importance of disseminating them, once the evaluation is over, so that other people and entities with a legitimate interest in the evaluated programs (including the beneficiaries and recipients) may become acquainted with them, take part in the process, and give critical feedback about the policy or program. €

- 1 The research supporting this article is based on studies and evaluations of schools and programs that were carried out by the Development Consultancy team, parts of which are cited in the "Reference" section. For a description of educational reforms and policies in Chile over the last 20 years, see García Huidobro (1999), Cox (2003), Raczynski and Muñoz (2007), MINEDUC (2012), Raczynski *et al.* (2013).
- 2 Raczynski and Salinas (2008, Table 1) show some of these factors in Chile and specifically underlines the inclusion, since 1997, of the evaluation of government programs in the Management-control and Budgeting-by-Results System of the Chilean Tax Ministry's Budgeting Department ([www.dipres.gob.cl](http://www.dipres.gob.cl)) and in public tenders for studies and consultancies aimed at evaluating social programs ([www.mercadopublico.cl](http://www.mercadopublico.cl)).

- 3 Some of these studies can be found at: [www.centroestudios.mineduc.cl](http://www.centroestudios.mineduc.cl)

### References

- Asesorías para el Desarrollo and Santiago Consultores (2000). *Evaluación del Programa de Mejoramiento de la Calidad de las Escuelas Básicas de Sectores Pobres*, p. 900. *Informe final de consultoría. Resumen ejecutivo*. Santiago de Chile: MINEDUC, 2001. ([www.asesoriasparaeldesarrollo.cl](http://www.asesoriasparaeldesarrollo.cl), Study Report section).
- Asesorías para el Desarrollo (2015). *Estudio sobre el funcionamiento de los servicios de Asistencia Educativa (ATE)*. Final consultancy report, Santiago de Chile: PNUD-MINEDUC.
- Bellei, C. et al. (2004). *¿Quién dijo que no se puede? Escuelas efectivas en sectores de pobreza*. Santiago de Chile: UNICEF.
- Bellei, C. and A. Osses (eds.) (2010). *Asistencia Técnica Educativa en Chile ¿Aporte al mejoramiento escolar?* Santiago de Chile: CIAE-Universidad de Chile.
- Concha, J. et al. (2001). *Superación de la pobreza y gestión descentralizada de la política y los programas sociales*, pp. 173-270, in Raczynski, D. and C. Serrano, *Descentralización. Nudos críticos*. Santiago de Chile: CIEPLAN-Asesorías para el Desarrollo ([www.asesoriasparaeldesarrollo.cl](http://www.asesoriasparaeldesarrollo.cl), book section). Decentralization. Critical nodes, part two).
- Cox, C. (ed.) (2003). *Políticas educacionales en el cambio de siglo. La reforma del sistema escolar de Chile*. Santiago de Chile: Editorial Universitaria.
- García-Huidobro, J. E. (ed.) (1999). *La Reforma Educativa chilena*. Madrid: Editorial Popular.
- Marcel, M. and D. Raczynski (eds.) (2009). *La asignatura pendiente. Claves para la revalidación de la educación pública de gestión local en Chile*. Santiago de Chile: Uqbar-CIEPLAN.
- MINEDUC (2012). Bill placed before the Senate increasing state subsidies for schools.
- PNUD-Asesorías para el Desarrollo (2004). *Sistematización y evaluación del Plan de asistencia técnica a escuelas críticas de la Región Metropolitana*. Final report to the Ministry of Education, Santiago de Chile.
- PNUD-Asesorías para el Desarrollo (2008). *Monitoreo, seguimiento y evaluación de la Estrategia de Apoyo a Establecimientos Prioritarios*. Final report to the Ministry of Education, Santiago de Chile.
- Raczynski, D. and G. Muñoz (2005a). *Supervisión educacional en Chile. Experiencias públicas y privadas*. Santiago de Chile: Ministry of Education
- Raczynski, D. and G. Muñoz (2005b). *Efectividad escolar y cambio educativo en condiciones de*

*pobreza en Chile*. Santiago: Ministry of Education, General Education Division.

Raczynski, D. and G. Muñoz (2007). *Chilean educational reform: the intricate balance between a macro and micro policy*, in W.T. Pink and G. W. Noblit (Eds.), *International Handbook of Urban Education*. Springer.

