# 7th International Conference

## IBSE in Elementary Schools

## Science Learning Assessment: Trends and Challenges

# COMPENDIUM

November 14th & 15th, 2013

Centro Cultural Universitario Tlatelolco, Mexico City

SEP
SECRETARÍA DE EDUCACIÓN PÚBLICA
CONACYT
FUMEC
INNOVEC
Innovación en la Enseñanza de la Ciencia A.C.

Compendium

Innovation in Science Education
INNOVEC

San Francisco 1626-205,
Colonia del Valle. C.P. 03100, Mexico City.
Print in Mexico. 2014.
www.innovec.org.mx

# 7th International Conference

## on Inquiry-Based Science Education in Elementary School

## Science Learning Assessment: Trends and Challenges

# Index

# Introduction

In the last decade, the application of the Inquiry Based Science Education System (IBSE-SEVIC) in Mexico has been made possible by the collaborative work of the Mexican Ministry of Education, state governments and Innovation in Science Education (INNOVEC), a nonprofit organization. As a result of this joint effort, there have been many opportunities for teachers, educational authorities and specialists, to discuss within a national and international perspective, the formative value of Inquiry Based Science Education.

One outcome of the discussion has been a shared consensus on the need to analyze the new educational assessment trends and paradigms, the challenges they face and the procedures required to determine how they contribute to students' formative processes under the inquiry approach.

Inquiry Based Science Education Systems emphasize that students attending elementary and secondary schools should both understand natural phenomena and develop the scientific attitudes and skills needed to successfully perform in the 21st century. Therefore, these educational programs are considered a valuable tools for achieving a quality education that reinforces not only academics but also civic values in students, enabling them to make decisions that will help improve their individual and social well being.

Within this framework, it is important to design new assessment models that are able to measure the level of understanding of concepts as well as the development of the students' skills and attitudes. This trend has become a priority to both Mexican and International educational systems.

The 7th International Conference "Science Learning Assessment: Trends and Challenges" will be a forum for the discussion and analysis of experiences and new proposals for assessing, among other things, the level of understanding and creativity achieved by students; the development of their critical thinking; their ability to apply acquired knowledge in problem solving; their ability to work collaboratively, their skills for addressing scientific oriented questions, their ability to design and conduct research and to interpret evidence and draw a scientific conclusion. In other words, an assessment that encourages their academic development without detracting from their joy for learning or negatively impacting their self-esteem.

Therefore, those who work with IBSE-SEVIC programs are sure that the Seventh International Conference, presented here, will greatly contribute to the creation of better assessment processes aligned with quality education.

**Jaime Lomelín Guillén**
President of the Board of INNOVEC

# Opening Ceremony

*"INNOVEC believes that the evaluation of education should be conceived as a common and routine process in schools. As a part of society we are especially interested in the assessment of learning as a way to improve teaching"*

On behalf of the Board of Directors of Innovación en la Enseñanza de la Ciencia (INNOVEC, Innovation in Science Teaching), I am very pleased to welcome you most cordially to our 7th International Conference on Inquiry-Based Science Education. This year the title of the event is "Science Learning Assessment: Trends and Challenges".

Now, I wish to recognize the members of the Board of Governors of the National Institute for the Assessment of Education and its chairperson, Ms. Sylvia Schmelkes del Valle, for her enthusiasm and commitment to our initiative to discuss this important topic.

I also want to recognize our Minister of Education, Mr. Emilio Chuayffet, and the Undersecretary of Elementary Education, Ms. Alba Martínez Olivé, who very favorably responded to our request for institutional support from the SEP (Ministry of Education), not only to organize this Conference, but also with funding that will allow us to continue working on our program in several states of the country. Thank you indeed for your support.

For more than ten years the efforts of INNOVEC in Mexico and the world have been devoted to fulfilling children's legitimate right to quality scientific education; and assessment is key to INNOVEC, not only because it is always important to pause and assess the impact, results and scope of our work, but also because science itself should provide us with evidence of our progress or of our weaknesses.

Therefore, INNOVEC believes that the assessment of education should be conceived as a common and routine process in schools. As a part of society we are especially interested in the assessment of learning as a way to improve teaching through the work of teachers and by designing and improving educational systems. Businesses, for example, value proactive involvement of their human capital to identify faults, recognize mistakes and suggest solutions to make right whatever isn't working well or does not satisfy client needs. Measuring our performance and effectiveness is part of our daily activities in any part of the productive chain. As consumers this is why we prefer quality products and services that satisfy our needs and tastes.

In assessment we need other models, of course, because we are educating to provide conditions that are optimal for human development; conditions where knowledge in and of itself is held by our civilization as a precious value, and where science, technology, art and humanism come together to attract every man and woman because they are all entitled to a good education. INNOVEC educates to promote an education that will provide Mexicans with more opportunities for employment and development.

To accomplish this we need effective assessment processes that will account for what we've learned while recognizing the great diversity of circumstances, contexts and challenges we face as we teach. In a country like Mexico we cannot use the same bar to measure our educational diversity, but we must establish minimal standards to track our progress.

I wish a very pleasant stay in our country to our colleagues from abroad. With that, I encourage us all to have a very productive meeting.

Thank you very much!*

* Transcription

## Dr. Mario Molina
Vice President of the Board of INNOVEC

*"...we should learn to evaluate the results of our efforts not only to determine if they are working, but also to optimize them. We must learn what is working best in order to disseminate it among ourselves, and also to understand the problems that need to be tackled."*

Thank you for your kind welcome and the opportunity to welcome you all to this event. I want to especially extend this welcome to all of our guests who have traveled from afar and to all of the teachers, and distinguished members of the presidium.

We all know that education is hugely important to our societies. Practically every country in the planet has recognized this, as of course we have done so in Mexico. We understand, however, that it is not enough to say we want more education, nor does it suffice to allot education an appropriate budget to make it operational. Indeed, that is all very significant, but quality in education is even more important. For that reason in Mexico we have been working on our experiential and inquiry-based science education systems (SEVIC, is the acronym in Spanish). They are part of an international program supported by the Science Academies of many countries, and therefore afford a unique opportunity, a revolution in education itself, because these new pedagogies truly bring about extraordinary improvements in the quality of education.

Speaking of Science Academies, I would like to recognize Dr. Lee Yee Cheong who came all the way from Malaysia, and is very closely involved in this through the science education program of the Inter Academy Panel. This Panel carried out the original studies into how we learn, how children learn and how society can also learn more effectively. Dr. Cheong also chairs the International Science and Technology Innovation Center for South-South Cooperation, under the auspices of UNESCO.

It is of paramount importance that we benefit from these new pedagogies, and fully tap into the potential for excellence these new systems have to offer. I will now refer to what we are doing here in Mexico. Many of our visitors perhaps are not aware of this, but our country has just reformed its educational system. One of the salient characteristics of this reform is the introduction of assessments.

It is inspiring the enthusiasm and huge gratification of the many teachers involved in the SEVIC program, when they realize that their students are truly learning. They are greatly gratified to watch children's innate curiosity at work as their students develop reasoning skills to learn, and enjoy science. There can be no greater satisfaction for a teacher! I myself have been a teacher for many years, so I appreciate the extraordinary feeling of watching my students understand something. Naturally, teachers involved in these programs (as perhaps many of you are), have no quarrel with evaluations because we understand how important it is to assess not only the teachers, but the system as well.

Hence the importance of this meeting, because if we understand the enormous potential of these pedagogies, we should learn to evaluate the results of our efforts not only to determine if they are working, but also to optimize them. We must learn what is working best in order to disseminate it among ourselves, and also to understand the problems that need to be tackled. One of these difficulties for many countries, for example, consists in providing science training to the teachers.

I consider, therefore, such assessment and evaluation systems extremely important, because conventional evaluations focus on conventional teaching which basically consists in memorization. The teacher merely recites concepts that the students must memorize, whether or not they understand them, whether or not they are bored by them, and how much students memorize is easily measured. The challenge lies in measuring student's understanding. This challenge is not just vastly important to elementary education, because new pedagogical systems are beginning to significantly influence other levels of education, all the way up to the universities, where it might be easier to measure how well students are learning.

There are now a whole series of experiments that have shown extraordinary results, including actively participating students, children experimenting and writing up their own reports, and discussing them with their classmates; children acquiring values and understanding what really matters to our society. We must be able to evaluate and measure all of this. Hence the major significance of this event you are all participating in.

With that I will conclude my remarks wishing you the greatest success in the work you do here. We hope to document the steps necessary to evaluate, measure and use the results from these assessments to further optimize these extraordinary science-teaching methods, which I believe are not just for science but will also have great repercussions on teaching in general.

Thank you very much for your attention.*

* Transcription

## José Narro Robles

Rector of the National Autonomous University of Mexico

*"...in our country we most certainly have to work to transform our current educational process. Such a transformation should be comprehensive."*

A very good morning to each and every one of you. I am very pleased to share the inauguration ceremony of the 7th International Conference on Inquiry-Based Science Education devoted to assessing the learning of science, its trends and challenges with such a presiding group of personalities and with all of you. I am also very happy to welcome the Undersecretary for Planning at the Ministry of Public Education. Kindly convey our regards to the Secretary, Mr. Emilio Chuayfett.

How fortunate to find ourselves here with somebody who fills us Mexicans with pride, and that is Dr. Mario Molina. It is very significant that a world-wide recognized scientist of his stature should devote his time, intelligence, skills and will to a matter such as the one that brings us here today. Thank you very much, Dr. Molina. I also want to thank the chairman of the board of INNOVEC, Mr. Jaime Lomelín, the representative of the Director General of CONACYT, the chairman of the National Council for the Assessment of Education, the Director General of the SEP, and engineer Fernández de la Garza. Thank you all for joining us, and our special appreciation to those who have traveled to be here. Thanks to all the teachers with us today.

Indeed, at the end of any educational process we require that permanent pairing that has accompanied the development of educational institutions throughout history: the students and their teachers. Happily today we see among us teachers interested and motivated by the topic at hand: science learning assessment.

In our so-called knowledge society we must not only convey knowledge, but also use the processes that generate it, to apply it and bring about both technological and societal innovation in every field. This is relevant here, because in our country we most certainly have to work to transform our current educational process. Such a transformation should be comprehensive.

I don't think that anybody in their right mind in this country or any other can believe that everything has been done and that it has all been done well. The educational process itself teaches us that we must promote change, because if we remain where we are, we'll fall behind. If we don't advance at the right speed, we will lag behind. Now we in Mexico need to change. We must further develop our incipient Reform, because nobody can believe that our Reform ends with a few constitutional changes and secondary legislation. Of course not! We need to review everything concerning the process. We have to review how teachers are trained and kept up-to-date, our methods, and the conditions under which education is given and developed. We need to find out if we have the space we require, if our programs and curricula are flexible enough. We need to embark upon a full examination of public education in our country and in fact, rethink it and come up with a new version for the 21st century.

I am fully convinced that our public education system has greatly contributed to the development of Mexico. Our country would be very different and certainly not better if Mexico hadn't had those systems, those facilities, that infrastructure and those public education curricula in elementary school, middle school and, of course, higher education schools. Nevertheless, together with our analyses and consideration for our progress, we must recognize that today we are no longer producing the results we want, nor are we truly insuring the right conditions for a change in the years to come. We need to revise this.

Today, in the fields of science teaching and learning we have a very good opportunity to do so. I say opportunity because in order to solve the problem we must first recognize it exists, and once we recognize that, we need to know what the problem consists in, where we fail, what needs to be corrected, how we can improve or solve these deficiencies. This is why we should commend INNOVEC, its board, president and members, and the Public Education Ministry authorities for supporting this International Conference. If we are to change we need to start very soon. If we are to improve in this field we must begin with elementary education. As Doctor Mario Molina mentioned we must make use of children's interest and conviction, their joy, creativity, freshness and spontaneity by providing them with facilities and stimuli to insure that they can develop the skills they are endowed with. I undoubtedly believe that if we do all this, we will improve very substantially.

Any evaluation we review, whether from the World Economic Forum, OECD data, PISA tests, and our own analyses show we are not doing well, and that we have a huge opportunity to improve. Any problem should be viewed with double vision: the difficulty at hand and the huge potential to solve it. When things are the poorest, we have the greatest opportunities to make progress very rapidly. I am not going to provide data here, all I will say is that we have great potential to improve.

I want to invite the entire audience, every participant to pursue that endeavor. I am absolutely certain that if we work together in an articulated manner, coordinating our efforts and wills we can make progress. I trust that the work done during this 7th International Conference will help our children, our youths, and our institutions.

You all know that the National Autonomous University of Mexico is a very large institution with more than 335-thousand students, including 12 -13 year old teenagers incorporated through our 6-year junior high and high school system and our doctorate students. The challenge is huge. We have undergone a self-critical analysis and recognized that we have not been efficient in our teaching of science and that we have to transform our programs, infrastructure, laboratories, and make use of new technologies.

I will now conclude by inaugurating this 7th International Conference wishing you the best success in your deliberations and assessment, so that we can learn from such an evaluation and recognize successes and failures in order to make proposals to correct and move forward. The Mexican children and young people will thank you for it. Thank you.*

* Transcription

# Keynote Conference

## Sylvia Schmelkes

President of the Board of Governors of the National Institute for the Assessment of Education (INEE)

There are three problems in the educational system in Mexico: the first one is coverage, the second one is inequity and the third one is quality. We know those problems because we have evaluated the educational system for several years.But that it is not enough, evaluation itself do not solve the problems. An evaluation can truly be used to solve problems but requires educational policy mediation. In other words evaluations serve to indicate what education policies need to do to improve. To make the right desitions we need to build an evaluation culture.

# DEVELOPING AN ASSESSMENT CULTURE WHAT FOR?

Very good morning to everyone. I want to cordially thank all of the institutions organizing this event for the honor of having been invited to open it with this presentation. I will not be discussing the evaluation of science teaching, nor will I speak about the National Institute for the Evaluation of Education (INEE). I will attempt to provide a context for your discussions during the next two days and that is the need for a culture of assessment and evaluation, and the way to build one.

### THREE PROBLEMS

First of all I'd like to point out that when we speak about the need to improve the quality and equity of education in our country, we must necessarily begin by recognizing that we have problems, in our case of three kinds. I will highlight them, as I believe they are the most important ones in our national education system. The first problem is coverage, which we hardly recognize because international indicators assigned to our elementary school coverage only look at school children between six and eleven years of age. Mexico has fewer problems there, because our coverage of children between six and eleven is quite high, around 97-98%. This keeps us from realizing that we still have problems in our elementary school coverage. In fact, there are 3.9 million children between 3 and 14 who are out of school. Most of them are pre-schoolers who should be in first grade of pre-school, in other words, 3 year-olds. They do not go to school because none is available; for the most part first grade pre-school does not exist in the poorest and most remote places, inhabited by indigenous peoples. In these areas there are 1.7 million children who do not attend first grade preschool despite the fact that it is mandatory since 2008. Even though most unschooled children belong to this age group, there are still 407-thousand children between 6 and 11, and 548-thousand children between 12 and 14 who do not go to school. These figures add up to what I'd say is an alarming total for our country: 3.9 million children who should be attending mandatory elementary school and are not.

When INEE began to analyze why these children don't go to school, we realized that the problem is not always lack of availability. Except for first grade pre-school, schools do exist where these children live, and supposedly, they have teachers as well. Therefore the phenomenon has to do with external factors, and desertion is an important one of them; children who once went to school but dropped out for various reasons, including poverty and the need to work, or for disability reasons. This means that many children with some kind of disability do not have access to school. Also, a significant number of children, especially secondary school students, find school meaningless and that it does not provide them with something useful for their lives now or in the future.

Inequity is a second problem. In my opinion it is one of the most serious problems because inequity in the educational system means that education is not fulfilling the very important role of generating equal opportunities for a quality life. In Mexico we face significant inequity issues that reflect the inequity in the country, which as you know, is one of the greatest in the world. This can be seen in education, and is linked to whether children live in an urban or rural area, to the degree of marginalization in the region where they live, to whether or not they speak an indigenous language and, obviously to their household income level. This is noticeable in how much schooling they get, and unfortunately is also apparent in actual learning as measured by current methods, which consist in knowledge tests.

This inequity can be partially explained by the fact that Mexico has yet been unable to offer the same educational quality to all of its children and youths. We are not investing the same amount in every child, much less are we making up for social-economic differences. I don't like to use other countries as examples, but I find what I am about to say particularly illustrative. In other countries, schools are allotted a budget equivalent to the average cost of a child in a given grade, multiplied by the enrollment. This insures an equal amount of resources for every child in the country. That is not equity but equality, which makes equity possible. Additionally, in many of these countries, extra money is provided to schools recognized as in poverty or where students speak a language other than the official language, or where they accept children with special educational needs or children with disabilities. This is because those countries recognize that such conditions make it more difficult to attain desired results, and therefore require additional resources. In Mexico we're not doing this; quite the contrary, we give less to those that need the most. Let me quote some old figures from the past administration. I heard this from the former director of the National Council for the Promotion of Education (CONAFE). He said that under community courses, a modality we use to serve children in scattered rural communities, the annual cost per child is 7 thousand pesos, whereas the average annual cost per child in an elementary school is 37 thousand pesos; that is 5 times more. From a different perspective, the cost of teaching a CONAFE student is 5 times less. This should give us an idea of how instead of first insuring equality and later insuring equity, we are doing the opposite. This explains why the problem is so serious. Another reason why we have inequity is that in Mexico we follow a homogeneous model. Our curricula are identical for the entire country, and the organization of schools, at least theoretically, is also identical for every one of them. This model weakens as it reaches the remotest regions. For example, instead of having six teachers, one for every grade, there are classes that include multiple grades, or even unitary schools in which case the principal who is in charge of managing the school must also take on teaching duties, because very tiny communities are not thought to merit having more teachers. Thus, when the principal has to leave the school to carry out administrative formalities, the children miss class. That is an example of a theoretically identical model that becomes weaker as it reaches these communities.