Raczynski, D. and D. Salinas (2008). *Aportes y desafíos en la evaluación de políticas y programas de mejoramiento educativo. Reflexiones desde nuestra práctica de evaluación en Chile*, *Revista Iberoamericana de Evaluación Educativa*, pp. 79-92.

Raczynski, D. et al. (2013). *Subvención Escolar Preferencial (SEP) en Chile: un intento por*

*equilibrar la macro y micro política escolar*, *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación (REICE)*. Volume 11. No. 2, pp.165 -193.

Raczynski, D. et al. (2014). *Diagnóstico de la capacidad institucional para el monitoreo de la calidad y provisión de soporte pedagógico en el sector educativo chileno*. Final consultancy report. BID.

## DOSSIER: COMPREHENSIVE AND CONTEXTUALIZED EVALUATION

### OTHER PERSPECTIVES

## Black boxes and black holes: evaluation and assessment of the Educational Reform in Mexico

In these lines, the author reflects on the distinction between evaluation of the Mexican Educational Reform and assessment of the latter, using the “black box - black hole” metaphor customarily applied to the evaluation of study programs and government policy respectively, and also talks about the importance of evaluating the analysis of the socio-institutional contexts of government actions.

ADRIÁN ACOSTA SILVA

Teacher-researcher at the University Centre for Economic and Managerial Sciences of the University of Guadalajara  
[aacosta@cucea.udg.mx](mailto:aacosta@cucea.udg.mx)

The “black box – black hole” metaphor is often used by social scientists, including government-policy analysts, to emphasize the enormous –possibly total– ignorance that exists about the factors or processes that determine the development, make-up and degree of complexity of social phenomena. For example, in political science David Easton’s classic study used the aforesaid metaphor to stress the importance of the processes whereby input (resources, interests, demand) is transformed into results (decisions, policies) in the context of an analysis of a political régime (Easton, 1965).

This metaphor has continued in use and, in some ways, experienced a vigorous comeback, in recent years, in the field of public-policy analysis, in the context of a growing theoretical and practical debate about the strategic role of evaluation as a factor in government actions - a debate that springs from a scrutiny of the very idea of evaluation as a means of improving government actions or making them more efficient, and questions its effectiveness as a mechanism which, *per se*, guarantees better quality and ensures that government policy will succeed in solving public problems. A brief review of the terms of this debate may help us to reflect on the nature, methods and aims of the Educational Reform in Mexico.

The said discussion concerns the basic tenets of the evaluation of policies and programs, one of which is, precisely, the cause-and-effect linkage between problems and policies. This tenet can be expressed as follows: since programs are instruments for the implementation of policies that are deemed to be public, their evaluation will help us to measure the efficiency of the public resources invested in order to solve problems. Also, it will eventually be possible to estimate or gauge the impact or effects of the actions carried out in accordance with problem-solving programs.

This optimistic view of evaluation tends to be associated with a simplistic narrative that assumes that public problems are a matter of one-on-one cause and effect, leading to a certain degree of reductionism regarding the complexity and diversity of the factors that affect the make-up of government reforms. Often the evidence for these other causative components (unidentified, unperceived or deliberately ignored) is “suppressed” in order to avoid changing the *ex ante* cause-and-effect relationship deemed to justify the policies (Chelimsky, 2012).

Though this reasoning would often appear to be implicit in the arguments legitimizing evaluation (Julnes, 2012), nevertheless, as pointed out earlier, an analysis of actual evaluations of public policy implementation in different areas shows that it is often wrong for two reasons – on the one hand because it tends to separate program analyses [“black boxes”] from context analyses [“black holes”] (Patton, 2012), and on the other hand because the logic of policy evaluation tends to be confused with the logic of policy valuation or assessment (Scriven, 2012).

Hence, the new policy-evaluation trends have gone from examining the “black box” of government policy to an emphasis on deciphering the “black hole” of public-problem contexts. Furthermore, faced with the fact that institutions tend to evaluate programs with a view to legitimizing government interests, one observes a dearth of intellectual and conceptual rigor in their evaluation of how the said programs affect public interests. The predominance of a clear neo-positivist, quantitative tendency can be seen in the area of evaluation, while the fundamental role played by qualitative assessment in the analysis of public policy has decreased.

The risks inherent in these causative associations, unsuccessful practices and conceptual confusions lie in the way that they favor legitimizing the results of government policies over assessing the possible causes of public problems. The solution, on the one hand, is to deem the causes to be explicit hypotheses, rather than incontrovertible facts or evidence, and, on the other hand, to combine the technical evaluation of programs or policies with an assessment of their components and social impact on the fundamental public problems identified. This implies: (a) acknowledging the complexity of both the contexts and the assessment of the public policies, and

(b) using different quantitative and qualitative instruments for the purpose of technically evaluating the programs and publicly (and possibly politically) evaluating the policies.

### Problems and prospects for evaluating the educational reforms in Mexico

The aforesaid considerations would appear to be feasible and relevant to the analysis of the performance indicators, effects and impacts of educational policy in Mexico. As one is aware, the Mexican 2012-2013 Educational Reform focused on the general aim of increasing the quality of education and making it fairer by reforming guidelines, amending articles 3<sup>o</sup> and 31<sup>o</sup> of the Constitution, issuing a presidential decree and establishing the autonomy of the National Institute for Educational Evaluation (Spanish acronym: INEE), and it was carried out in accordance with the five guiding principles of “School at the center”, “Professional teaching service, quality and fairness”, “Institutional management”, “Social inclusiveness and coexistence”, and “Publishing program” (SEP, 2013).

The evaluation of these initiatives forms a large part of the Mexican government’s effort to regain authority in the area of education, and, since the present federal government took office, has been the political *raison d’être* of the Educational Reform policies for the present presidential term. In this regard, government policy has focused on different topics, areas and aims: improving educational infrastructure (high-quality and full-time schools); reforming the mechanism for hiring teachers (a new labor relationship associated with a teacher-evaluation-and-incentive system as a key tool for improving teaching quality and learning outcomes throughout the public-school system; improving school management and increasing its autonomy while also strengthening links with central federal and state-level authorities; recognizing that schools play a part in promoting social cohesion and thus giving children and young people a sense of belonging and social identity; bolstering state textbook publishing as a way of supporting educational endeavors throughout the country.

Since these programs imply different causal relationships, they may have different impacts on the make-up of the public problems that they seek to solve. One needs to take

stock of the logic of contextual black holes and also of the logic of assessing their impact on public problems in order to clearly identify the difficulties of implementing the reforms in the different contexts of the Mexican education system and the impacts that the said reforms have on the nature of the educational problems in each of the areas to be reformed.

For example, the Professional Teaching Service (Spanish acronym: SPD) is an apt subject for study and contextual and evaluative analysis. The basic assumption behind the program is that, since teachers are the crucial factor in the endeavor to raise educational quality, improving teacher entry requirements, retention and promotion will have a positive impact on teaching, learning and the development of students’ cognitive and technical skills. However, the relationship(s) between the said program and the other programs that make up the Educational Reform is/are not clear.

Notwithstanding the temptation to evaluate the program’s efficiency (in terms of how many new teachers join the system, what professional qualifications they have, how much they are paid, and what incentives they receive) and link all this to the impact on teaching-learning (in terms of terminal student efficiency, academic achievement, effective learning outcomes that can be compared with others both in Mexico and abroad), this may lead to a technocratic illusion that typifies the prevalence of government interests over public ones. A more consistent and far-reaching effort to assess the said program would have to link quantitative measurements with a qualitative valuation of the contextual features (types of school, available infrastructure, accumulated managerial, teaching and academic experience, type of parent participation, teaching staff) and with the strategic values that are meant to be examined (relevance, fairness, cohesion) so as to achieve higher-quality approaches for ascertaining the impact of the programs on improving performance in Mexican public schools and the causes and effects underlying the said impact.

The biggest challenge faced by the reforms is that of assessing their social and public effectiveness, in addition to measuring the efficiency of the actions taken by the government. This means acknowledging the complexity of assessment in the field of education: a complexity that implies

identifying different methods and approaches, combining quantitative tools and qualitative instruments for assessing the impact of public -i.e. government and social- actions aimed at improving our country’s elementary-education system. The use of approaches such as policy meta-evaluation; valuative program ethnography; a system for monitoring students throughout their elementary education, from pre-school to upper-secondary; analysis of the public value of grams at the state and regional levels, or of Mexican and foreign efforts to design and implement policies based on evidence about student achievement (e.g. TALLIS results), all form part of the basic methodological mechanism for creating information systems that are powerful enough to evaluate and assess the Mexican government’s Educational Reform policies in their different socio-educational contexts. To that end, the support of entities such as the Information and Educational-management System (Spanish acronym: SIGED), set up under the Reform, and technical input from the INEE, constitute strategic institutional resources for building a model for evaluating and assessing the policies aimed at achieving the long-term aims of the said Reform – i.e. raising the quality, and increasing the equity, of the Mexican education system, based on an acknowledgment of its wide diversity and great complexity. €