And yet we have this homogeneous model despite the great diversity in our country: 10 % of our population (depending on how it is measured, but 10 is the average number) belong to an indigenous group. 7-million speak an indigenous language, but 15-million describe themselves as indigenous from one of 68 perfectly differentiated ethnical and linguistic groups. This is in addition to other cultural and geographical differences in our country. North and South are different, while living in a mountainous area is not the same as living on the coast. Despite all of this diversity we deal with differences in exactly the same way, so some benefit from the model, while others suffer for it because they do not fit in with the characteristics of the model, which explains inequity. Regionally speaking we also find a strong correlation between the gross internal product of the states and the level of schooling completed by their populations. There is a strong relationship between cultural, social-economic conditions and achievements in schooling and learning. This is the second problem I wanted to talk about.

A third problem is quality, which tends to get more attention. We overlook providing the same access, and concern ourselves relatively little over equity, but quality is discussed much more frequently. Somebody already mentioned we are unhappy with what our children are learning. We have PISA data that are based upon a conception of what students need (15-year old students specifically), to respond to the demands of modern society in the short term. We have these in addition to the results of our own tests. In both cases, we have very high percentages of students whose outcomes are below the baseline, or below the baseline defined as necessary to face the demands of modern life. Even in secondary school mathematics, the figure is 52% in our own tests, and a very similar one in PISA data. Obviously in the case of sciences, data do not differ significantly. Another thing to be observed in this process is that in general terms when we look at the historical evolution of our students' performance in different kinds of tests, we see a certain improvement in elementary school, but in secondary school reality becomes more static making it more difficult to attain important increases in secondary education.

Looking beyond what affects school children it's important to consider that we have a population over 15 termed "adult" population that lacks elementary education. Indeed, a third of our national population does not have elementary school education, and that third of the population represents half of the individuals 15 years old or more who haven't finished elementary schooling. In other words, this constitutes another important challenge we tend to forget because we assign very few resources to adult education, and this challenge needs to be recognized. Thus, in secondary education our inefficiency amounts to 21%. That is, for every 100 children who start school, 79 finish three years later and 21 are left out.

*"Inequity in the educational system means that education is not fulfilling the very important role of generating equal opportunities for a quality life. In Mexico we face significant inequity issues that reflect the inequity in the country."*


## CAN ASSESSMENT SOLVE THE PROBLEMS?

I think it is very important to clarify that assessment on it's own will not solve problems. In fact, during the past decade Mexico has been using a very intense assessment. It has assessed students and teachers over and over again, and yet this has not been used to improve education, but rather for accountability. It has been used to reward teachers, which is the same as punishing them, because teachers who are not rewarded are necessarily punished. Thus, the fundamental purpose of this test has not been improvement.

Assessment and evaluation do not solve problems, they provide a dimension of the problem while educational investigation explains it. So assessment and evaluation can truly be used to solve problems, but it requires educational policy mediation. In other words evaluations serve to indicate what education policies need to do to improve. Assessment and evaluations alone cannot solve problems and the mediations they call for refer to improving working conditions. A teacher cannot be assessed and be successfull if he is working environment does not provide the minimal conditions to teach.

He cannot be expected to pass a further test or evaluation with no improvement if his working conditions. That would be unfair. It is necessary to modify and improve working conditions, and to consider the context. A fair assessment will consider the context recognizing that it poses problems that affect schooling and learning. I find it important for us to understand that the issues at the root of educational problems call for inter-sector intervention, and that they have to do with context: children having to work, who do not have proper nutrition, children who have some kind of disability and can't start school. These are the conditions we have to modify.

A fundamental mediation, which is perhaps the most important one of all lies in teacher training, especially training in service, although early training continues to be important. The kinds of mediations necessary for truly better curricular quality and education equity include the manner in which resources are distributed, modifications to curricula and educational materials, a homogeneous vision of the programs, and the need to make them flexible. Of course, programs and policies must be designed according to identified causes of the problems and ways they will be tackled. This in turn will lead to determining what requires innovation; i.e. what we need to do differently in order to get different results. So, assessment leads to knowing what has to be innovated at schools, in classrooms and, obviously in educational systems.

## FAIR AND RELIABLE ASSESSMENT

Now, not all assessment can support decisions about educational improvement. Assessment that can really become instruments for improvement should have certain characteristics: they must be essentially formative in their proposal. This means a purpose of the assessment should be improvement, an intention, an explicit will behind every assessment. It should focus on known problems. For many years now, we have been involved in education and evaluation research. We know what the problems are, so we must focus: on evaluating access, quality and equity in education. In order to obtain evidence of the causes of these problems we must evaluate the main components of the education system that are known to impact the problem and the basic players. Next, we must assess the students, teachers, institutions, programs and policies. We evidently need to promote the development of education and evaluation research that will allow us to look deep into the causes we must tackle, because assessments and evaluations alone are not enough.

Additionally in order to further improvement, assessment and evaluation must be credible and therefore, as technically sound as possible, by this I mean that in technical terms we are not yet fully developed either. We are improving upon assessment and evaluation techniques that will make it increasingly possible to actually measure what we want to measure. However, we must strive to do so as reliably as possible with available techniques and methodologies.

It is also necessary to begin with the complexity of the act of education. Evaluating education is complex because education is complex and cannot be simplified nor analyzed with an assessment that simplifies for example, by means of a single tool. We know that complex facts and problems are multi-factorial, so we must consider all those factors that have a bearing on the outcome: context, cultural diversity, are highly relevant data.

For an evaluation or assessment process to be credible, it must be fair; and if it is to be fair it must insure it is neither awarding nor punishing without evidence. We must be certain this can be communicated, that because our decisions are properly supported we can indeed convince. Diversity must be recognized and considered so that an evaluation will let everyone improve. Assessments cannot become just another homogenization tool, but should start from recognizing that our reality is diverse and requires differentiated processes in order to deal with problems. They should also recognize our inter-cultural relationships and enhance them. Those of us working in intercultural education must make a distinction between inequality and diversity. The former is something to be tackled, while the latter is something to be enhanced and promoted because it enriches us all. From that perspective inter-cultural relationships should be enhanced to avoid the natural tendency of any evaluation to homogenize. It is necessary to prove our ability to promote decisions, programs and policies that actually offer to improve a known educational reality.

## RISKS

Also important to recognize is that assessments have their limits. These natural limits are imposed by the development of theory, methodology and assessment techniques. Indeed, at this time not everything can be assessed. We'd like to evaluate everything and oftentimes we can come close to evaluating as much as possible, but must recognize beforehand that there are things we are not certain we can evaluate; values, for example. Our methodologies to assess values are still weak, so there are limits and it's important for us to recognize them. It is also important to consider that assessments pose risks, and an evaluation culture (which I'm about to discuss) entails recognizing and consciously avoiding risks such as the inappropriate consequences leading to perverse outcomes that assessments can sometimes bring. Here I'd like to refer to how the ENLACE test was used. In the early stages of its design, the test was intended to assess our national education system and identify where there were problems in order to provide feedback to schools, teachers, students and their parents. All of a sudden someone thought it should be used to assess teachers, and this inappropriate use lead to the perverse effects known by all: tests were sold, fraud, teaching for the test, children permanently doing exercises from the ENLACE tests, and so on. We forgot what truly mattered. For example, since multiple-choice tests do not measure writing, we forgot to teach it. These perverse effects are precisely due to evaluations, and therefore constitute a risk we must conscientiously seek to avoid.

Measuring what can't be measured will necessarily lead to discretional decisions. If we attempt to use an instrument to measure something that technically cannot be measured, then our decisions will be subjective, discretional and hence become a risk to be avoided.

The risk of reducing education to what can be assessed; that is, assigning importance only to what can be measured would be a very dangerous. Assessment should not become an informal or hidden curriculum of any education system. This would be terrible because teaching only what can be measured, and as I have already mentioned, homogenize for the sake of simplification would lead to a banal interpretation of the complexity of the universe, and this would be a terrible mistake.

*"Assessment and evaluations do not solve problems, they provide a dimension of the problem while educational investigation explains it. So these processes can truly be used to solve problems, but they require educational policy mediation."*

## THE CULTURE OF ASSESSMENT

Now I will turn to the topic I was asked to address today: the culture of assessment. I will begin by defining our understanding of culture in this context. It is not an anthropological definition, but understands culture as a socially shared way to envision and understand a reality in order to judge and act in consequence. This refers to culture in general, not an assessment culture, which, in my own terms, is our socially sharing the fact that assessment must be technically sound and fair, for only sound and fair assessments allow individuals and groups alike, to make better decisions and understand the problems at hand, what needs to be solved, or what practices need to be changed or improved. That is an assessment culture. It must be shared socially, otherwise it is not a culture.

Therefore we must share the conviction that an evaluation must be technically sound and fair.

On the other hand, an assessment culture is also understood as one that recognizes that a technically sound and fair assessment constitutes a tool to insure everyone's right to quality education, which consists in learning whatever is necessary to lead a dignified life. When I say everyone I am including the concept of equity. Here I want to quickly refer to the **4A's** that assess the right to education: **A**vailability, which means that there is a school with a teacher, and that availability is insured; **A**ccessibility; this means that once there is availability we must not impose barriers upon demand and the population who want access. Of course there are physical barriers for children with motor disabilities, but there are other barriers like financial barriers. These come in the form of fees, or mandatory uniforms, or excessive amounts of school supplies that end up imposing financial barriers to school.

The other two A's refer to quality. **A**daptability goes against curriculum homogeneity and common modes of imposed school organization. It fundamentally has to do with an education that is meaningful, relevant and pertinent to different population groups in their current and future lives. The final **A** is **A**cceptability, which defines quality from the students' perspective: their liking school because they know they are learning and feel respected, welcome and safe. Schools should be acceptable to learners otherwise they cannot be as well disposed to learn.

Thus, we should hold this great framework of the right to an education as something we truly want to improve. We must insure it is fully complied with using assessment and evaluation as a tool to accomplish this.

An assessment culture demands certain things. Assessment must be necessarily public, otherwise how are we to generate the shared awareness that assessments actually work? Making them public means making known the purposes of our assessments, why we design assessments, and respect for private information when individual outcomes are not called for. This affords a great deal of certainty. Personal information is not released when results are made known, and such personal data are added only when necessary such as in the case of a university admission examination. In this case such information is released to the individual only. Not all assessments make it necessary to release data on individuals, so privacy must be respected.

Assessments must also be transparent. This means that they can be dissected to find out what was done, and determine the process. Finally, assessments must be subject to appeal. Evaluators should not have the final word. People who are assessed have a right to know their results, challenge them if need be and request verification. In other words, a true assessment culture requires this from an evaluator so an assessment can be considered fair and its virtues shared.

An assessment culture must be built. It cannot be generated by decree. It's impossible to change a culture by decree. So an assessment culture must be built and strengthened as assessment effectively proves its ability to bring about improvement. Indeed, that is the condition that will allow us to effectively build an assessment culture.

A balanced assessment culture implies respecting reliable assessments and rejecting those that are partial or unfair. Such a culture also requires assessments research to learn the causes of the problems discovered and described by assessments. Which calls for research to complement them. A constructive assessment culture involves recognizing that assessments in themselves do not bring about improvements but lead to demanding necessary mediation. That is part of an assessment culture: the notion that assessments are not a magic wand to improve quality, once again, because they alone will not improve it.

## KINDS OF ASSESSMENTS

As you know we have two kinds of assessment: summative assessments that measure individuals at the end of certain processes, and formative assessments that provide feedback and improve the quality of those processes. Summative assessments are necessary because they are used to make significant decisions, so we may build upon the credibility at the foundation of an evaluation culture. Summative assessment designers should guarantee the highest degree of validity and reliability that they are using every tool necessary to calibrate for multi-factorial problems or facts. The users of the outcomes, for example decision-making education authorities, need to accept diversity and transparently consider the conditions of the context. They are also supposed to responsibly decide the formative use of these assessments for second and third opportunities. Also important is that assessments should not be merely viewed as cold figures that allow us to make a decision in any context because contexts must also be considered as a process.

Assessments should allow for second and third opportunities while providing feedback so outcomes in those second and third opportunities can actually be better. Subjects of summative assessments are supposed to accept that the best decisions resulting in consequences for them are based upon objective and fair assessments. This means the outcomes (in final examinations, admissions exams, teachers exams) must also be accepted responsibly. We must come to accept that a technically sound and fair assessment is the best way to make those decisions.

Formative assessments, on the other hand, focus on processes and are intended to give feedback to improve. Designers of these assessments should insure that they are endowed with the greatest potential to give feedback that will bring about improvement. The recipients of these results (once again, education authorities, classroom teachers, and others), from external assessment must have the assurance that these assessments are guaranteeing the mediations necessary for improvement, because the assessments will not provide for improvement. Subjects of assessments should accept any support derived from formative assessments and take any measures necessary to commit to subsequent changes.

## WHAT ASSESSMENT MEANS FOR THE DIFFERENT PLAYERS

So what does this mean for every one of the players in the education process? For students it necessarily means that they understand the purposes of the education process in question; in other words what they want to do with their education, what needs to be assessed. They must also accept the decisions derived from valid and fair assessments. Because these assessments are formative, students must be willing to follow the recommendations derived from their outcomes so that they are indeed used for improvement. That is part of the culture in which one expects students to say, "I'm glad! This assessment is giving me feedback and telling me what I need to do to improve." When teachers evaluate their students they are expected to clarify the purposes of their teaching, design assessments that give feedback, not pop quizzes for mere scoring. If teachers are to assess, it will be for feedback purposes, and perhaps more importantly to accept the feedback from their students results, because it will tell teachers whether or not they are being effective, whether they are achieving what they set out to accomplish, whether they are being effective with all students, and if not that they need to correct what they are doing in order to improve it.

When teacher assessments are fair and valid, they must be recognized as a means towards professionalization. Teachers should accept what assessments entail with regards to training and preparing for innovation. The idea is for teachers to change the way they do things in order to get better results.

Schools or educational institutions are expected to use the outcomes from assessments of the students, teachers and of the institutions themselves to identify problems, plan a collegiate solution to the most important ones, and commit to make any necessary changes, monitor and evaluate progress and, thereby begin a new cycle of improvement.

It is very important that as a result of assessments, educational systems establish appropriate mediations to deal with the problems discovered and estimated through assessment, especially those I consider to be crucial: access, inequity, lack of quality in education. And, how is this to be accomplished? Well, by means of support and accompaniment processes, by training teachers, teaching teams, and principals, by modifying their working conditions, as well as the conditions that affect schooling in specific contexts.
There are also implications for society at large, because societies must become interested in knowing the results of assessments, and these assessments need to become tools to strengthen our democracy. Therefore, society must put pressure on authorities to generate mediations and correct the causes of

any deficiencies detected. Naturally, society must do its share to contribute to improvement processes.

To close, I want to point out that I believe a balanced assessment culture understands that assessments are not infallible. They can make mistakes and should be revised. I think this is absolutely essential because at times we tend to turn assessments into fetishes granting full credibility to mere tools. Assessments are not infallible. They can be mistaken and can be perfected. They will never give us ultimate answers, so we must continue to advance our conceptual, methodological and technical ability to assess what we want to assess. This of course requires evaluating assessments themselves for their effects, impact and ability to predict and provide solutions to known problems, as well as those identified through assessments and, obviously make all of this public.

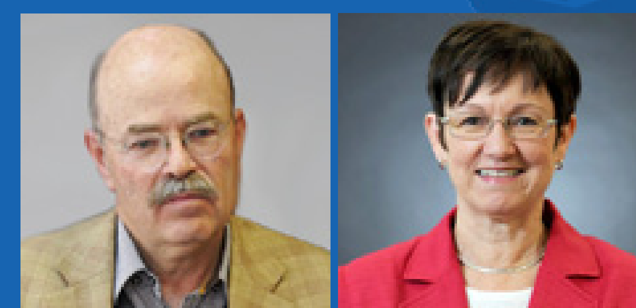
Thank you very much.*


* Transcription



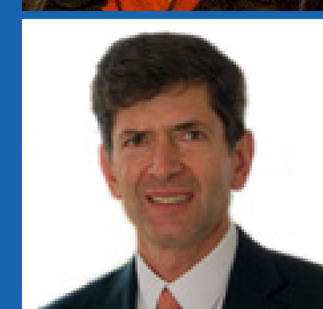
# Panel 1 Discussion

## HOW IS SCIENCE LEARNING BEING ASSESSED IN BOTH NATIONAL AND INTERNATIONAL CONTEXTS



**Speakers**
- Eduardo Backhoff Escudero
- Shelley Peers

**Panelists:**
- Alejandra González
- Ulrika Johansson
- Louise Hayward



**Moderator:**
- Carlos Mancera

# Cognitive models in science education assessment

Eduardo Backhoff Escudero

This paper is intended to justify the use of cognitive models to assess student learning, especially science learning. It will describe the importance of assessing learning both inside and outside the classroom, exemplify the use of cognitive models to assess sciences, and characterize science tests made by the National Institute for the Assessment of Education (INEE) and the PISA test of the OECD. Finally I wish to conclude on the need to move towards computer-managed tests supported by cognitive models.

Assessments of school achievement provide useful information so that teachers, principals, students and parents can make decisions that will improve learning. The intended use of an assessment will determine the components required for each stage in its design, preparation, and interpretation.

## KINDS AND PURPOSES OF ASSESSMENT

Inside a classroom, good teachers use different methods to assess their students; for example: tests, student observations, written homework, conversations with students, and other instruments to understand what they have learned. Since the purpose of this kind of assessment is to help pupils learn, it is called formative assessment.

It has been well documented that students learn more when they receive feedback on the particulars of their schoolwork, which is one of the important premises. However there are other kinds of evaluations. Summative assessments in classrooms also serve to make educational decisions. Basically they allow teachers to determine if a student has attained a certain level of competence after having completed a stage in his education. These assessments are also known as achievement tests. Some of the better known forms of summative assessments used by teachers in their courses include year-end examinations, institutional examinations or mid-term exams.

There are other assessments made outside the schools and classrooms. They are usually large scale in their use and application and are given by external staff. They also provide institutions relevant and comparative data on student achievement, which classroom-exams are unable to do. Third-party exams are far between and results are provided at later dates. They rarely provide timely information to teachers and students that can be used to make classroom decisions. They do, however, provide valuable information to institutions and national education systems.

As you may have seen recently in the press, decision-makers are beginning to consider large-scale education achievement tests a powerful tool to change what goes on inside classrooms and schools. In fact, all over the world evaluations are now seen not only as a means to measure performance but to change it. There are examples of this in Mexico such as the ENLACE test (National Assessment of Academic Achievement in Schools) and PISA test (Programme for International Student Asssessment), among others.

Even when the criteria of an assessment are met, care should be taken to avoid generalizing results and arriving to conclusions that are not supported by assessments. For example, a teacher whose students score high on a test, is not necessarily better than one whose pupils get lower scores. The same thing can be said about schools, the quality of inputs such as student backgrounds and available educational resources, all of these should be considered when interpreting assessments or their results.

## VALIDATION OF ASSESSMENT

A crucial feature of an assessment is its validity. Assessments of learning must be valid, reliable and equitable if they are to be useful. However we have yet to reach a universal agreement on what validation is, and that is fundamental to assessments. I will now share two of the most widely used interpretations or definitions. The first one comes from the United States and says that for something to be valid it must be supported by empirical evidence, and by the theory that backs up the interpretations of the assessment results. In other words, the interpretations and uses of assessment are emphasized. There is a very well known European school from Holland that says that a test is valid to measure an attribute - for example, a competence, literacy or knowledge – if and only if that attribute exists to begin with, and the variations in measurements are causally produced by variations in attributes. This means that if an attribute exists and it is reflected in the assessments, it is a sign that the tests are valid. Validity is a core issue that is often overlooked, but it is very important and will be highlighted throughout this paper.

## ASSESSMENT AS EVIDENCE

Once again, an assessment is a tool designed to observe student behavior and generate useful information in order to conclude what they know and what they can do. An assessment tells us what kind of evidence is available to explain the competencies of the individuals examined. Our beliefs about the nature of learning will impact the kinds of evaluation data pursued and the inferences that can be made.

Peregrino et al. provide a very important model, which I find fundamental in this proposal. He says there are three elements he terms the assessment triangle. I've included this as a model to make it clearer for you that there should first be a cognition model that I'll define later, then an observation model and finally an interpretation model. The cognition model encompasses the other two elements, while the observation model encompasses the interpretation model.

A student cognition model should contain two levels of specificity: a general model on how learning occurs and another to explain learning in a specific domain; for example, understanding fractions. An observation model should seek and be based upon the beliefs and assumptions on the kinds of evidence of student competencies an evaluation should provide. In other words, an observation model must be in line with the cognition model, while the interpretation model should serve to find meaning in the information provided by the assessment, and appropriately interpret it within the context of the model we have previously specified.

The cognition component, that is, the intellectual part in the design of an assessment refers to a theory or group of beliefs about how students represent their knowledge. Assessment will become more effective if its designer begins with a very explicit

*"An assessment tells us what kind of evidence is available to explain the competencies of the individuals examined. Our beliefs about the nature of learning will impact the kinds of evaluation data pursued and the inferences that can be made."*

and clearly conceptualized cognitive model of learning– something that rarely happens in assessments. Such a model should reflect the most plausible scientific explanation of the way students represent knowledge and become experts in a given domain.

I find the following example very much represents what I'm saying. In the 1970's Siegler examined how people develop an understanding of the components underlying the principle he called "torque", which I'm about to explain. He showed children of different ages the kind of balance in the figure below. It includes a rotating support called a "fulcrum", which allows the balance to tilt

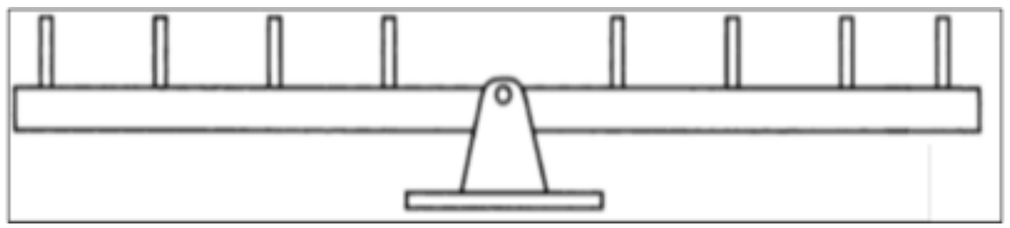
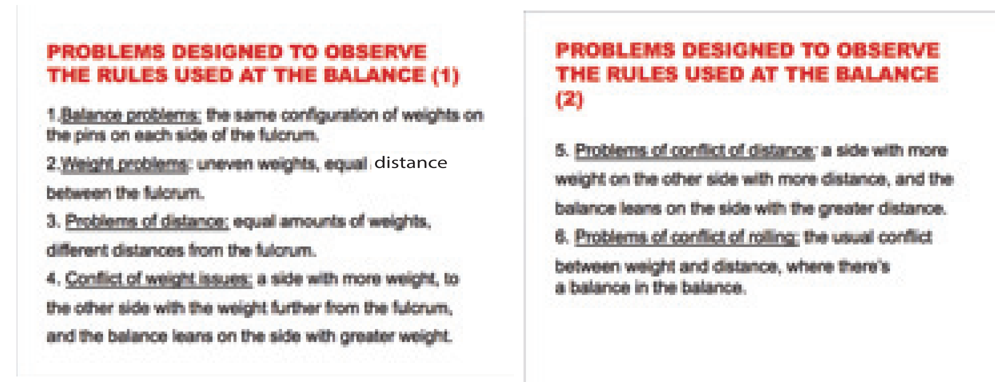


FIG.1
Backhoff Escudero, E. PPT. November, 2013.
http://innovec.org.mx/home/images/PresentacionesVIIConferencia/backhoff.pdf

to one or the other side. This experiment says that the balance can tilt to the left or right or remain level according to how the weights are arranged in the pins. The task of the students was to predict if the arm of the balance would tilt to one side or another. In this case the variables that influence the results are very important: the amount of weight on each side, and the distance from the fulcrum. Solving these problems requires knowing how to proportionally combine weight and distance. This is a widely known physics experiment.

In his investigation Siegler found four very important rules. The first one says: if the weight is the same on both sides, the forecast is that the balance will be level. If the weight is different, the forecast is the balance will tilt on the heaviest side. This rule is practically common sense and many students use it.

The second rule says that if one side carries more weight, the prediction is that the balance will tilt on that side. If both sides have equal weights, the balance will tilt on the side where the weight is furthest from the fulcrum. This is another simple rule students use a lot. Finally there are two or three rules more. To save time I will tell you the next to last rule: If the weight and distance are equal, the prediction is that the balance will remain level. If one side has greater weight or distance, and both sides are equal in the other dimension, then the forecast is that the balance will tilt to the side where the unequal dimension is greatest. I'll leave it at that, because the important thing to note is that Siegler's experiment produced four well-defined rules that can explain students' answers in the experiment or their language proficiency. For that purpose he designed six problems. The problems were basically defined and designed to determine which of the four rules students will use.
Here they are:
This is the balance and these are the problems posed by Siegler:



PROBLEMS DESIGNED TO OBSERVE THE RULES USED AT THE BALANCE (1)

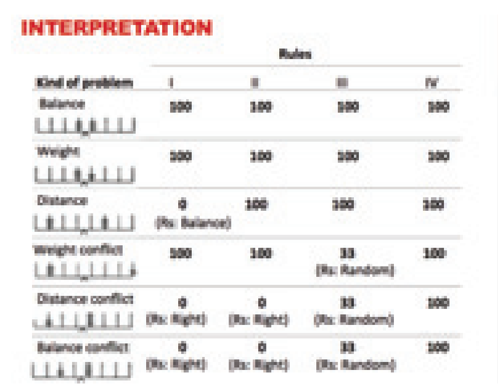
1. Balance problems: the same configuration of weights on the pins on each side of the fulcrum.
2. Weight problems: uneven weights, equal distance between the fulcrum.
3. Problems of distance: equal amounts of weights, different distances from the fulcrum.
4. Conflict of weight issues: a side with more weight, to the other side with the weight further from the fulcrum, and the balance leans on the side with greater weight.

PROBLEMS DESIGNED TO OBSERVE THE RULES USED AT THE BALANCE (2)

5. Problems of conflict of distance: a side with more weight on the other side with more distance, and the balance leans on the side with the greater distance.
6. Problems of conflict of rolling: the usual conflict between weight and distance, where there's a balance in the balance.

FIG. 2
Backhoff Escudero, E. PPT. Noviembre, 2013.
http://innovec.org.mx/home/images/PresentacionesVIIConferencia/backhoff.pdf

He observed that when students followed the first rule, they usually succeeded in all balance problems. They succeeded in the weight problem, but not in the distance problem. They succeeded in the weight conflict problem, but not in the distance and balance problem. The table shows that students who understood the relationship between weight and the distance from the fulcrum were able to solve all of the problems posed regardless of conflicts or different weights at different distances.



INTERPRETATION

| Kind of problem | Rules I | II | III | IV |
|---|---|---|---|---|
| Balance | 100 | 100 | 100 | 100 |
| Weight | 100 | 100 | 100 | 100 |
| Distance | 0 (Rs: Balance) | 100 | 100 | 100 |
| Weight conflict | 100 | 100 | 33 (Rs: Random) | 100 |
| Distance conflict | 0 (Rs: Right) | 0 (Rs: Right) | 33 (Rs: Random) | 100 |
| Balance conflict | 0 (Rs: Right) | 0 (Rs: Right) | 33 (Rs: Random) | 100 |

FIG. 3
Backhoff Escudero, E. PPT. November, 2013.
http://innovec.org.mx/home/images/PresentacionesVIIConferencia/backhoff.pdf

I will quickly show you an example. An item at the Excale test, by the

## COMMENTS

*A current demand that will require greater innovation is to reconcile large-scale student and teacher assessments with evaluations in the classroom. More importantly, these evaluations have different purposes and therefore different approaches. If a large-scale evaluation wants a snapshot of an education system it will necessarily be different to a class evaluation since the purpose of the latter is to understand what students have and have not learned.*

*In many cases evaluation, science development and teaching have followed very different paths. There are scant experiences (although the ones that do exist are highly interesting), in which scientific experts for example in biology, physics or chemistry, have at the same time advanced in the teaching and evaluation of scientific learning. Therefore, we lack well-developed theories as to how students learn that could be taken to classrooms in order to derive important information from any assessments made by the teacher to correct anything students' may have not learned in time.*

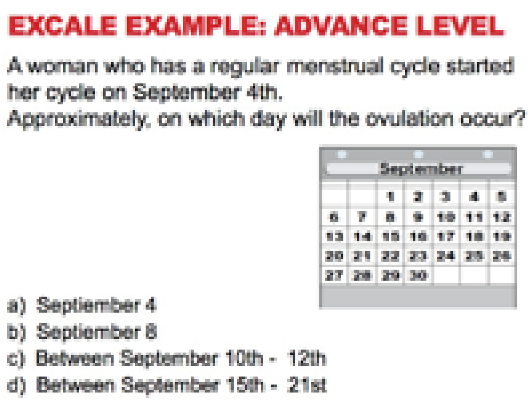


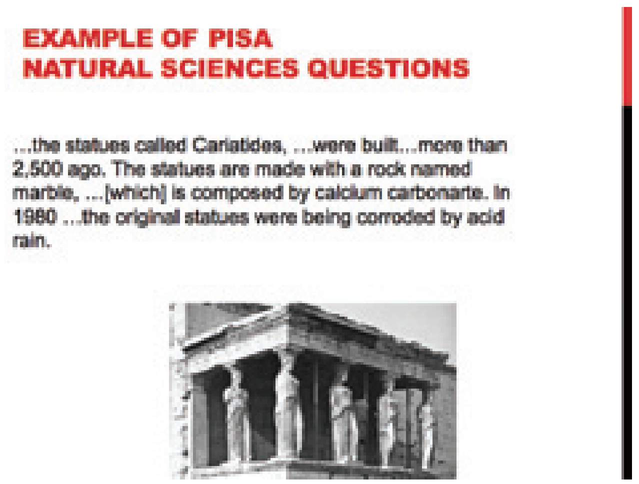

*FIG. 4*
*Backhoff Escudero, E., PPT. November, 2013. http://innovec.org.mx/home/images/PresentacionesVIIConferencia/backhoff.pdf*

National Institute for the Assessment of Education (INEE), says that if a woman with a regular menstrual cycle began menstruating around September 4, what day will she ovulate? The test provides a calendar and students must determine and select the approximate dates when ovulation will occur.

Another very different example is a sciences item that appears in PISA and reads:"The statues known as Caryatides were built more than 2,500 years ago. They were carved out of a kind of rock called marble, which is made of calcium carbonate. In 1980 the original statues were corroded by acid rain." This is the basis for the questions. One says: Normal rain is slightly acid because it has absorbed some carbon dioxide from the air. Acid rain is more acid than regular rain because it has also absorbed gases such as sulfur oxide and nitrogen oxide. Where do these nitrogen and sulfur oxides in the air come from? This question is much more complex than the one before. As you can see, it is not multiple choice, so students must write their answers and use their cognitive skills. These two items differ widely one from the other, and both are very different from Siegler's when he states early that, "learning occurs this way, a student may or may not assign, or have one of the four rules investigated previously, and therefore any student who has one of these rules and uses it will be able to solve such problems."

To conclude, the strengths of assessments rest on their adhesion to learning theories. Their limitations become apparent in that they are not able to capture the vastness and wealth of the competency being assessed. This is what evaluations are usually criticized for: they do not capture the important things. There is a concern as to whether currently used assessments capture today's emphasis on the complexity of learning: levels of learning, reasoning, comprehension, application or higher levels of application. Regrettably many evaluations do not focus on the cognitive aspects indicated by investigations. They are not designed to capture critical aspects of understanding in students' learning.

*FIG. 5*
*Backhoff Escudero, E. PPT. November, 2013. http://innovec.org.mx/home/images/PresentacionesVIIConferencia/backhoff.pdf*




Finally, despite the great differences between PISA assessments, and the assessments of the National Institute for the Assessment of Education (the former are not aligned to curricula, whereas the latter are) neither one of the examples I showed you uses explicit cognitive models such as the ones proposed by Siegler.

The use of these explicit cognitive models would require:

1. Adopting an effective cognitive model to assess each one of the scientific competencies of interest; and, 2. Move beyond the pencil and paper model to a computer-based model that would make it possible to evaluate cognitive skills in both small-scale and large-scale applications. Today we have what we call automatic test or item generators that make it possible to produce examinations based on cognitive models. This would represent considerable progress in science

assessment. This time, however, I will not go into explaining automatic item generators, but if you invite me next year, I'll be happy to. Thank you.*

* Transcription

## COMMENTS

*The Mexican Academy for Science has a program called Science at Your School (Ciencia en tu Escuela) devoted to the professional development of teachers. For ten years, our evaluation culture has not been easy nor fully understood by everyone as it has established itself with time (...) For example, inquiry-based learning requires discussion. Language forms thoughts, so it is important for us to discuss and also to learn to listen, observe and develop patience.*

*The program assesses materials and advisor performance. Advisors in turn assess teacher learning activities and provide a summative assessment through a project. (...) Another thing we evaluate are teachers' attitudes towards sciences and mathematics by means of a questionnaire used before and after interventions. The results have been surprising. For example, there are teachers who tell us they find it very difficult to approach mathematical problems or science challenges and therefore hold tight to their textbooks. At the end of the interventions, teachers tell us their perception has improved and that they feel better equipped to approach problems and devote more time to experiment.*

# How is science learning being assessed in both national and international contexts. Primary Connections case - Australia

SHELLEY PEERS

An investigation made in 2012 on the results of the Primary Connections Program for the Teaching of Science in Australia, showed that inquiry-based science education had a positive impact upon elementary school teacher and student performance. Engage, Explore, Explain, Elaborate and Evaluate are the names of the five stages in the effective model known as the 5E's.

Thank you very much to the organizers for their invitation to be here and share this conference with you.

My program sits within the Academy of Science. This will give you some context of schooling in Australia that might help you interpret what I am going to talk to you about today.

There are eight states and territories, and most of the people live in the coastal areas, and most of them live on the Eastern seaboard. There are just under 8 thousand primary schools; 126 thousand primary school teachers, and on average there are 30 students per class.

The program that I am director of is called Primary Connections Linking Science to Literacy, and this is the aim of the program:

"To improve student learning outcomes for primary school students in science and literacy, by developing professional learning programs supported with quality curriculum resources, to improve their confidence and their competence for teaching science."

The program has been going now for about eight years, and it has two main parts, which is very similar to a lot of programs in other parts of the world for scientific education. We have now developed a professional learning program that has ten modules for working with teachers and facilitators. There is also now a suite of 31 units of work that cover all the first six years of primary schooling.