### References

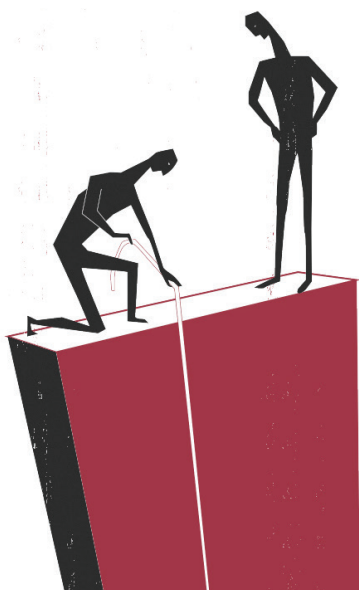
- Easton, David (1965). *A Framework for Political Analysis*. Englewood Cliffs, N. J.: Prentice-Hall.
- Julnes, G. (2012). Managing valuation, In G. Julnes (Ed.). *Promoting valuation in the public interest: Informing policies for judging value in education. New Directions for Evaluation*, 133, Spring 2012, pp. 3-15.
- Secretaría de Educación Pública (2013). *Programa Sectorial de Educación* (Education Sector Program) 2013-2018. Mexico: SEP.
- Scriven, M. (2012). The logic of valuing, In G. Julnes (Ed.). *Promoting valuation in the public interest: Informing policies for judging value in education. New Directions for Evaluation*, 133, Spring 2012, pp. 17-28.
- Patton, M. Q. (2012). Contextual pragmatics of valuing, In G. Julnes (Ed.). *Promoting valuation in the public interest: Informing policies for judging value in education. New Directions for Evaluation*, 133, Spring 2012, pp.



## Towards a possible future: the expectations for, and challenges facing, educational evaluation in Mexico

In the interview transcribed below, **Felipe Martínez Rizo**, a researcher at the Autonomous University of Aguascalientes, reveals the historical background of the first evaluations carried out in the U.S.A., talks about their impact on that country and the pros and cons of large-scale international evaluations, and envisions the future of educational evaluation in Mexico.

Felipe Martínez Rizo, a former rector of the Autonomous University of Aguascalientes, holds a bachelor's degree in the Social Sciences from the University of Louvain in Belgium. He asserts that educational evaluation has a very promising future and that "while the word, 'evaluation', was formerly used, almost exclusively, to refer to the analysis of student learning outcomes,



the new instruments and technical tools that are currently available enable us to evaluate entire education systems.

He points out, with regard to evaluation and student learning outcomes, that it is necessary to make a distinction between large-scale evaluations on the one hand and classroom-level evaluations on the other. "The latter," he affirms, "have always been carried out by teachers, while large-scale evaluation is a fairly new phenomenon, since the tools needed for it have only recently become available," going on to comment that "Standardized tests that can be given to large numbers of students and computer graded have only recently been developed. In our country, they were very timidly introduced in the 1970's and 1980's, with their use expanding in the 1990's".

### Large-scale evaluations: a historical overview

"The first large-scale evaluations," explains Rizo, who is a specialist in quality, planning, evaluation and inequality in education, "were introduced at the national, rather than international, level, originating in the U.S.A. and then spreading to Anglo-Saxon countries such as Australia and England, where trends from the US tend to arrive first, before being adopted in the rest of the world in the 1970's.

"The International Association for the Evaluation of Educational Achievement (IEA) was established in the 1970's due to a very specific event – i.e. the launching by the Soviet Union, in 1957, of Sputnik, the first artificial satellite, which was seen as a threat by the U.S.A. because it showed that it might lose the next world war, since another nation was winning the space, and hence the arms, race. It was in this way, due to the West's concern about losing the Cold War, those international educational evaluations in the areas of mathematics and the sciences began to appear".

"In Latin America the first international tests, pertaining to UNESCO's *Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación* (Latin American Laboratory for Assessment of the Quality of Education: Spanish acronym: LLECE), were implemented in the 1990's, though national tests had previously been implemented in countries such as Mexico, Chile and Columbia".