This has been developed with Australian government funding of $11.2 million dollars over those years. Because we are using taxpayers' money, it has been a requirement that we monitor and evaluate the program. So there has been a program of commissioning independent, external evaluation and research, and there are over 20 research reports that can be accessed in English on our website: https://primaryconnections.org.au

Today what I am going to talk about is one of those major pieces of research that was done in 2012. The graph on next page shows the uptake of our program in Australia.

You can see over the period that teachers have embraced the kind of approach that we have been working with them and that now is used in one way or another, more so in some schools than in others, in 62% of Australian primary schools.

But this is what I will go through. I will talk to you about why we did this particular piece of research. It is based around the learning and teaching model that underpins our program called the 5Es. I will talk briefly about what that is. I will talk about what the research did, what we found, the conclusions,



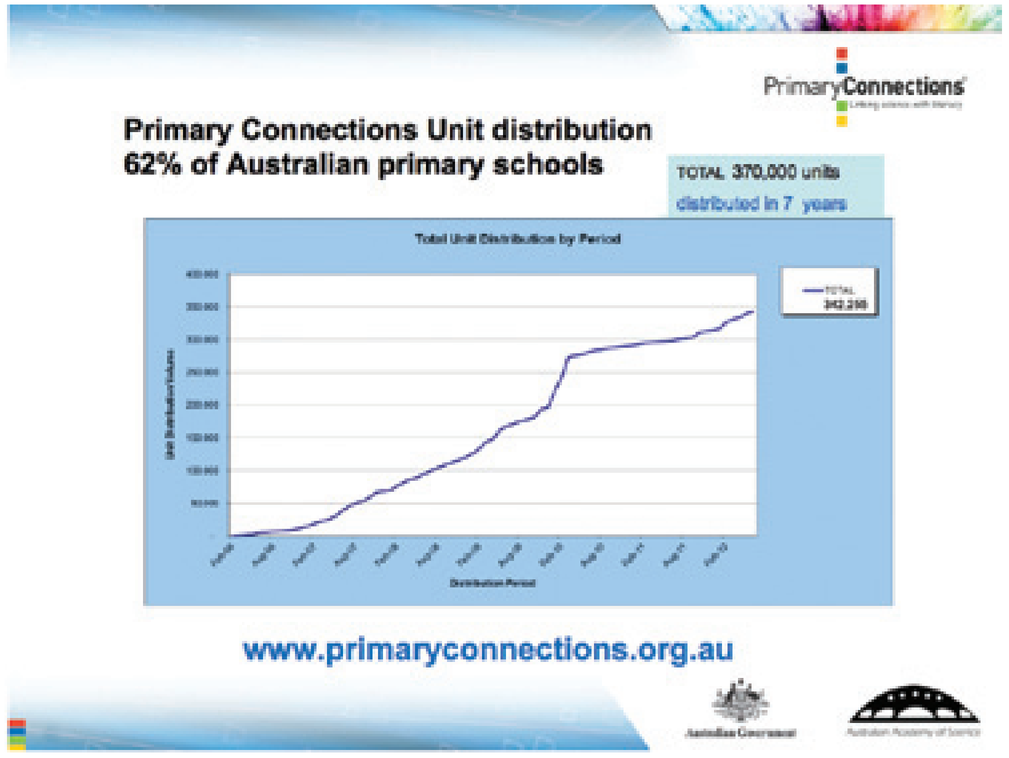



*Fig. 1*
*Fig. 1 Peers, S. Primary Connections case – Australia. PPT Presentation. http://innovec.org.mx/home/images/PresentacionesVIIConferencia/peers.pdf*

and then some implications from that. First of all, why did we commission this particular piece of research? This report was commissioned so that we could be accountable for what we were doing. We had spent a lot of money, and more importantly from my perspective, a lot of energy working on developing this program, and we needed to continuously monitor the impact that we were having; and also to continue to improve teaching.

We had a very large amount of data from the process we used in developing the program. Our 31 units of work once developed were always trialed in schools across Australia. As part of the trialing process, the teachers provided to us written feedback about each one of these units of work. This was then used to modify the materials. So we had this huge amount of feedback that gave great insight about our teacher practice. We approached to a researcher from Southern Cross University in Australia to do a review of the impact of the 5Es learning cycle that we used as the underpinning of our program.

So that you can understand some of the findings, I'll very quickly go through what the 5Es model is.

## THE 5E MODEL

The 5Es is an instructional model—a sequence of planning learning that is designed to assist students with their reasoning to help them develop better understanding. Roger Bybee in the United States of America developed the original 5Es model. We had worked somewhat with that over the eight years, and we'd done an enhanced model of it, but there is too much in that to get into it today. So to know the 5Es is sufficient. I believe that here in Mexico it is common to use a model that has four phases, and there are many such models that are based on inquiry approaches. This just happens to be the one we use, but as you will hear, we found it quite effective.

**ENGAGE** students' interest
**EXPLORE** hands-on activities
**EXPLAIN** using science ideas
**ELABORATE** conduct student-planned investigations
**EVALUATE** students' learning outcomes

The key thing to remember about this and to interpret some of the feedback from the researchers is to know this was explained in this sequence. You see that EXPLAIN is in the middle and most importantly, that ENGAGE comes first. So ENGAGE, EXPLORE, EXPLAIN before students can acquire what they've been learning, to do some ELABORATION and then there is EVALUATION.

From this model, we have various expectations about what it is that teachers and learners will talk about and do in their classrooms, and our research is about it.

## THE RESEARCH

The research questions: We wanted to know what teachers were understanding of the 5Es model, how did they implement it? And were there any factors that either obstructed it or helped them in carrying out effective teaching?

What the researchers did was a rather major piece of work. They undertook a qualitative content analysis of this written feedback from over 200 teachers that had been compiled over seven years, from 2005 to 2012, of 16 particular units. This consisted of almost 3 thousand teacher statements, so this was quite a major task.

They used multiple lenses for the analysis. There were three main ones in fact. We looked at the purposes of the phases in the 5Es that I talked about just before. We then looked at the teacher-learner roles that we expect in an inquiry approach and that is based on some work by Wynne Harlen. And the third frame of reference was to look at the feedback from the teachers in relation to what we had come to know as the components that support effective teaching and learning in inquiry-based approaches.

## FINDINGS

As well as looking at this document analysis there was also a fairly small survey done of about half of the teachers that had been involved in providing this analysis, but I won't go through that bit today.

The research was reported in these five categories, so I will look at each of those.

**I.** General implementation of the 5Es model
**II.** Implementation of the purposes of the phases of 5Es
**III.** Perspectives on teacher and learner roles (Harlen, 2009)
**IV.** Components of effective learning in science (Sis Tytler 2003)
**V.** Other issues arising

The first one: the general implementation of that model for inquiry instruction. What I will do is that in each of these five ways of reporting I will look at what was happening that was going well, and what were the things that we found teachers were finding challenging. I hope that will help you get some insight into the challenges that we had with inquiry approaches.
Ok, for the general implementation we found that the structure of using a framework for teachers to do their planning worked very positively and certainly encouraged a lot of student autonomy in undertaking their learning. In fact we even found that in some of the classrooms that the students were able to tell the teacher that "No, this is the engage phase. I'll be doing what you want later on. I'm going to engage for now." So the students started to take control over their learning.

*"I think it is important that we learn far more from the challenges than we do from the successes. And I encourage my staff to look at the positive feedback once, and the negative feedback nine times."*

There were also a lot of cases where the teachers took this planning model from their science teaching and used it in other areas of the curriculum (So it was good to see), and they started to develop an appreciation that for effective science learning being "hands on" and being engaged to learn is not enough.

Next I will talk about the challenges we found but first I would like to make a comment. I think it is important that we learn far more from the challenges than we do from the successes. And I encourage my staff to look at the positive feedback once, and the negative feedback nine times. So for every ten times we looked at their feedback, we had more to learn from what the teachers told us wasn't going well, than we did learn from what gone very successfully.
But of course, as writers of the units there's a strong temptation to nine-times read what went well and what the teachers thought was great, and only once look at what went wrong. And perhaps even then say, "Well the problem lies with the teacher". So we had to encourage our team to flip that the other way. If ever there was a negative piece of feedback we endeavored to change whatever it was that they did.

Ok. The challenges that we found were the time that it took teachers to implement these units. The first units we ever wrote, we ended up cutting them in about in half. And what happens in modern science units? Frequently, very passionate people put in everything they think is important about science, and the teacher in the classroom cannot cope with that. So we learned a very big lesson in our first year: that everything we did had to be achievable, and it had to leave the student and the teacher with a sense of success.

The other thing that we found was that when teachers got under stress they started omitting phases of the teaching and learning model. There is quite a bit of research that the learning is not as effective when that occurs.
The next thing that was looked at was the implementation of the purposes of the phases.

Now the things we were doing very, very well were that teachers were very good at engaging students and students were very good at being engaged. Students were very good at exploring. They love to tinker, have a go at things. Also we found that in the EVALUATE phase, students were very good at reviewing their understanding.
However the things that were less well done were in the ENGAGE phase. It's a case in which teachers must raise the questions from the students. Teachers found it very hard. From the feedback we learned that it was because they were afraid of what they had to do with those questions if they got them from the students. But it's a key part of inquiry that when students feel an ownership over the question that their learning is far deeper.

The other thing that we found was that students found it very difficult to compare their ideas with the ideas of other people; and of course that's an essential part of learning. If you want to improve what you think, you need to compare what you currently think to someone else. And the students needed a lot of support for that.

Ok, the third one, Perspectives on the teacher and learner roles. From a constructivist perspective we found that there were high levels of active learning, but there were low levels of the students providing evidence for their reasoning. We found that this was due a lot to the fact that teachers weren't sure how to do this as well. So we had to increase the amount of support that we provided to teachers in that area. The other thing we found was that students

## COMMENTS

*In the media in Australia, frequently there are statements: "The problem with school's science education is we've got to get the students interested in science." The problem isn't students' interest in science. It's what we do to the kids when we teach them science in schools. And clearly the problem gets worse as the students go through school.*

*There is quite a large study by the Institute of Engineers in the UK. Massive numbers of engineers were interviewed about when they first developed an interest in science. And there were some quite substantial studies to show that unless students get switched on to the idea of science as a disposition and a way of thinking by the time they are eleven years of age, they probably never will. And by the time they hit fourteen years of age it is too late. So I think you can see from this how critical it is what we do with students in the elementary years.*

found it very difficult to modify their ideas in the light of new evidence. They would have an idea before they would do a test. They would carry out the test and the results would not confirm their ideas. But they would say "I'm must've done it wrong" and they found it very difficult to entertain the idea that maybe their original idea wasn't right. Which again is the basis of learning. So, we improved the amount of scaffolding we would give to teachers, so they could help students engage with those ideas.

From the perspective of an inquiry perspective, we found this was a strong ethos, but the negative side was that there were far more teacher-guided investigations than there were student directed investigations. And we find it takes teachers two or three years before they have a level of comfort to engage with that.

Another perspective is a language perspective. Australian kids love to talk. They didn't have any problem with this. But of course it's about the quality of talk. There was a strong focus on students expressing themselves, but what was less well done was listening to others about their ideas to help frame their learning.

And finally, from an assessment perspective, we found we had some very good results. However, the results were inconsistent. A very important one was that teachers have a lot of confidence in assessing concept outcomes, about the content; but they were doing less well in assessing students inquiry skills or the processes of science.

Also, using informative assessment. Teachers were very good at assessing where the students were at, but they didn't know what to do with that information. So, again, we built that very explicitly into the units of work we will provide.

And, the components of effective teaching and learning. The strong ones to be found were that students were actively engaged with ideas. They were starting to be engaged with ideas about evidence. Also, they were developing deep understanding, so there was a move away from just doing things to actually learning. And in one way or another all the elements of the framework we use about effective learning and teaching were found to be evident.

About other issues arising: teacher beliefs impact on the way that they chose to operate in a classroom. We found that this resulted in teachers modifying the approach, before they understood the approach. We found that they would abandon or adapt lesson steps, or even lessons; or that they would switch around the order of the phases.

Once teachers understand the model, we think that it's part of the teachers' professionalism that they must take control of the teaching in the classroom. But teachers were doing this without fully understanding while they were making these choices. So, the quality of learning was less.

Another very surprising thing we found was that teachers expectations of which of the activities their students would find interesting was frequently wrong, both ways: things that the teachers thought the students would find boring and laborious, the students would love. And things that teachers expected students would find exciting, they found boring. So we had a lot to learn from that too, about the way we constructed the units.

## CONCLUSIONS

Moving on now to conclusions: overall from this study what we found was that primary teachers in Primary Connection using this 5Es model had had a very real and positive influence.

The other thing we found was how important language was as a tool of learning, both in terms of the students representing what they know and for them to learn things as well. For developing the resource materials, (we found) very strongly that learning science needs to be more than doing activities. Also, the amount of work that we expect needs to be confined. We found that a unit of work that took more than ten hours, the teachers started leaving things out, saying that they just couldn't get through that volume of work. And the other thing you could see it from the earlier comments. The purpose of the phase in which the learning is occurring must be very clear and must come out in the literature.

## MESSAGES FOR PRINCIPALS AND FOR EDUCATION SYSTEMS

All our research the only area that we found as being a challenge has been principals, because principals don't have time to engage in exactly what it is that we are trying to help teachers do, they don't appreciate how much time it takes nor how much support teachers need.

The other thing is they focus more on literacy and arithmetic, so there is quite deal of work that needs doing to help principals understand about inquiry approaches. I think it needs time. They need a chance for professional learning. They need time for preparation and they need time to reform their practice. So the message for principals and systems is that initiatives need to be sustained and they shouldn't be run off and they shouldn't stop and start. That doesn't help teachers.

## MESSAGES FOR TEACHERS

You need to encourage quite a range of science inquiry skills, and you need to provide support; especially for fair testing from the initial stages. Students have a strong sense of what is fair. A very young child can tell you what's fair and what's not fair when food is being served at the table. Their sense of fairness is very strong. So what that means in a science context is teacher support. And there is

## COMMENTS

*I will tell you about a European project that has been very successful for NTA in Sweden. It is called Fibonacci.:http://fibonacci.uni-bayreuth.de/*

*One outcome from this project was a booklet called "Tools for Enhancing Inquiry in Science Education". From these tools we developed a teacher-training course that we are trying now in Sweden. The teachers participate in three meetings and they start with these three questions that we have chosen:*

*- How do I collect the pupils' thoughts and experiences?*

*- How do I work for the pupils to realize and see how they will observe when they are doing experiments?*

*- How can the pupils show their learning and feel confidence in learning?*

*We worked with these questions in collaboration learning. One thing that we have in mind is to try to focus on teachers to ensure that pupils are at the center of the process all the time. We try to foreground delivery methods to go from that foregrounding professional knowledge and skills, and meet the double* the need to explicitly introduce this idea of evidence in a way that everybody can use in order to make decisions.

Learning through programs like Primary Connections, or perhaps there is a program in your country as well is, number one: we need advocates, we need champions, we need leaders in schools. Because it is hard work we need those people to drive initiatives forward.

We need funding, and the funding must be ongoing. We need to know who all the key players are in the positions of power, and we need to know their needs, and we need to help them understand exactly what it is we are trying to do. We need to be able to communicate our vision so people understand what we're trying to do and bring all the players along with you. We need to plan for reform as a conscious improvement. It's hard work but it's very rewarding. Keep in mind that tackling student engagement is perhaps the easiest part. They have an innate curiosity about the world; however it's teachers who need the support. And a lot of research has shown that teachers are the single most important school-based factor to impact on students learning.

I'll say that to you again because I think that it's really important. Teachers are the single most important school-based factor to improve student learning. So the focus needs to be on teachers.

## IMPLICATIONS FOR MEXICO

I think the challenge for all our Mexican colleagues will be how you are going to synthesize all the information you're going to get at this conference, and how will you decide a plan of action that will transform science education for you. What can you learn from both the successes and failures of others? What's holding you back from your goals? Do you know what your goals are and where you want to go? And will you be able to work out how you are going to get there?

Thanks for the opportunity to share with you.*

*"Learning through programs like Primary Connections, or perhaps there is a program in your country as well is, number one: we need advocates, we need champions, we need leaders in schools. Because it is hard work we need those people to drive initiatives forward."*

* Transcription

## COMMENTS

*I think the trends and challenges that I saw in former presentations are partly different, but we have a common goal in mind. What I thought was interesting was listening to people as learning to live with complexity, rather than finding simple solutions.*

*In terms of what we've learned in Scotland, there like you and like many countries internationally, Scotland has been involved in major curriculum and assessment reform. We call our reform Curriculum for Excellence. The model we used this time was successful mainly because of three issues. The first of these is that innovation has to have educational integrity. People have to believe that innovation is about what matters. Is about giving children better opportunities in learning. The Reform has to have professional integrity. So, it has to matter to the teacher. It has to deal, for example, with something matters to the teacher. It has to be done with people not to people. And the third part that matters is that is has to have systemic integrity. We believe that education innovation is like a watch: if one wheel doesn't move, everything stops. So everyone has to be part of that process and work together.*

# CONCLUSIONS

Science learning is a fundamental part of integral education for the 21st century. It should be challenging and interesting both for teachers and students, as well as for supervisors and education authorities. Far from considering science an insurmountable challenge, it should be looked upon as daring, subject to reflection and analysis for the general enrichment of society.

Educational practices should be based upon respect and promote collaboration, inclusion, and professional development among teachers. Science education is not just for experts or gifted children, it is everyone's right and it should be available to all. There are examples in other countries of how breaking paradigms can lead to effective classroom science education, especially in elementary grades, thereby promoting a positive disposition towards science in students.

Of course we must be careful. There are technical challenges that should be considered, because in addition to affording sound and rigorous science teaching and evaluation such challenges become guides to better understand what is to be taught and in what way.

Our challenge is to build environments for collaboration and reflection on assessment of science education in order to determine what we are doing well and what requires improvement. This seminar is a step in that direction.

# Panel 2 Discussion

## STRATEGIES FOR THE IMPROVEMENT OF SCIENCE LEARNING OUTCOMES ACCORDING TO THE INQUIRY BASED SCIENCE EDUCATION


**Speakers:**
- Wynne Harlen
- Rosa Devés



**Panelists:**
- Guillermo Solano
- Anne Goube
- Petra Skiebe-Correte

**Moderator:**
- Norma Sbarbati Nudelman

# Strategies for using Assessment for the Improvement of Science Learning Outcomes

WYNNE HARLEN

This paper is concerned with the use of assessment to help inquiry-based learning in science. As we will see, the goals and practices of assessment used to help learning (formative assessment) relate closely to those of inquiry-based science education (IBSE) and this relationship can help the understanding and implementation of IBSE. The paper begins by looking at the collection of pedagogical strategies that make up formative assessment and then consider their application to IBSE. Throughout it is important to clarify the terms being used, so we begin with the meaning of assessment.

## ASSESSMENT PURPOSES AND USES

Assessment is the word that is used to refer to processes of generating, collecting interpreting and using evidence to make judgements about students' learning for a particular purpose. In some countries and languages this is called 'evaluation', but generally the word 'evaluation' is used, as by the OECD, to refer to 'judgements on the effectiveness of schools, schools systems and policies' (Nusche et al 2012). Where there is likely to be any confusion we refer to student assessment or assessment of learning outcomes.

Student assessment in education serves several purposes, which fall into three main categories:

[1] to help build students' understanding (formative assessment)

[2] to provide information on students' achievements to parents, to students' next teacher as they move through the school or into high school, and (at the end of high school) to further and higher education institutions and employers (summative assessment)

[3] to hold individuals and institutions to account (assessment for accountability).

The concern in this session is with the first of these – the use of assessment to help learning, particularly in inquiry-based science education (IBSE). Assessment used in this way is called 'formative assessment' or 'assessment for learning'. It involves the on-going monitoring of students' progress towards learning goals in order to provide feedback both to the teacher and the students. Assessment for this purpose is not undertaken at one particular point in a lesson or series of lessons on a topic – as is the case for summative assessment which summarises achievement for reporting at certain times – but involves gathering and, where possible, using data about learning as it takes place.

## FORMATIVE ASSESSMENT AS A RECURRING CYCLE OF EVENTS

The actions and decisions involved in formative assessment can be represented as a cycle of events (figure 1, based on Harlen 2006). 'A', 'B', and 'C' represent activities related to the goals of the lesson or series of lessons. The goals determine what evidence to gather. This is then interpreted, used in deciding how to improve learning, leading to action in the form of subsequent activities. The processes are best explained through an example shown in the green box of next page.



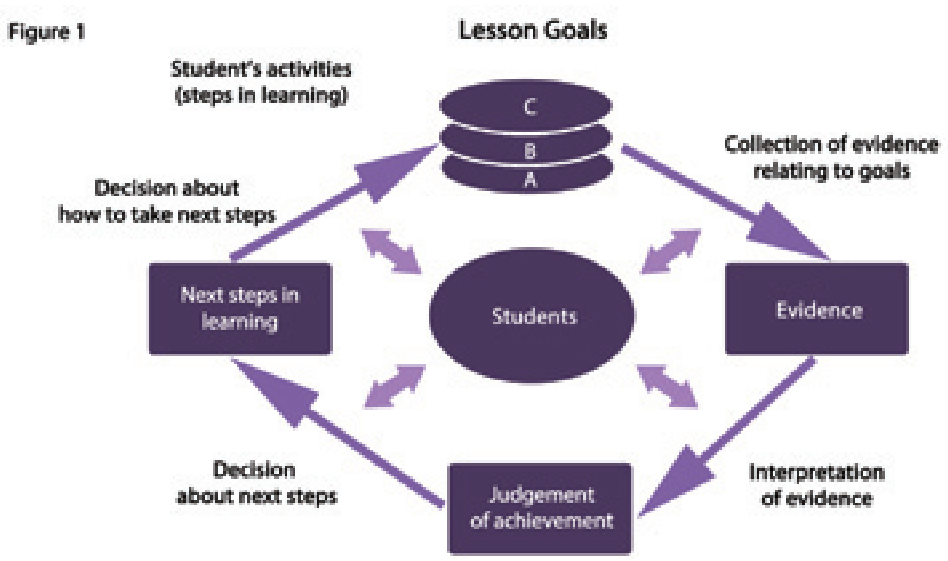
In a typical class of 10 year olds, an inquiry relating to condensation and evaporation starts from the observation that moisture forms on the outside of a drinks can just after it is taken out of the fridge. The goals of the lessons in which this phenomenon is investigated might be for students to use inquiry skills to find out about what this moisture is and where it comes from. In activity A they might plan to test their initial ideas about what the moisture is (many expect it to taste like the drink inside). Having decided it is water, another question is posed – where does it come from? They plan and conduct a test of their ideas about this (very frequently they expect it to come from the liquid inside, filtered by the metal of the can). The teacher observes what the students do, encourages them to discuss with each other, listens to their discussion, probes their thinking using open and person-centred questions (questions that ask for what students' ideas not for 'the right answer').

The teacher interprets this evidence in terms of where they have reached in relation to the goals, both in their ideas about the phenomenon being investigated (the presence of water vapour in the air) and the skills they are using in attempting to find an answer to their question. This informs decisions about the next steps to be taken. Perhaps the students have some knowledge of water vapour in the air, but have no idea about when and why it turns into water on the cold surface. So further questions are raised, leading to activity B and the cycle of data gathering and interpretation is repeated. The effects of decisions about next steps are assessed in the on-going process, which results in some progress in relevant ideas and in skills.

In Figure 1 students are at the centre of the process, since it is they who do the learning. The two-headed arrows linking students to the various parts of the assessment cycle indicate that students both receive feedback from the teacher and also provide information in what they do and say as feedback into teaching. They participate in decisions, where appropriate, through self- and peer-assessment

Of course, the process is not as tidy and formal as this representation appears to show. The actions indicated by the arrows in figure 1 are not 'stages' in a lesson nor necessarily the result of pre-planned decisions made by the teacher. They represent the thinking involved in focusing on what and how students are learning and in using this to help further learning. In some cases it may be possible for teacher and students together to decide on immediate action. In other cases, the teacher may take note of what help is needed and provide it at a later time. In this way, using assessment formatively can ensure that there is progression in learning and that students are developing understanding of what is involved, not just in these activities, but in learning, and are beginning to take some responsibility for it. But does this actually improve their achievement?

Figure 1




There is a growing accumulation of evidence that formative assessment does lead to improvement in levels of achievement (eg Black and Wiliam 1998; Black et al 2003; Brookhart 2007; Hattie and Timperley 2007; Shute 2008; Wiliam 2009; Minner 2010).

Many of the studies of the impact of formative assessment on learning highlight the central role of students in their own learning. The involvement of students in assessment of their own and each others' work is among the approaches that are most successful in raising achievement. There are examples of successful strategies for involving students from the age of five upward in assessing their work.

As well as empirical research, support for involving students in decisions about their learning also derives from theories of learning.

Current views reject the notion of learning as a matter of absorbing information and ready-made understandings from the teacher or

*"IBSE means students progressively developing key scientific ideas through learning how to investigate and build their knowledge and understanding of the world around."*

textbook. Instead, learning is seen as involving the active participation of learners in using existing ideas to try to make sense of new experiences. There is also recognition of the value of doing this with others so that ideas are constructed in the course of sharing and collaboration. Nevertheless, learning goes on inside students' heads and they must be willing to undertake it and to make the nereachjcessary effort. This being so, the way to help learning is to give the students as much opportunity as possible (appropriate to their age and stages) to know what they are aiming for in their work and how to go about it. This does not, of course, mean telling students the answers, but helping them to understand the questions. It is the difference between saying 'follow these instructions to compare X and Y' and 'find out the best way of comparing X and Y'. This may seem an obvious point but it is in fact quite uncommon for students to be able to articulate what the teacher intends them to learn from a particular activity, as opposed to what they are supposed to do.

In order to assess their work students need to realise what 'good work' means. For example, what makes a good plan for an investigation? What makes a good report of an inquiry? There are many ways of helping this understanding without imposing standards that may seem arbitrary and meaningless to students (Harlen and Qualter 2014). One approach is through groups of students brainstorming the reasons for considering one report to be better than another. Bringing together ideas from all groups leads to a list of criteria which all agree are important. This creates a useful checklist for self-assessment of their reports. Such a list seems reasonable and understandable to the students because they produced it.
In summary, the key practices of formative assessment are:

- Students being engaged in expressing and communicating their understandings and skills through classroom dialogue, initiated by open and person-centred questions.

- Students understanding the goals of their work and having a grasp of what is good quality work.

- Feedback to students that provides advice on how to improve or move forward and avoids making comparisons with other students.

- Students being involved in self-assessment so that they take part in identifying what they need to do to improve or move forward.

- Dialogue between teacher and students that encourages reflection on their learning.

- Teachers using information about on-going learning to adjust teaching so that all students have opportunity to learn.

## FORMATIVE ASSESSMENT AND IBSE

Formative assessment will only benefit inquiry-based learning in science if what is assessed, how it takes place and how the results are used reflect the principles, practices and goals of IBSE. As a recent OECD review of assessment practices recommends, assessment should be aligned with the learning goals set out in the curriculum (OECD 2013).
What should be assessed?
The learning goals of IBSE are implicit in this definition:

IBSE means students progressively developing key scientific ideas through learning how to investigate and build their knowledge and understanding of the world around. They use skills employed by scientists such as raising questions, collecting data, reasoning and reviewing evidence in the light of what is already known, drawing conclusions and discussing results. This learning process is all supported by an inquiry-based pedagogy, where pedagogy is taken to mean not only the act of teaching but also its underpinning justifications (IAP2012).

In more explicit and operational terms, when they are learning through inquiry, over a period of time, students will be undertaking actions and practices listed in Box 1.

**Box 1. Key students actions and practices in IBSE**

- Gathering evidence by observing real events or using other sources.
- Pursuing questions which they have identified as their own even if introduced by the teacher, and raising further questions.
- Making predictions based on what they think or find out.
- Suggesting ways of testing their own or others' ideas to see if there is evidence to support these ideas.
- Using and developing skills of gathering data directly by observations and measurement and by using secondary sources.
- Working collaboratively with others, communicating their own ideas and considering others' ideas.
- Assessing the validity and usefulness of different ideas in relation to evidence.
- Reflecting self-critically about the processes and outcomes of their investigations.

Of course, there are aspects of learning science, such as knowledge of scientific vocabulary, conventions and use of equipment, that are best learned through direct instruction and assessed through short classroom tests and quizzes devised by the teacher at appropriate times. Thus not all science teaching and not all assessment will be concerned with the specific outcomes of learning through inquiry. However, when understanding is the aim, inquiry has a key role in students' science education and it is necessary to face the challenge of assessing the actions and practices listed in Box 1. This brings us to the question of how such evidence can be collected.

### How is evidence collected for formative assessment in IBSE?

As mentioned earlier, evidence can be collected during classroom activities by the teachers through observing, questioning and interacting with students. This is likely to involve:

- Asking questions of a particular form – questions that probe students' understanding, ideas and reasoning (What are your ideas about …? What do you think is the reason for…? What do you think will help you to find out…?)

- Encouraging discussion, dialogue and argumentation, in which students have to give reasons for their statements and claims and use evidence to support their conclusions.

- Observing students, since for young students particularly what they think is expressed in what they do. (Noticing what variables students change in an investigation and whether appropriate variables are controlled; listening to the words they use and whether they use scientific words correctly; etc.)

Data collected in this way have to be interpreted in terms of progress towards the goals, which requires teachers to have an understanding of how students progress in their conceptual development and in the development of inquiry skills. With this understanding teachers use the data about the students' ideas and skills to decide how to proceed – what are the next steps and what intervention, if any, is needed. This brings us to the question of how to use the results.

### How are the results used?

The purpose of formative assessment is to inform any action that is needed (and action may not be necessary) whilst learning is taking place. This is where feedback comes in - feedback to the students and feedback to the teacher. Feedback to students has been identified as 'one of the most powerful influences on learning and achievement' (Hattie and Timperley 2007) but whether or not it has a positive effect on learning depends on several factors. Feedback is most obviously given by teachers to students orally or in writing, but also, perhaps unconsciously, by gesture, intonation and indeed by action, such as when assigning tasks to students. Research (Butler 1988) shows that written feedback is most effective when it is in the form of comments which indicate what students need to do to improve their work. Marks or grades don't do this and when both grades and comments are given, the students seize upon grades and ignore any comments that accompany them. When grades are absent they engage with what the teacher wants to bring to their attention. The comments then have a chance of improving learning as intended by the teacher.

In relation to the content of feedback the evidence from research and practice indicates an important difference between feedback that gives information about next steps and how to take them, and feedback that is expressed in terms of how well the student has done (this includes praise as well as criticism) rather than how well the work has been done. This applies to feedback given orally as well as in writing. By all means praise good

COMMENTS

*It is important to be clear what the objectives of teaching are. However, we must be careful that these objectives should not simplify learning in a lists about "what should be" and "what not". Learning is broad and diverse, is a process that even though you must be driven by objectives, should not be limited by them.*

*From my perspective, one of the main differences between formative and summative assessment is as follows: the formative evaluation reviewed how a learning process is occurring in the moment in which it takes place; While the summative evaluation tends to look back in time, intends to see what the student has managed to learn, or not. Both share the goal of identifying what was learned or not, or that part of the learning process that was effective or not, in order to find the points that have to be strengthened to reach the learning objectives. Both types of evaluation provide the teacher valuable and complementary perspectives.*

**Wynne Harlen**

COMMENTS

work, but recognise that this does not help further learning.

Just as important as giving effective feedback to students is that the teacher uses the data about how students are tackling their inquiries to feedback into their own actions and plans. Teachers have to plan their lessons in advance and make decisions that may not always turn out be the best in practice.

Using formative assessment provides the information needed to revise and change teaching decisions if necessary. In so doing the teacher can adjust the challenge of the students' work so that so that there is the right mixture of the familiar and the novel, so that the students are neither bored by work that it too easy nor confused by too great a challenge. It may be necessary for a teacher to change plans when students are struggling rather than risk a sense of failure. In this way the feedback enables teachers to regulate teaching to maximise learning.

## STRATEGIES FOR IMPROVING SCIENCE LEARNING OUTCOMES

The practices of formative assessment relate to learning in any domain and will not tell us what action to take in relation to developing scientific understanding and science inquiry skills. For this we need to turn to what is known from experience and about how learning takes place. In the case of inquiry skills some general strategies are summarised in Box 2, while Box 3 indicates some strategies for helping students in developing scientific ideas.

**Box 2. Some general strategies for helping progression in inquiry skills**

A. Provide opportunity to use inquiry skills in exploring materials and phenomena at first-hand.
B. Ask question that require the use of the skills (and allow time for thinking and answering).
C. Provide opportunity for discussion in small groups and as a whole class.
D. Encourage critical review of how activities have been carried out.
E. Provide access to the techniques needed for advancing skills.
F. Involve students in communicating in various forms and reflecting on their thinking.

**Box 3. Some general strategies for helping development of scientific ideas**

- Extend experience so that non-scientific ideas are challenged.
- Scaffold the introduction of alternative ideas for students to test.
- Give opportunities for new ideas to be applied in different situations.
- Develop reasoning about changes that are being judged only from appearances.
- Develop inquiry skills so that relevant evidence is used in drawing conclusions.
- Create links between events with a common explanation, making ideas 'bigger'.
- Discuss with students the meaning they are giving to words related to science concepts.

## IN CONCLUSION

Formative assessment and IBSE are approaches to teaching that have much in common. Both start from the existing ideas and skills that students being to the classroom, formed from earlier experience both within and outside school. Both promote students' active learning – not just physical activity in manipulating objects, but mental activity. Both focus on progression in learning and both are underpinned by a view of learning as constructed by students in the company of and through interaction with others. But they are not identical and each has a particular contribution to make to effective science education. Formative assessment, whilst not concerned with students gathering and testing evidence, contributes important points about feedback and student self-assessment. Thus the implementation of IBSE has much to gain from simultaneous implementation of formative assessment. We need both.*

* Document for the Presentation

## COMMENTS

*Every evaluation involves communication. We are not used to seeing this, but it is present all the time. When you ask a question, a student answers and you react to that answer, that is a process of communication.*

*There is not much research on formative assessment among populations with cultural and linguistic diversity. (At Boulder University) we are trying to find key aspects in formative assessment that could fail if teachers lack better training for such diversity.*

*What we have found up until now is that many teachers definitely either do not speak to their students or do not speak to all of them in the same way. Many teachers speak more often to students they know will give the right answers. But they need to focus more on the students who do not participate, because it is more valuable to know if students actually know something and if not, understand why not.*

# REFERENCES


Black, P., Harrison, C., Lee, C., Marshall, B. and Wiliam, D. (2003). Assessment for Learning: Putting it into Practice. Maidenhead England: Open University Press.

Black, P. and Wiliam, D. (1998) Assessment and classroom learning, Assessment in Education, 5 (1) 7-74.