"The large-scale international tests appeared later, when several countries expressed an interest in doing something together, which is why the **Trends in International Mathematics and Science Study** (TIMSS) and the LLECE tests chose to analyze the curricula of the countries participating and seek a common denominator. In contrast, the Program for International Student Assessment (PISA),<sup>2</sup> initiated in the year 2,000 by the Organization for Economic Cooperation and Development (OECD), had a different focus, setting out to evaluate life competencies rather than the contents of the different countries' curricula".

### Classroom-based evaluation and multiple-choice tests.

"In the United States," continues Martínez Rizo, "large-scale evaluations began to appear at the start of the last century, due to the widespread dissatisfaction with the quality of education and the evaluations carried out by teachers at that time. In response to the need for instruments that made comparison possible, on November the 18th, 1900, a non-profit organization specializing in the creation of tests called the College Board, which currently collaborates with 6,000 educational institutions all over the world, was set up in that country to facilitate access to higher education."

"In the late 1900's, the technology for processing multiple-choice questions, which became the standard evaluation format and also became common in Mexico, was developed and began to grow."

"In the 1960's, a course about how to design multiple-choice questions and score them using a curve was given in our country's teacher-training colleges, with the result that tests based on the said type of question became accepted as the norm and teachers were encouraged to copy them. However, over time, the teacher-training colleges, which are a badly neglected area of initial training in our country, have ceased using both large-scale evaluations and formative classroom tests."

Martínez Rizo, who was awarded the 1991 Aguascalientes Prize for Development in the Arts and Humanities, acknowledges: "With regard to the limits of large-scale tests, while some people take very radical positions regarding such tests, others take more

nuanced stances, asserting that they are useful because they are the only ones that can be used to compare whole education systems. However, they should not replace good teacher-implemented evaluation but, rather, complement it. The most important evaluation is the evaluation carried out by teachers in their classrooms, which is why it should be much richer and have a formative emphasis.”

Speaking about the ways in which Mexico could benefit from the results of large-scale tests, he comments: “Following the implementation of the National Teaching Career Program following the signing of the National Agreement for the Modernization of Elementary Education (Spanish acronym: ANMEB) in 1992, the use of large-scale tests increased in primary schools, and subsequently, in secondary ones, because they were a way for teachers to earn points and have access to Teaching-Career stimuli. Later, when the census-based National Evaluation of Academic Achievement in Schools (Spanish acronym: ENLACE) appeared in 2006, it was administered to 10 million children and gave the impression of being the first of its kind, since, although around 8 million tests had been given the year before, they were not universal, being administered only to the students of teachers who wanted to get a stimulus, with the exception of teachers in private schools, in some government schools, and in indigenous schools. Nevertheless, the results in question weren’t used to improve curricula or study plans, but merely as a means for teachers to earn Teaching-Career points. They didn’t have a big impact because they didn’t undergo any other kind of analysis and weren’t disseminated”.

### The establishment of the INEE and its impact on public-policy formulation

“Starting in 2003”, says Martínez Rizo, “when the National Institute for Educational Evaluation (Spanish acronym: INEE) began, before the Amendment to the Law that granted it independent status, to be responsible for carrying out evaluations and disseminating their results, the latter had a certain degree of impact stemming from the administration of the Examination of Quality and Academic Achievement (Spanish acronym: EXCALE), since there was interaction with the curriculum-design area of the Ministry of Public Education (Spanish acronym: SEP) regarding the publication of the results.

“If the evaluation data are published, it’s a more serious matter for the education authorities to say “No, we don’t agree and we won’t do it”, since this could lead to very strong social pressure if they don’t have very sound reasons for refusing. That’s why evaluation is also a tool that can empower society to demand accountability”.

“Until 2004, the tests didn’t have any other impact, but they did have it between 2004 and 2006, because there was a good relationship with the SEP and the results were taken into account, though, in contrast, from 2007 on, ENLACE was only used to earn Teaching-Career points. This leads us to conclude that tests aren’t either good or bad per se; it depends on how they’re used”.

With regard the INEE, Martínez Rizo, who served as the latter’s first general director from 2002 to 2008, comments: “The setting up of the Institute is one of the many positive steps that Mexico has taken in the field of education, since, to date, it’s the only Latin American country whose constitution bestows such autonomy on the entity in charge of evaluation”.