Brookhart, S.M. (2007) Expanding views about formative classroom assessment. A review of the literature. In J.H. McMillan (ed), Formative classroom assessment: Theory into practice. New York: Teachers College Press, 43-62.

Butler, R. (1988) Enhancing and undermining intrinsic motivation: the effects of task-involving and ego-involving evaluation on interest and performance. British Journal of Educational Psychology, 58, 1-14.

Harlen, W. (2006) Teaching, Learning and Assessing Science 5 – 12. 4th edn. London: Sage.

Harlen, W. and Qualter, A. (2014) The Teaching of Science in Primary Schools 6th edition. London: Routledge.

IAP (2012) Taking Inquiry-Based Science Education into Secondary Education. Report of a Global Conference. http://www.sazu.si/files/file-147.pdf

Minner, D.D., Levy, A. J and , Century, J. (2010) Inquiry-Based Science Instruction—What Is It and Does It Matter? Results from a Research Synthesis Years 1984 to 2002, Journal of Research in Science Teaching, 47 (4) 474-496.

Nusche, D. Laveault, D., MacBeath, J. and Santiago, P. (2012) OECD Reviews of Evaluation and Assessment in Education: New Zealand 2011. OECD Publishing.

OECD (2013) Synergies for Better Learning: an international perspective on evaluation and assessment. Paris: OECD.

Shute, V.J. (2008) Focus on formative feedback. Review of Educational Research, 78, 153-189.

Wiliam, D. (2009) An integrative summary of the research literature and implications for a new theory of formative assessment, in (eds) H. L. Andrade and G. J. Cizek, Handbook of Formative Assessment, New York: Taylor and Francis.


# COMMENTS

*I recommend teachers to develop a chart for children self-assessment, in order to identify how they develop different skills that are very important in IBSE. Each child should have a chart with skills like this:*

| IBSE UNIT | | | | | | |
|---|---|---|---|---|---|---|
| **SKILLS** | **Class Session** | **Class Session** | **Class Session** | **Class Session** | **Class Session** | **Class Session** |
| Conduct investigations | ● | | | | | |
| Observe | ● | | | | | |
| Measure | x | | | | | |
| Compare | x | | | | | |
| Select and classify | x | | | | | |
| Interpret | x | | | | | |
| Collaborate with other students | ● | | | | | |
| Record/fill up personal note book | x | | | | | |
| Explain to others /Discuss | ● | | | | | |

**Anne Goube**

*At the end of each class session students mark with a green dot the skill they developed the most successfully, with an orange dot does that are on its way, the ones that are not perfect yet, and red to mean that they were not successful on this. The interesting thing is that at the end of the Unit all the red dots become orange, and the orange dots become green. So that kind of chart gives a very positive feedback to the student. Even if you have many children in your classroom there is a track that you can follow and see how each individual student is doing. If you use in IBSE a chart like this, it can show the parents what has been done during the class. It means that as a teacher you had been working on skills that are very important like collaborate with other students, explain to others and discuss. Even if the learning goal is not achieved at the end of a class session, it would come at the end of the Unit. Ihope this assessment resource can be useful to you.*

COMMENTS

# Formative Assessment as means to induce reflection and the generation of positive learning environments

Rosa Devés

We will address formative assessment as a process that is able to support learning processes through: i) the stimulation of evidence-based reflection on what we learn and how we learn and ii) the generation of environments that can favor learning by promoting dialogue and collaborative work. Both dimensions are particularly relevant to science education because, reflection on learning and fruitful interactions with others, are core elements of the scientific practice.

These two aspects of formative assessment will be analyzed from the evidence gathered in the course entitled "Scientific Inquiry at the School", which is offered to students of different programs at the University of Chile. The course aims to familiarize students with the fundamentals of inquiry-based science education and to provide opportunities for them to link this knowledge to their own learning processes. In particular, in this presentation we give an account of the effect of formative assessment in the learning outcomes of the students and the increased understanding of the teaching team about the formative process.

## THE COURSE ENTITLED "SCIENTIFIC INQUIRY AT THE SCHOOL"

The course was developed as part of a General Education program and has been offered four times during one semester since 2010. Both its conception and execution are closely linked to the Inquiry Based Science Education Program (ECBI) that has been applied for a decade in elementary public schools in Chile. The teaching staff consists of one academic and three specialists from the ECBI program, that work in collaboration with classroom teachers from seven schools.

In its first two versions, the course was offered only for undergraduate science students and in following years the composition of the group has progressively diversified, both in terms of discipline, and the level of study. In 2013, the class was formed by eight undergraduate students enrolled in Biochemistry (2), Biology (1) Dentistry (1) Early Childhood Education (4) and by two students of the Ph.D. Program in Biomedical Sciences (a medical doctor and a medical technologist). Undergraduate students were at different levels of their careers from 2nd to 5th year, five of them were women and five were men. Therefore, the group of students in the 2013 version (that is analyzed here) was highly diverse. This condition was understood as an opportunity to generate a more fertile and interesting learning environment for the development of all students.

*"There is an abundant evidence that the use of formative assessment can be an important facilitator of learning for students as well as for teachers"*
**(Wynne Harlen, 2013)**

In the course the students face the challenge of creating and implementing an inquiry science lesson at the elementary level, while they learn about the bases of the inquiry approach, its objectives, its pedagogy, and the systemic challenges involved in its implementation. Since the course is not specifically aimed at students of pedagogy, but falls within the scope of general education, an underlying aim is to offer to them opportunities to reflect on their own learning processes.

A central aspect is that students learn in different contexts, the university and the school, and this encourages them to analyze the processes and outcomes of their work, exercise collaborative work, communicate and share with other people experiences and ideas (their peers, their teachers, school teachers and children). All of this while they become aware of the responsibility that involves working closely with the public school system.

## APPLICATION OF FORMATIVE ASSESSMENT

Because of the educational principles that guide the course, as well as because of the learning that it seeks to promote in a diverse group of students, assessment follows the guidelines of formative assessment. From the beginning, we have been progressing from a more intuitive application of these principles and practices to a systematic application based on the principles and strategies that are presented and discussed in the book "Assessment and Inquiry-Based Science Education" ( Wynne Harlen, 2013).

Assessment is structured around the following concepts and practices:

• Inquiry Methodology
Assessment is embedded in an inquiry teaching and learning process that is consistent with the objectives of the course. Teachers act as facilitators by offering the students different opportunities to develop understanding, skills and attitudes.

• Person – centered assessment
The questions and challenges are posed so that the students confront their visions and ideas with those that arise from their interaction with other people and contexts (other students, teachers, children and teachers of the school system) as well as from reflection of their own inquiry experience.

• Progression
At the beginning of the course, assessment is primarily oriented to bring out the experiences, beliefs and interests of the individual students. As the course progresses, focus is transferred to the learning goals related to the principles and practices of the inquiry methodology. From another perspective, the student's attention is guided from a reflection which is initially centered on themselves (previous experiences, concerns), to the challenge of generating science learning in children (the development and implementation of the inquiry lesson). The entire process is documented in the student portfolio in which students gather the evidence of the work they have done, their readings, observations and the reflections derived from the analysis of the evidence. After completion of the course, review of the entire process is encouraged, so that students can recognize their achievement and progress and also perceive the future challenges.

• Dialogic Relationship
In all activities dialogue is encouraged (student-student, student-teacher, teacher-teacher). Students are also given the responsibility to establish effective dialogues with the teachers at the schools in the process of preparing the inquiry lesson to be implemented.

*"Teachers facilitate the learning process. They offer to the students various opportunities to develop compehension, and different capabilities."*

These conversations which describe their ideas, predictions, observations and experiences, are an important source of information about the progress of their learning and the evolution of their thought.

• Continuous improvement and feedback
Feedback on the work done by the students is delivered through written comments and / or discussions, either individually or in groups. Feedback also includes questions or suggestions about ways to further the development of concepts, skills and attitudes. Whereas, during the development of the course some assessment also includes a mark (because it is required by the system), students are called to complete or further their work and are informed that these marks are only referential to assess progress. In this way they are encouraged to engage in continuous improvement as well as in the review of their work. After completion of the course, assessment is delivered in the form of a text that refers to the progress and achievements of the students from the evidence gathered in the portfolio, and an overall mark.

• Systematic and continuous reflection by teachers.
Teachers, like the students, maintain a dialogic relation, and after each class meet to decide the next steps of instruction, including a revision of the challenges that will be presented to students. The comments, as well as the marks, are discussed and decided among the whole team. There is concern to gather evidence on progress in different dimensions ranging knowledge, skills and attitudes.

## RESULTS

In the following sections we present observations regarding the results of the application that can be deduced from the records in the student portfolios. These are focused on the two aspects of formative assessment set out above: the ability to induce reflection and the generation of positive learning environments. The oral presentation will illustrate these conclusions with evidence in their own words.

Students expressed that as a result of experience they have been able to:

• Deepen their conceptual understanding using "experience as a key to generate thought."

• Widen their visions about education by learning from the experiences of other students and teachers.

• Become aware of the importance of planning and rigor in pedagogy and develop responsibility.

• Visualize the complexity of the teaching practice and profession, and the importance of collaboration to face this challenge.

• Experiment personal changes; among these, strengthening their commitment to the teaching profession and to education in general.

• Detect the progress in understanding from the analysis of the progress of their work.

• Acquire "new ideas or ways of understanding" as a result of the contact with the reality of school and group reflection of these experiences with students from different disciplines, "I learned to use my imagination more, to develop empathy towards children, to discover new tools."

• Recognize the value of paradigms that are different from yours and experience "radical changes in vision"; this is more evident from the social sciences towards the natural sciences.

• Recognize the importance of being allowed to make mistakes, to have a place to "learn from others, not to fear" and to rethink many times the ideas in the search for greater understanding.

• Develop self-criticism and self-evaluation with the stimulus of group discussion.

The practice of formative assessment has also helped the teachers to recognize situations that would not have been evident without the information that comes from this approach.

## COMMENTS

*There is something profoundly different between formative assessment and summative assessment, and that difference lies in teachers' affections. When knowledge is reported in summative assessments students could be judged. This is unfortunate but it happens. By contrast, in formative assessment (a teacher) accompanies someone in his learning. Thus the student is truly at the center and this changes the perspective of affection, so communication changes. Additionally, something happens to the teachers because accompaniment makes it necessary for them to know (their students). Understanding is more than just measuring. Understanding is much more complex.*

These can be summarized in the following points:

• The greatest gain in learning achievement, in successive versions of the course, has come, in part, from the growing diversity of the group and the use of formative assessment to benefit from this diversity. Thus the course which initially was aimed for students of scientific careers, has become a learning experience of higher potential, as a result of diversifying the student disciplinary fields. It has also been beneficial that the group is made up of students with different levels of university experience.

• Interaction with the school system is one of the most powerful sources of learning and the reflection around these experiences with others and with the support of literature, expands ideas and consolidates new knowledge.

• Both students that come from the natural sciences, as those who come from the social sciences (education), state that at the beginning the language seemed foreign to them; at this point the feeling is that the language that is spoken and heard corresponds to the language "of the other." This shows that all students must leave their comfort zone to enter a different learning space. The support that can be granted through formative assessment, in this process of adaptation, is critical.

• The two characteristics of formative assessment that we have addressed are interdependent: concrete experience and reflection are enriched when others are involved, and the ability to establish meaningful and productive relationships is strengthened through reflective dialogue about experiences.

• Formative assessment can lead not only to improved educational processes in the specific area in which it is applied, but it can also contribute to educational knowledge, through inquiry and the conceptualization of expert pedagogical work in a community of learning environment.*

*"Formative assessment (...) can also contribute to educational knowledge, through inquiry and the conceptualization of expert pedagogical work in a community of learning environment."*

* Document for the Presentation

# REFERENCE


Wynne Harlen (2013), Assessment in Inquiry based Science Education: Issues in Policy and Practice. Published by the Global Network of Science Academies (IAP) Science Education Programme (SEP). www.interacademies.net/activities/projects/12250.aspx
TWAS-Strada Costiera, 11-34151, Trieste, Italy.


## COMMENTS

*People or children learn when they know what they have to learn, what the lesson goals are, and they also need to know how to get to that goal. So, we should not hide what has to be learned. If you and the children know the goals you can do formative assessment, because you know what you are looking for. If you know the goal, you also know what to look for, if children reach the goal or not, and if they have to change some practices in order to reach that goal.*

*Knowing the goal also allows children's self assessment. It gives them the possibility to identify what is good work and what is not, and to assess their own work compared to peers' work.*

# CONCLUSIONS

Formative assessment, which means continuously reviewing student progress towards learning goals has proven to be a great ally of inquiry-based science education as it potentiates its impact. Both have a lot in common: they promote active learning built by students interacting with others.

It is therefore important to promote formative assessment in classrooms recognizing the core role students play in their own learning, motivating self-evaluation and peer-evaluation in order to make it more effective and significant. Achieving this requires that teachers as well as students are clear on the objectives of teaching so as to identify whether activities will lead to a satisfactory completion of established goals.

It is also fundamental to insure an appropriate environment in the classroom respectful of diversity, and where it is also recognized that learning processes are accompanied by good and bad choices, the latter of which are highly challenging and enriching.

# Panel 3 Discussion

## INQUIRY BASED SCIENCE EDUCATION ASSESSMENT PROCESSES IN NORTH AND LATIN AMERICA


**Speakers:**
- Hubert Dyasi
- Jorge Alejandro Neyra (represented by: Raymundo Edgar Martínez Carbajal)



**Panelists:**
- Cristina Aguilar Ibarra
- María Figueroa
- Patricia Rowell

**Moderator:**
- Reyes Tamez Guerra

# Next Generation Science Standards and implictaions for assessment in the United States of America

HUBERT DYASI

This article comments why inquiry-based science education is seen as 'cornerstone' in education in the United States of America; why 'the trio of concepts, practices, and epistemology is at the heart of the efforts to revise K-12 science standards; the claim that "the US new science education Standards build upon research-based cognitive models of how learning unfolds over time; and the implications of these developments for assessment.
.

## WHY INQUIRY-BASED SCIENCE EDUCATION IS SEEN AS 'CORNERSTONE' IN K-8 SCIENCE EDUCATION

The 1996 National Science Education Standards (NSES) describes science inquiry as "the diverse ways in which scientists study the natural world and propose explanations based on the evidence derived from their work" (National Research Council, 1996, p. 23). The NSES argued very strongly that since science inquiry is such a central feature of science content it should be a major component of students' science activities to enable them to "develop knowledge and understanding of scientific ideas, as well as an understanding of how scientists study the natural world" (loc. cit.). It further stated that inquiry-based science "is basic to science education and a controlling principle in the ultimate organization and selection of students' activities" (p. 105). Inquiry-based science education is essential because it reflects science as practiced in the real world as a way of achieving knowledge and understanding about the world.

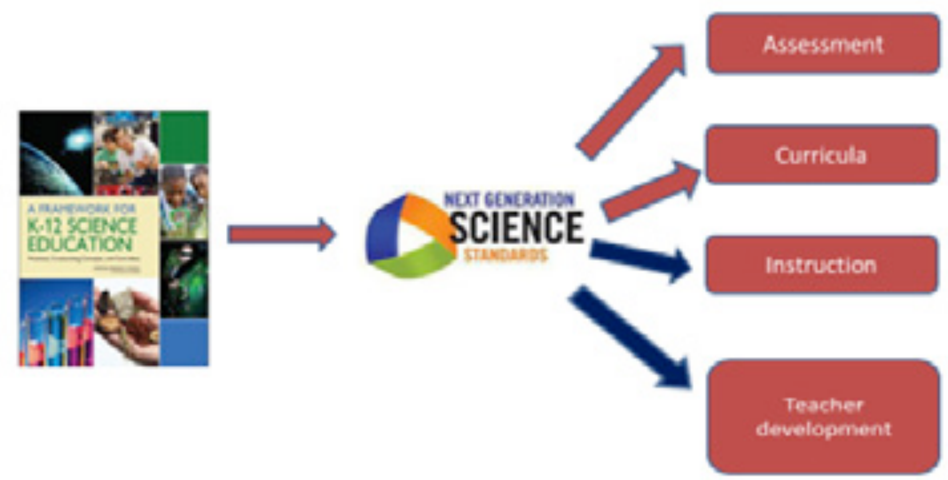
When skillfully practiced in classrooms, inquiry-based science education creates, as it does in professional science labs, a community of inquiry where students engage directly and intellectually in the practices of science. They interact with each other, with their teacher as co-inquirers into phenomena, and participate in critical but friendly discussions and arguments about their inquiries in pursuit of scientific understanding and explanation. These

*"These experiences make students' science knowledge and capacities visible. They also help to bring out in students what numerous research studies have confirmed, that they have the intellectual capability to learn science and are capable of causal reasoning. "*

experiences make students' science knowledge and capacities visible. They also help to bring out in students what numerous research studies have confirmed, that they have the intellectual capability to learn science and are capable of causal reasoning. Also, they can discriminate between reliable and unreliable sources of knowledge and demonstrate that they have the cognitive capacity to engage in serious ways with the enterprise of science. (National Research Council, 2007, p. vii).

Inquiry-based science education in the US got a big boost when the National Assessment of Educational Progress (NAEP) made it central in its 2009 science framework. That framework outlines principles of science inquiry, concepts, and applications of science on which NAEP survey tests would be based (National Assessment of Educational Progress, 2009: Science Framework for the 2009 National Assessment of Educational Progress). Inquiry-based paper-and-pencil questions were later included in NAEP tests and more recently NAEP has introduced computer simulations and actual hands-on components in its large-scale norm-referenced tests to assess national trends in inquiry-based science. Presuming that these measures are valid and reliable, then for them to be also fair students' inquiry-based science education should be at the core of students' science learning in schools. Similarly, at the international level the Science Education Programme of the Global Network of Science Academies has published detailed descriptions and significance of the inquiry-based science approach and of its centrality in science education at all educational levels beginning in early childhood education. For example, in Inquiry-Based Science Education: An overview for educationalists, it puts forward justifications for inquiry-based science education. It claims that inquiry-based science education enables learners to

**A Framework to guide changes in science education in K-12**




Rick Duschl, NSTA Web Seminar presentation, p.11

understand aspects of the world around them, both natural and those created through application of science; it develops a basic understanding of what science is, how it works and what are its strengths and limitations; it cultivates skills for communicating experiences and ideas in science; it enriches the linguistic and representational skills needed for effective expression of thoughts and ideas in science to advance arguments, justifications, and constructed explanations; it nourishes an ability to continue learning leading to further development of concepts, skills, attitudes, knowledge and understanding is regarded as more important than accumulating large amounts of factual knowledge (Wynne Harlen & the IAP Working Group, 2009, pp. 21-22).

The promotion of inquiry-based science education recently reached its highest point in the US with the publication of the National Research Council's research-based document,

## A FRAMEWORK FOR NEW K–12 SCIENCE: PRACTICES, CROSSCUTTING CONCEPTS, AND CORE IDE

The Framework is a greatly evolved view of inquiry-based science education and portrays the scientific enterprise as knowledge- and theory-building. It presents a coherent vision of K-12 science education in three ways.

" First it is built on the notion of learning as a developmental progression. It is designed to help children continually build upon and revise their knowledge and abilities, starting from their curiosity about what they see around them and their initial conceptions about how the world works. The goal is to guide their knowledge toward a more scientifically based and coherent view of the sciences and engineering, as well as of the ways in which they are pursued and their results can be used.

Second, the framework focuses on a limited number of core ideas in science and engineering both within and across the disciplines....

Third, the framework emphasizes that learning about science and engineering involves integration of the knowledge of scientific explanations (i.e. content knowledge) and the practices needed to engage in scientific inquiry and engineering design. Thus the framework seeks to illustrate how knowledge and practice must be intertwined in designing learning experiences in K-12 science education." (National Research Council, 2012, pp. 10-11).).

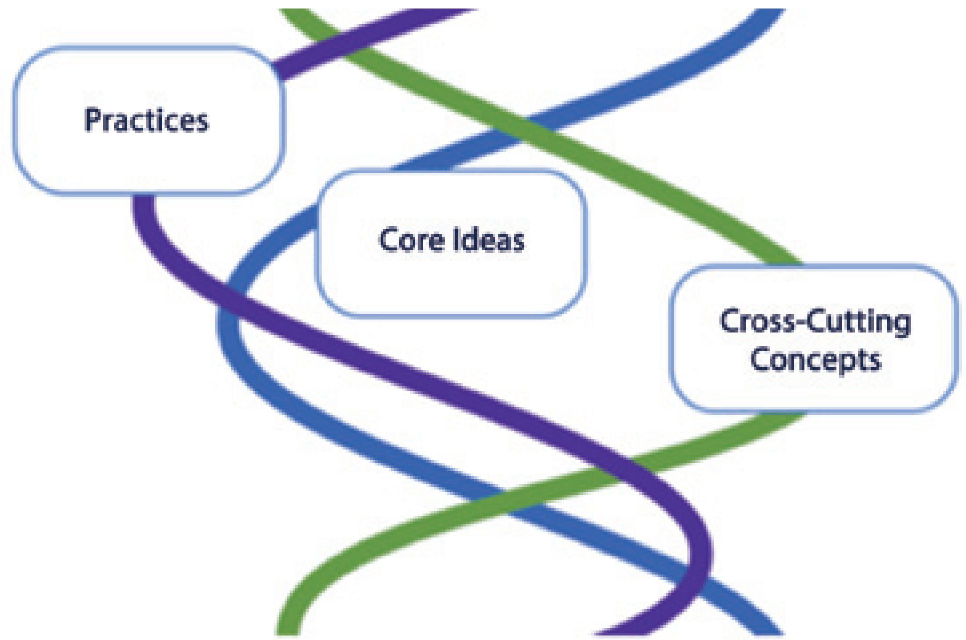



Diagram from Helen Quinn and Heidi Schweinburger; Seminar presentation.

The Framework has three intertwined dimensions, namely scientific practices and engineering design (e.g. asking questions in the case of science and defining problems in the case of engineering; constructing explanations in the case of science and designing solutions in the case of engineering), crosscutting concepts (e.g. patterns, cause and effect, structure and function) and disciplinary core ideas (e.g. matter and its interactions - physical sciences; from molecules to organisms in life sciences; earth's place in the universe in earth and space sciences; and engineering design - engineering, technology, and applications of science) (See APPENDIX 1 for full list).

All three dimensions are organized in learning progressions simultaneously within a grade level band and across grade level bands. Science learning progressions are:

"(...) empirically grounded and testable hypotheses about how students' understanding of, and ability to use, core scientific concepts and explanations and related scientific practices grow and become more sophisticated, with appropriate instruction…These hypotheses describe the pathways students are likely to follow to the mastery of core concepts. They are based on research about how students' learning actually progresses – as opposed to selecting sequences of topics and learning experiences based only on logical analysis of current disciplinary knowledge and on personal experiences in teaching. These hypotheses are then tested empirically to assess how valid they are. " (Corcoran, Mosher, & Rogat, 2009, p. 8; see also National Research Council (2007), and Duncan & Rivet, 2013)

Next Generation Science Standards (NGSS) is a reflection of the Framework (Achieve, 2013)http://www.nextgenscience.org/. The example below illustrates how a standard in the K-2 grade range combines practices, core disciplinary ideas, and crosscutting concepts in a single statement regarding students' performance expectations around the disciplinary core idea of earth's systems. The example also shows connections to the nature of science, to the same core idea at higher grade levels, to language arts or literacy, and to mathematics.

## WHY 'THE TRIO OF CONCEPTS, PRACTICES, AND EPISTEMOLOGY IS AT THE HEART OF THE EFFORTS TO REVISE K–12 SCIENCE STANDARDS

Disciplinary science core ideas serve as anchors around which to build more and deeper understandings of subject matter. Research comparing performance of experts and novices in all fields of study found that "experts, regardless of the field, always draw on a richly structured information base… Deep understanding of subject matter transforms factual information into usable knowledge." Research further shows that a pronounced difference between experts and novices is that experts' command of concepts shapes their understanding of new information: it allows them to see patterns, relationships, or discrepancies that are not apparent to novices… their conceptual understanding allows them to extract a level of meaning from information that is not apparent to novices, and this helps them select and remember relevant information… Experts are also able to fluently access relevant knowledge because their understanding of subject matter allows them to quickly identify what is relevant." (National Research Council, 1999, p.12)

By focusing on acquisition of science concepts it is believed that successful implementation of NGSS can help students progress towards developing and using a conceptual framework of science that will enable them to progressively become better than novices and think more like scientists. A student who has acquired a conceptual framework is better able to apply what was learned in new situations and to learn related information more quickly. (loc. cit.). For example, a student who has a conceptual understanding of local weather can relatively easily understand weather phenomena in another part of the earth.

Science practices are the vehicle that carries scientific inquiry forward. They help focus attention and generate questions that can be answered scientifically yielding demonstrated evidence. Engagement in all the practices is the lifeblood of doing science and of the advancement of the field; without them, there would be no progress in science as creator of testable models, explanations, and theories. There would be no scientific community as we know it.

It is important for students to have evidence-based criteria for making judgments about their science inquiries. But they must know and

COMMENTS

*Just as in inquiry-based science education, the concept of assessing learning outcomes has not been fully understood in its exact dimension. A proper learning assessment proposal for inquiry-based education should begin with a conceptual framework and the purposes of this model, in addition to an updated conception of assessment supported by national and international progress on the subject. An assessment of this kind requires conceiving students as human beings capable of building scientific literacy and resorting to their intellectual capabilities. It also calls for proper planning and rigorous methodology that focuses on measuring students' progressive understanding of scientific ideas and in the development their skills.*

*Evaluation can come in various modalities according to what is to be evaluated. The objectives and characteristics of the inquiry-based science education model determine the importance of keeping an eye on formative assessment. What do we mean by formative assessment? Felipe Martínez Rizo characterizes it by stating that in a broad sense formative assessment refers to the feedback teachers usually give their students. According to this idea, formative assessment has been present in classrooms*

**K-ESS2-1 Earth's Systems**

Students who demonstrate understanding can:

K-ESS2-1. Use and share observations of local weather conditions to describe patterns over time. [Clarification Statement: Examples of qualitative observations could include descriptions of the weather (such as sunny, cloudy, rainy, and warm); examples of quantitative observations could include numbers of sunny, windy, and rainy days in a month. Examples of patterns could include that it is usually cooler in the morning than in the afternoon and the number of sunny days versus cloudy days in different months.] [*Assessment Boundary: Assessment of quantitative observations limited to whole numbers and relative measures such as warmer/cooler.*]

The performance expectation above was developed using the following elements from the NRC document *A Framework for K-12 Science Education*:

| Science and Engineering Practices | Disciplinary Core Ideas | Crosscutting Concepts |
|---|---|---|
| **Analyzing and Interpreting Data**<br>Analyzing data in K–2 builds on prior experiences and progresses to collecting, recording, and sharing observations.<br>• Use observations (firsthand or from media) to describe patterns in the natural world in order to answer scientific questions.<br>- - - - - - - - - - - - - - - -<br>-**Connections to *Nature of Science***<br>**Science Knowledge is Based on Empirical Evidence**<br>• Scientists look for patterns and order when making observations about the world. | **ESS2.D: Weather and Climate**<br>1. Weather is the combination of sunlight, wind, snow or rain, and temperature in a particular region at a particular time. People measure these conditions to describe and record the weather and to notice patterns over time. | **Patterns**<br>• Patterns in the natural world can be observed, used to describe phenomena, and used as evidence. |

*Connections to other DCIs in kindergarten: N/A*

*Articulation of DCIs across grade-levels:*
**2.ESS2.A ; 3.ESS2.D ; 4.ESS2.A**

*Common Core State Standards Connections:*

*ELA/Literacy -*

| | |
|---|---|
| **W.K.7** | Participate in shared research and writing projects (e.g., explore a number of books by a favorite author and express opinions about them). (K-ESS2-1) |

*Mathematics -*

| | |
|---|---|
| **MP.2** | Reason abstractly and quantitatively. (K-ESS2-1) |
| **MP.4** | Model with mathematics. (K-ESS2-1) |
| **K.CC.A** | Know number names and the count sequence. (K-ESS2-1) |
| **K.MD.A.1** | Describe measurable attributes of objects, such as length or weight. Describe several measurable attributes of a single object. (K-ESS2-1) |

adopt standards of evidence or epistemological norms inherent in various science disciplines so that they know what is credible evidence in a given science discipline. On those bases they can distinguish between what is wrong and what is right; they can also engage in forceful arguments about relevant science models and theories knowing very well what is significant and what is trivial in the discipline.

The 'trio' of dimensions was adopted to address perceived weaknesses in K-12 science education in the US. Some of the major weaknesses, are that K-12 science education "is not organized systematically across multiple years of school, emphasizes discreet facts with a focus on breadth over depth, and does not provide students with engaging opportunities to experience how science is actually done" (National Research Council, 2012, p. 1). The Framework and NGSS suggest that these weaknesses can be removed by first creating and implementing science education standards that integrate science and engineering practices, core disciplinary ideas, and crosscutting concepts embedded in science learning progressions. Emphasis on the integration of the 'trio' and focusing on fewer science core ideas helps to avoid shallow coverage of many topics without going into depth; prevents treatment of science as disjointed topics; meaningless repetition of topics; and "helps clarify what is most important to spend time on and avoid proliferation of detail to be learned with no conceptual grounding" (p. 11). The dimensions are an essential component of overcoming the weaknesses by articulating broad research-based sets of expectations of students in science.

The three dimensions help to clarify science. For example, while the sciences share features and abilities described under practices, their disciplinary core ideas differ or emphasize different perspectives, hence the separate core ideas according to different groupings of disciplinary knowledge. Crosscutting concepts transcend individual sciences and buttress the idea of the presence of unifying concepts across the various science disciplines. It is vitally important for students to grasp relationships among science discipline-based core ideas and crosscutting concepts as they engage in scientific investigations characterized by science practices.

## RESEARCH-RELATED JUSTIFICATION

The Framework proposals, which have now been formulated into NGSS, are based on solid research findings in the learning of science. In 1999 the National Research Council published a ground-breaking report entitled How People Learn (National Research Council, 1999). The report consolidated research findings on learning in a variety of fields of study. In research that compared experts to novices, findings showed that regardless of field, experts draw from a very rich knowledge base to ask questions, to notice patterns, and to present reasonable arguments. An existing command of concepts in an intellectual framework shapes the way new information is understood. It stands to reason, therefore, that in science education schoolchildren should have opportunities in school to successfully work and learn to think like scientists. In this respect a special report, How People Learn: Bridging Research and Practice (National Research Council, 1999) highlights three relevant research findings relating to children, upon which classroom practice can be based:

1) Students come to the classroom with preconceptions about how the world works. If their initial understanding is not engaged, they may fail to grasp new concepts and information that are taught, or they may learn them for purposes of a test but revert to their preconceptions outside the classroom.
2) To develop competence in an area of learning, students must have both a deep foundation of factual knowledge and a strong conceptual framework.
3) Standing and progress in problem solving. (pp. 10, 12-13).

Later, Taking Science to School: Learning and Teaching Science in Grades K-8 (National Research Council, 2007) examined and reported an enormous amount of research studies and findings from cognitive and developmental psychology, science education, and the history and philosophy of science to synthesize what is known about how children in grades K through 8 learn the ideas and practice of science. The study reported that "students learn science by actively engaging in the practices of science; a range of instructional approaches is necessary as part of a full development of science proficiency" (p. 3). Also, "children's rich but naïve understandings of the natural world can be built on to develop their understandings of scientific concepts. At the same time, their understandings of the world sometimes contradict scientific explanations and pose obstacles to learning science. It is thus critical that children's prior knowledge is taken into account in designing instruction that capitalizes on the leverage points and adequately addresses potential areas of misunderstanding." (p.3) Equally important, it has been demonstrated that even young children "demonstrate causal reasoning, and are able

COMMENTS

*recognize when learning goals have been achieved or require revision or improvement. In formative assessment feedback is key. The important thing here is not to find out if a student masters a subject or not, but to effectively contribute to students' progress in learning and achieving a given goal—namely, student learning.*

❖

*If we consider that inquiry-based science teaching and formative assessment are congruent and complementary, we can admit they are "made for each other" and therefore, it is highly advisable to recommend and reinforce incorporating formative assessment in classrooms where learning is through inquiry.*

*I have emphasized formative assessment, but it would be unfair not to mention summative assessment. Good summative assessments, understood as final exams, should be ruled by technical requirements and procedures derived from investigations on evaluation. Therefore they can provide valuable general information to complement data from formative assessments on academic performance as well as the strengths and weaknesses of the model.*

to discriminate between reliable and unreliable sources of knowledge" (p.vii.), regardless of gender, race, or socioeconomic circumstances. These attributes, however, do not come up spontaneously, they must be nurtured with "carefully structured experiences, instructional support from teachers, and opportunities for sustained engagement with the same set of ideas over weeks, months, and even years." (p. 3)

## LEARNING EVOLVES OVER TIME

Findings from the research literature on children's learning and development "can be used to map learning progressions in science. That is, one can describe the successively more sophisticated ways of thinking about a topic that can follow and build on one another as children learn about and investigate a topic over a broad span of time (e.g., 6 to 8 years)." Extensive supportive evidence has been generated in studies of learning progression and in of children's intellectual and social development. (Cite some studies here)

IV. Implications for assessment

The Framework and NGSS pose challenges and opportunities for both assessment for learning ("to help students while they are learning," Harlen 2013, p. 16) and assessment of learning ("to find out what they have learned at a particular time," Harlen 2013, loc. cit).

## ASSESSMENT CHALLENGES

In 1996 The National Science Education Standards or NSES (National Research Council, 1996) devoted a large section to assessment standards. A major feature of the Standards is alignment of assessments with all the components of the NSES vision of science education and also alignment of different kinds of assessments with their own purposes. It also indicated that different assessments should complement one another. Subsequently, a three-year study titled Knowing What Students Know: The Science and Design of Educational Assessment (National Research Council, 2001) laid out principles of a requisite assessment system to meets the needs of the science education envisaged in the NSES. It recommended an assessment system that is comprehensive (encompassing a range of assessment practices that provide a variety of evidence to support decision-making, formative and summative assessments, assessments that move students towards attainment of expectations including assessment of quality of instruction); coherent (models of learning underlying the assessments and different assessments used across the system are compatible); continuous (ongoing and seamlessly integrated with instruction); integrated ("carefully designed to fit into a larger, coherent educational system that provides resources and professional development to ensure that teachers have the capacity to do what is expected of them based on the standards in place"); and be of high quality ("meets relevant professional standards") (National Research Council, 2001). The Framework recommends this system approach to assessment processes.

Challenges posed by NGSS

NGSS is organized around science learning progressions and highlight performance expectations regarding all dimensions described in the Framework. The K-2 standard shown below illustrates this complexity, especially as it seems to demand not just assessment of achievement but also a presumption of adequate opportunity to learn.