In his view, the negative attitudes that currently prevail are not well founded: “We have to be aware of the educational problems that we have, but also of the positive things. We’ve made a lot of progress, and, if we look back and ask ourselves whether everything was fine and dandy before, we’ll be forced to admit that it wasn’t. With all our limitations, and the challenges that we face, we are clearly better, in every way, than we were thirty years ago. €

Interview by María Cristina Tamariz Estrada

- 1 Evaluation of mathematical and scientific knowledge in fourth- and eighth-grade students in various countries, developed by the IEA.
- 2 They set out to evaluate the extent to which students about to exit compulsory education have acquired the knowledge and skills that they need in order to fully participate in the knowledge society, and are administered each 3 years, in specific subject areas. All the previous PISA evaluations have focused on a specific area: Reading (2000), Mathematics (2003), Science (2006), Reading (2009), Mathematics (2012), and Science (2015). Source: [www.oecd.org/pisa/pisaenespaol.html](http://www.oecd.org/pisa/pisaenespaol.html)

## DOSSIER: COMPREHENSIVE... OTHER PERSPECTIVES

### The school as comprehensive educational-evaluation center

The Educational Reform stresses the efficient functioning of schools because it is there that government-provided education takes place, asserts the economist, **Carlos Mancera Corcuera**, who, in the interview transcribed below, expresses his unabashed admiration for teachers.

Interviewed in his office, Carlos Mancera, who holds a degree in Economics from the *Instituto Tecnológico Autónomo de México* (ITAM), talked about the implications of, challenges to, and expectations for, educational evaluation in the context of the Educational Reform, and discussed the role of the National Institute for Educational Evaluation (Spanish acronym: INEE) in the setting up of the National Educational Evaluation System.

“The Reform,” comments Mancera Corcuera, “places emphasis on the school as the center of the education system, so it’s essential that evaluation give priority to supporting effective school management”.

“Since the Reform contemplates a series of policy components that must coincide in order to improve the running of schools, one needs to ascertain what’s happening in the said areas. However, little is known about these day-to-day internal processes, and, to evaluate them, we need to administer questionnaires to school principals, teachers, parents and students, and have evaluators and other specialists visit the schools to gather the kind of information that can’t be obtained via a questionnaire. These people would have to be qualified observers or evaluators who are not hired just to make sporadic visits, but to study the schools on a regular, almost routine, basis, using the appropriate instruments in order to gather and analyze information and issue reports. The more complete the observations carried

out in the schools are, the better the analyses will be, as will the recommendations made, and judgments issued, about what is observed”.

### **Minimal normalcy and the learning environment**

Mancera Corcuera, who was Deputy Director for Scientific and Technological Policy of the National Science and Technology Council (Spanish acronym: CONACYT) in 1994, having joined the Ministry of Public Education (Spanish acronym: SEP) in 1992, the year when the National Agreement for the Modernization of Elementary Education (Spanish acronym: ANMEB) continues: “For a school to function well, it must be able to make decisions without enjoying total autonomy.”

When asked which rules govern the relationship between schools and education authorities, what limits there are on what schools can do, and at what point the supervision by state-level education authorities comes into play, he replies: “There’s very little information about, and almost no evaluation of, such relationships. For example, in the area of material resources, schools that have received such resources under the Reform Program (formerly the Program for Excellency in Schools Aimed at Reducing Academic Backwardness) can use them as they require, but we need to make sure that these new powers, which are very much welcomed by the schools, are governed by certain general rules.”

“As far as teaching practices are concerned, a school that maps out an improvement path can’t just follow, word for word, the contents of the national curricula and study plans; it needs to develop a capacity for creative teaching, so that teachers can do their job and, in their elementary-level students, foster the development of pertinent capabilities that must be defined at the national level, rather than limited to the local one. We can’t vary our definitions of the required mastery in reading and writing according to context, because this would make them subject to relative criteria, leading to the unacceptable conclusion that it would be alright for children living in underprivileged circumstances to half learn to read and write. All children should be able to perform these tasks well, and what needs to vary according to context is the way they’re taught”.

When asked what should and should not be contextualized, Mancera Corcuera, who was the SEP’s Undersecretary for Planning and Coordination from 1994 to 2001, answers: “Minimum-normalcy conditions shouldn’t be a matter of context. Regardless of the context, no school should open late or not open at all, or accept the idea that its students don’t need to master essential capabilities. Achievement of the learnings expected from all Mexicans in elementary schools shouldn’t be a matter of context. Rather, we need to contextualize the teaching processes in order to ensure that academic aims are achieved”.

### **Comprehensive evaluation: from photograph to movie**

When asked about comprehensive evaluation, he stresses that the latter has inherent value because it takes a photograph, but insists that it is richer when it makes a movie: “One of the big challenges for this type of evaluation, in the context of the Educational Reform, is to turn it into a program that, at the very least, ensures that certain basic features are covered in order to make an orderly movie. There should be a script that makes the various components coherent and meaningful. The purpose of evaluation shouldn’t be to make comments, but, rather, to help provide everybody involved in education –students, parents, teachers, school principals, authorities, researchers and society in general– with information about what’s happening in the field of education.

“If the purpose of evaluation is to achieve improvement,” says Mancera Corcuera, who is currently a partner of a company that specializes in education and culture, “then there’s all the more need to make a movie. Doing this isn’t easy; it’s not a matter of sticking a series of pretty photographs on a card and then showing them side-by-side. We need to make sure that the photographs can be compared with each other at different times so that, when we compare the results of one set of evaluations with those of the previous one, we can make sure they’re meaningful. This is very complicated and it’s going to take the INEE quite some time to develop such a movie-like evaluation system.”

### **The different types of evaluation**

When asked about the different types of evaluation, our interviewee responds:

“Student evaluation can be relatively fast because there are antecedents and we know very well how to do it. We can get the first frame of our movie in Year 1. Why not?”

“As for school evaluation, once the system has been set up, which can take several years, starting with the first evaluation we’d get the first frame of our movie, while also taking photographs that will enable us to get a better idea of what’s happening in our schools.”

“The big new challenge is evaluating teacher performance. The INEE has already made significant progress in this direction. It’s a process that needs to be carried out methodically, respecting teachers’ rights and being sensitive to the contexts in which they work. This is important not only to ensure that evaluation helps teachers, but also because they have rights in this regard under the Educational Reform. Since this type of evaluation has impacts, it would be wrong to suppose that there would be a period of initial trials in order to subsequently make radical corrections. We need to get it right from the outset. Even if the launch is successful, it’ll take at least five years to shoot the frames for the movie, plus a complete evaluation cycle in order to make a comprehensive assessment.”

“We’ll have to see whether the instruments that have been designed for the Competitive Examinations for entry to the Professional Teaching Service are the right ones for the medium term. Of course, they’re the right ones for the current situation, since they’ve enabled us to measure capabilities in applicants for teaching jobs that were unknown before, when there was no entry exam. But the examinations need to evolve in order to measure the results in a more precise way and compare them with those obtained by the teachers during their first two years of service, when they were still in the learning phase and on probation.”

“In accordance with the provisions of the Law Governing the Professional Teaching Service (Spanish acronym: LGSPD), the evaluation of teacher performance needs to be done in context by expert observers and evaluators. This sort of evaluation requires more than just standardized



instruments. Teachers are obviously entitled to be evaluated within certain contexts, and hence it essential that competent evaluators be used”.

With regard to curricula and study plans, it's very important, especially in the area of elementary education, that the INEE evaluate curricula and study plans to ensure that they have clear sequencing, are technically sound, satisfy the needs of the context and fulfill educational aims. Our practice with regard to national curricula and study plans has generally been to prescribe what must be done to the letter, without room for any fine-tuning, in all the different types of school.”

“Now we have a chance to rethink curriculum design and aim at a suitable balance between meeting national educational needs on the one hand and on the other hand giving schools the freedom they need, and affording teachers the role they deserve, in the decision-making process. Curricula and study plans are created so that teachers can know what to teach, but it's a mistake to overfill them with contents that smother the initiative of teachers and schools.

“The creation of curricula and study plans is the job of the federal education authorities, and, in their carrying out of this task, it is essential that the basic contents and learning aims that go to make up the said study programs be clear, since only thus will there be a match between the latter and the evaluations that it falls to the INEE to administer. It is important that the INEE, when evaluating the aforesaid curricula and study plans, ensure that their basic components are arranged in the proper order of importance and serve as a foundation for evaluating student learning and, insofar as is fitting, teacher performance and school administration”.

“With regard to educational policies and programs, it's not just a matter of curricula, but also of those programs that conjoin organized strategies and actions so as to achieve aims. The INEE will have to be selective, since it will only be able to carry out certain evaluations. Regardless of the specific programs that it opts to evaluate, it should not lose sight of the whole range of them if it is to ascertain how the different components that must come together in order to strengthen our schools interact.”

“In this regard, its functions are different from those of the National Evaluation Council (Spanish acronym: CONEVAL), whose oversight is limited to budgeting matters, though the two institutions may complement each other in some areas. The INEE is able to take a comprehensive, very productive, look at the whole education system, and it should exploit this advantage when designing its public-policy evaluations.

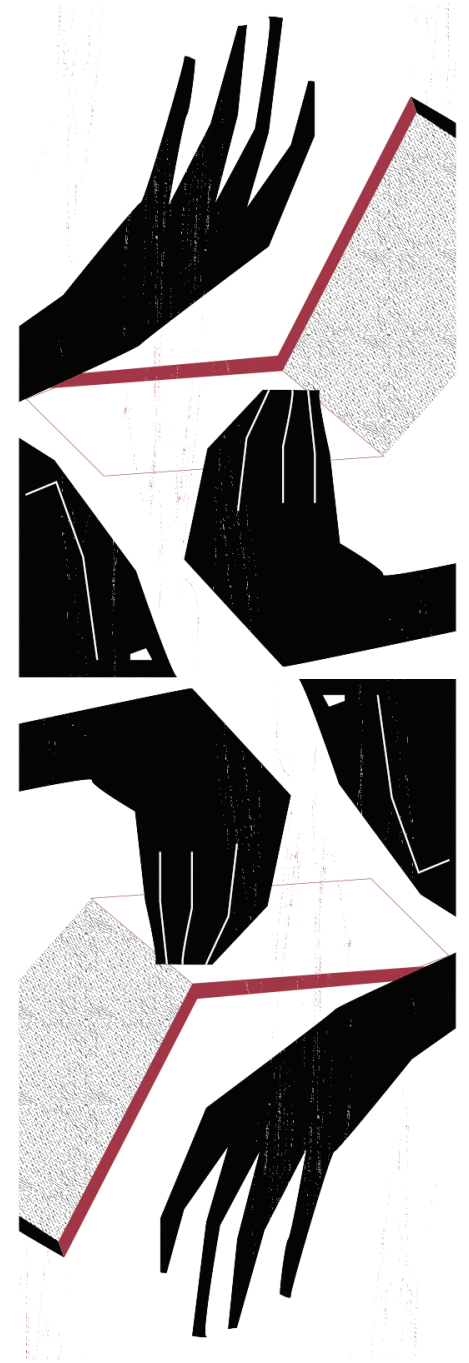
“Although the INEE is the unquestioned leader in the area of educational evaluation, its actions should be complemented by those of other bodies such as the CONEVAL. There is an enormous range of people and organizations who should play a role in the evaluation of public policy and educational programs.”

### **Educational reform: policy factors and the creation of an atmosphere that is conducive to dialogue**

“We should strive to set shared aims and foster the exchange of ideas about how to implement educational policy among all those involved. Discussion and the expression of different viewpoints are the way to achieve the high-quality education that vouchsafed in Article 3 of our Constitution. This doesn't happen when power games and interests that have nothing to do with education encroach on the fragile terrain of education, and, though it isn't up to the INEE per se to solve these problems, it can, nevertheless, help to create a climate fostering serious analysis to the extent that it engenders pertinent, reliable evaluations whose results are properly disseminated. However, since not only technical skills are needed to achieve such evaluations, but also communications skills, it's essential that the INEE listen to, and heed, the different comments that teachers, social groups, researchers and other people make about the evaluations. It's only natural that some people should harbor doubts about educational evaluation, especially with regard to teacher performance in very difficult socioeconomic contexts, but at the same time it's in such areas that evaluation can make a more valuable contribution to improving educational quality. Evaluation in difficult contexts constitutes an opportunity for the INEE to convince people about the importance and advantages of the said activity for improving education throughout our country.

The diffusion of data that results from it will provide everyone with a better basis for respectful, constructive dialogue that fosters the improvement of our education system.” €

Interview by María Cristina Tamariz and Óscar González Ramírez



Gazette of the National  
Educational  
Evaluation Policy  
in Mexico