All these desired characteristics of assessment pose enormous challenges to the assessment process. There needs to be assessment tasks that integrate all the dimensions, indicating position of a learner "along a sequence of progressively more complex understandings of a given core idea, and successively more sophisticated applications of practices and crosscutting concepts" Tasks will also have to be developed to assess students' abilities to make "connections between the different strands of disciplinary core ideas (e.g. using understandings about chemical interactions from physical science to explain phenomena in biological contexts)" (Wilson, 2013, See also APPENDIX 2).

One of the most challenging assessment issues of NGSS is assessment of children's "imperfect" but "productive" science conceptions. Duncan & Revit (2013) take a position that "it is important to differentiate between scientifically inaccurate ideas that are conceptually unproductive and understandings that are inaccurate, yet productive, and that can foster learning of more sophisticated understandings. The former are simply wrong; the latter can be seen as incomplete, overly simplistic, or tied to only a few limited contexts." It might be useful in most cases, though generally unattractive, to value productivity as long as the conception is not just plainly wrong.

## ASSESSMENT OPPORTUNITIES

NGSS performance expectations are endpoints of each grade level band and suggest upper boundaries for assessment. Teachers and curriculum developers, therefore, have to develop and implement science learning activities and performances that build to corresponding NGSS summative assessment boundaries. Even though both the Framework and NGSS are not a curriculum, by blending three dimensions of practices, core ideas, and crosscutting concepts, they convey a kind of vision of learning or pedagogy in science classrooms (cf. Quinn, Keller, & Moulding; Web Seminar on Framework, July 2011; Quinn & Schweingruber, 2011). It is for this reason that I think they offer best opportunity for development of explicit, robust classroom- and school-based and "classroom-blended" practices of assessment for learning.

**K-PS2-1 Motion and Stability: Forces and Interactions**

Students who demonstrate understanding can:

K-PS2-1. Plan and conduct an investigation to compare the effects of different strengths or different directions of pushes and pulls on the motion of an object. [Clarification Statement: Examples of pushes or pulls could include a string attached to an object being pulled, a person pushing an object, a person stopping a rolling ball, and two objects colliding and pushing on each other.] *[Assessment Boundary: Assessment is limited to different relative strengths or different directions, but not both at the same time. Assessment does not include non-contact pushes or pulls such as those produced by magnets.]*

The performance expectation above was developed using the following elements from the NRC document *A Framework for K-12 Science Education*:

| Science and Engineering Practices | Disciplinary Core Ideas | Crosscutting Concepts |
|---|---|---|
| **Planning and Carrying Out Investigations**<br>Planning and carrying out investigations to answer questions or test solutions to problems in K–2 builds on prior experiences and progresses to simple investigations, based on fair tests, which provide data to support explanations or design solutions.<br>• With guidance, plan and conduct an investigation in collaboration with peers.<br>- - - - - - - - - - - - - - - - - - - - - - -<br>**Connections to the Nature of Science**<br>**Scientific Investigations Use a Variety of Methods**<br>• Scientists use different ways to study the world. | **PS2.A: Forces and Motion**<br>1. Pushes and pulls can have different strengths and directions.<br>2. Pushing or pulling on an object can change the speed or direction of its motion and can start or stop it.<br>**PS2.B: Types of Interactions**<br>• When objects touch or collide, they push on one another and can change motion.<br>**PS3.C: Relationship Between Energy and Forces**<br>• A bigger push or pull makes things speed up or slow down more quickly. *(secondary)* | **Cause and Effect**<br>• Simple tests can be designed to gather evidence to support or refute student ideas about causes. |

COMMENTS

*We founded the Pequeños Científicos (Little Scientists) project in 2010. This is the Colombian inquiry-based science education project (ECBI, acronym in Spanish) that seeks to renew experimental sciences teaching and learning through Colombian primary schools. How was this done? Through teacher training and using a guided inquiry strategy.*

*My research question was: how do science learning outcomes of the students participating in ECBI programs compare with those of students who are not participating in such programs?*

*Who did I study? I had three schools that followed the ECBI strategy and two that did not. The subject I focused on was The Systems of the Human Body. I had 5th grade students, 365 of which participated in a written test and 147 who did performance tests. For this comparison I selected several schools after focusing on the results of their assessments in tests such as Enlace, which is used at the end of the school year. I considered these results because I did not want my ECBI schools to have excellent outcomes and my non-ECBI schools to perform poorly. I wanted both kinds of schools to be relatively similar for the*


Assessment for learning (or formative assessment) has been explained in many publications by Wynne Harlen, Page Keely (2011), the National Research Council (especially National Research Council, 2001), the National Association of Science Teachers, just to mention a few. Harlen's latest writing on this subject is a recent report for the Science Education Programme of the Global Network of Science Academies (IAP), entitled Assessment & Inquiry-based Science Education: Issues in Policy and Practice (Harlen, 2013). Her quotation of Dylan William clarifies formative assessment:

Practice in a classroom is formative to the extent that evidence about student achievement is elicited, interpreted and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence what was elicited. (William, 2009, p.9 quoted in Harlen, 2013, p. 17). She describes elements of formative assessment and gives illustrated descriptions, strategies, and vignettes of formative assessment in action. She discusses its efficacy of efficacy in the implementation of inquiry-based science, teachers' questions, feedback to students and into teaching, and student self- and peer-assessment. Her model of assessment for learning is an integral part of a grades 3-5 Inquiry Project led by Sue Doubler of TERC in Massachusetts. The project focuses on matter and is built around the dimensions and science learning progressions of the Framework. It is funded by the National Science Foundation.

Embedded assessments are usually an integral part of curriculum materials. These assessments are varied and can include "externally developed replacement units (curriculum materials + assessments); externally developed, item banks of tasks; portfolios, collections of work samples; tasks specified externally; moderation techniques (that) can be used to enhance the comparability of these assessments so they could support the desired inferences/comparisons needed for a monitoring purpose" (Wilson, 2013).

Over the years, a selected group of K-12 schools in New York State have dedicated themselves to school-wide use of embedded performance assessments such as collections of students' and teachers' work samples and learning records, 'moderation' techniques, and portfolios. They adopted these assessments as additional to district and state mandated monitoring and accountability assessments.

The Full Option Science System (popularly known as FOSS) at the Lawrence Hall of Science at Berkeley (California) has also developed an impressive embedded assessment system. It adapted already existing curriculum materials to include the multi-dimensions of the Framework in a classroom context.

It is possible and desirable to design and implement large-scale assessments based on the Framework and NGSS. DeBarger, Penuel, & Harris (2013) have reported on a large-scale assessment project as part of their middle school Project-Based Inquiry Science. "PBIS units align well to the core ideas in the Framework and student learning of content integrated with science practices" (p.4). Their assessment tasks focused on modeling and assessed students on "ability to construct a model and use the model to explain a phenomenon; ability to construct a model and use the model to make a prediction about a phenomenon; and ability to evaluate the quality of the model for explaining a phenomenon." Description of their work is at http://www.k12center.org/rsc/pdf/s2_debarger.pdf

At the national level, two major events designed to tackle the challenges posed by NGSS took place in 2013. During September 24-25, the Center for K–12 Assessment & Performance Management at The Educational Testing Services held an Invitational Research Symposium on Science Assessment in Washington, DC. The symposium explored "the skills and competencies called for in Next Generation Science Standards (NGSS) and the measurement challenges and opportunities they pose. Commissioned papers were presented on the design of both summative tasks and formative systems that adhere to the vision of instruction underlying the NGSS. In addition, the policy and practice work ahead, including the difficult trade-offs to be made in the designs of comprehensive science assessment systems, were discussed." Most of the presentations are available on line at http://www.k12center.org

The second development is an NRC study committee co-chaired by James W. Pellagrino (University of Illinois at Chicago) and Mark R. Wilson (University of California, Berkeley) that is completing its work on Developing Assessment of Science Proficiency in K-12. Wilson (2013) reported on the roles of three components of an assessment system that are being considered, namely: (a) some classroom-based assessments designed to support classroom activities/instruction; (b) some assessments designed to monitor science learning; (c) some process indicators to track opportunity to learn, exposure to high quality teaching and appropriate resources, and other factors that influence outcomes for students.

## CONCLUSION

The promise of the Framework and of its related NGSS is high and so also is the promise of some assessment developments underway. But for the promises to satisfy the desired characteristics outlined in Knowing What Students Know: The Science and Design of Educational Assessment, and in other existing and forthcoming publications there has to be dramatic shifts from the prevalent culture of teaching and learning science at the classroom level in the United States and from the culture of over-testing, to an educational culture that can be accurately described as inquiry-based and rich in authentic scientific practices and discourse. There will also need to be: "more connection and communication between teachers and assessment designers in different science areas; more coherent development of ideas over time -- not disconnected lessons; multiple experiences with each practice; more discourse-rich classrooms; more opportunities for classroom-based assessments that develop student knowledge and understanding and capacities to use what they know. (Quinn & Schweingruber, 2012)*

* Document for the Presentation

## COMMENTS

*Students were given different materials, paper towels, and water. They had to determine a strategy to tell us which paper towel absorbed the greatest amount of water. All of students had different tools and not all of them followed the same strategy. I was interested as much in the process as in the outcome, and focused on process as well as strategic literacy.*

*It is crucial to work on the reliability of evaluations. It took us a year and a half to develop the tests. If we are to measure and make decisions regarding these measurements we must have reliable, solid and fair tests.*

*Results: IBSE students always performed better than our control groups and this was true in all kinds of knowledge: declarative knowledge; procedural knowledge, schematic knowledge in sciences and strategic knowledge.*

# REFERENCES


Achieve (2013). Next Generation Standards: For States, By States. Washington, DC: National Academy Press.

Corcoran,T., Mosher, F.A. & Rogat, A. (2009). Learning Progressions in Science. Phildelphia: Consortium for Policy Research in Education.

DeBarger, A.H., Penuel, W.R., & Harris, C.J. (2013). Designing NGSS assessments: Evaluating the efficacy of curricular interventions. http://www.k12center.org/rsc/pdf/s2_debarger.pdf

Duncan, R.G. & Rivet, Anne E. (2013). Science learning progressions. Science, 339 (6118), 396-397.

Duschl, R. (2012). Preparing for NGSS: planning and carrying out investigations. http://learningcenter.nsta.org/products/symposia_seminars/NGSS/files/PreparingforNGSS--PlanningandCarryingOutInvestigations_10-9-2012.pdf

Harlen, W. (2013). Assessment & Inquiry-Based Science Education: Issues in Policy and Practice. Trieste, Italy: Global Network of Science Academies (IAP).

Harlen, W. & the IAP Working Group (2009). Inquiry-Based Science Education: An overview for educationalists. Trieste, Italy: Global Network of Science Academies (IAP).

Keely, P. (2011). Formative assessment in science:Using students' ideas to inform instruc-tions and promote learning. NSTA Web Seminars, May 2, 2011.

National Assessment of Educational Progress (2009). What does the NAEP science assess ment measure?. http://nces.ed.gov/nationsreportcard/science/whatmeasure.aspx

National Research Council (1996). National Science Education Standards. Washington, DC: National Academy Press.

National Research Council (1999). How People Learn: Bridging Research and Practice. Washington, DC: National Academy Press.

National Research Council (2001). Knowing What Students Know: The Science and Design of Education Assessment. Washington, DC: National Academy Press.

National Research Council (2007). Taking Science to School: Learning and Teaching Science in Grades K-8. Washington, DC: National Academy Press.

National Research Council (2012). A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas. Washington, DC: National Academy Press.

Quinn, Helen & Schweingruber, Heidi (2012). The on-going role of the K-12 science framework. http://www.k12center.org/rsc/pdf/s1_quinn%20.pdf National Research Council, 2001

Wilson, M.R. (2013). Update on the work of the NRC committee on science assessment. http://www.k12center.org/rsc/pdf/s2_wilson.pdf


# APPENDIX 1

**Three Dimensions of the Framework**
(Source: National Research Council, 2012, p3)

Scientific and Engineering Practices
1. Asking questions (for science) and defining problems (for engineering
2. Developing and using models
3. Planning and carrying out investigations
4. Analyzing and interpreting data
5. Using mathematics and information and computet technology
6. Developing explanations (for science) and designing solutions (for engineering)
7. Engaging in argument
8. Obtaining, evaluating, and communicating information

Crosscutting Concepts
1. Patterns
2. Cause and effect
3. Scale, proportion and quantity
4. Systems and system models
5. Energy and matter
6. Structure and function
7. Stability and change

Disciplinary Core Ideas
Physical Science
- PS1 Matter and its interactions
- PS2 Motion and Stability: forces and interactions
- PS3 Energy
- PS4 Waves and their application in technologies for information transfer

Disciplinary Core Ideas
Life science
- LS1 From Molecules to Organisms: Structures and Processes
- LS2 Ecosystems Interactions, Energy, and Dynamics
- LS3 Heredity:Inheritance and Variation of Traits
- LS4 Biological Evolution: Unity and Diversity

Disciplinary Core Ideas
Earth and Space Science
- ESS1 Earth Place in the Universe
- ESS2 Earth Systems
- ESS3 Earth and Human Activity

Disciplinary Core Ideas
Engineering and Technology
- ETS1 Engineering Design
- ETS2 Links Among Engineering, Technology, Science, and Society

# APPENDIX 2

**Challenges for Assessment**
(Source: Mark Wilson, ETS Symposium September, 2013)

- Developing rich assessment tasks that evaluate the intended practices and content and crosscutting concepts
- Having the platforms and resources to administer these kinds of tasks
- Scaling the tasks in the presence of multidimensionality and linkage across dimensions
- Scoring the tasks
- Developing informative, useful reports of test results
- Implementing "moderation" strategies in the U.S.
- Making use of information from classroom assessments for accountability purposes
- Making use of process indicators
- Assembling the components into a coherent system

# Project on the progress and assessment of the results in the State of Mexico

Raymundo Edgar Martínez Carbajal

The purpose of this presentation is to share the experience acquired in the development of the project for the assessment and follow up of the Program to Promote the Inquiry and Experience Based Science Education Systems in the State of Mexico, known as SEVIC in our entity and in some of the results achieved.

It is not easy to speak highly of the results achieved in the most populated state in the country, but considering that the purpose of the SEVIC program in the State of Mexico is to contribute to the scientific education of boys and girls, it is of utmost importance to highlight the information that helps us assert that the Program is moving forward although its coverage is somewhat limited.

Information is obtained through the assessment, follow up a continuous monitoring process of all actions undertaken. This process was designed for the very beginning to obtain information about the progress degree in the accomplishment of the established objectives and goals.
In order to contextualize the actions carried out, population and enrollment references of elementary education in the State of Mexico and its educational policy are provided. The assessment and follow up processes are highlighted, describing their stages, the amount of information for documenting the experience of all players and the challenges we face to assess the outcomes and to strengthen the Program.

## ABIDING BY THE STATE EDUCATIONAL POLICY

According to the numbers in the 2010 Population and Housing Census, the State of Mexico has 15'175,862 inhabitants, from which 4´353,914 are boys and girls from 0 to 14 years old, that is, 29.05% of the total population. The State Educational System provides elementary education to 3´396,157 students, from which 580,341are in the pre-school level, 1´961,234 in elementary school and 854,582 in secondary school.

In view of these numbers the State of Mexico faces each school cycle the challenge of improving the education provided to its population and in the Plans for Development of the State of Mexico there is full awareness of the need for students to improve in such subjects as Spanish, Mathematics and Science (PDEM, 2005-2011). All in view of the problems posed by the high number of people that live in the State of Mexico and their lack to reflect upon their actions and communicate their decisions based on scientific knowledge and evidence (PDEM, 2011-2017).

This situation demands the inclusion in every Plan for Development of programs and goals to improve education in the above mentioned subjects. In the current administration of Governor Eruviel Ávila Villegas the importance of strengthening the programs that contribute to the improvement of competencies such as reading, mathematics and science has been recognized in order to understand and solve the problems posed by living in a society (PDEM, 2011-2017).

Within this context and as of 2009 SEVIC was implemented in public elementary schools in the State of Mexico and the Agreement entered into by the Ministry of Public Education, the non-profit association called Innovation in Science Education (Innovación en la enseñanza de la ciencia) and the State Government was reasserted and thus continue with the effort of these three institutions to achieve the goal established for SEVIC in our state:
Contribute to the scientific education of boys and girls and to the improvement of their ability to learn, work in teams and actively and intelligently participate in the analysis and solution of problems.

## OPERATIONAL STRUCTURE

The SEVIC program is related to the state educational policy and as every governmental program is based on legal guidelines. It is also related to the administration´s structure taking care at the same time of the requirements and features of the SEVIC proposal, both in its pedagogical and operational aspects. With respect to operation a Trust was created for the management of its resources, also a Technical Committee for decision making with the participation of various administrative departments (management, planning, rules and regulations, science and technology, comptroller's office), the educational subsystems and INNOVEC.

The operational structure, with the participation of various departments involved, helps respond very precisely to the needs and regulatory guidelines and thus transparency, relevance and viability of all actions are guaranteed.
With respect to the growth of the Program there are two departments called State Coordination and Operational Coordination that report to the Office of the Assistant Secretary of Elementary and Normal Education of the Secretariat of Education in the State.

## QUANTITATIVE GROWTH

The support of the state educational policy and the fact of being anchored to the institutional administrative structure both help in the planning of financial and human resources requirements, facilities for the organization and advance of all actions, attachment to the institutional objectives and goals as well as the assessment of progress made and the achievement of all plans contemplated.The preceding has helped in the operation and the gradual growth of the Program as evidenced by the numbers shown in its four years of operation.

SEVIC started in the 2009-2010 school cycle with 44 elementary schools, 705 teachers and 25,855 students. In the 2012-2013 school cycle work was carried out in 161 schools, with 2,568 teachers and 90,322 students who worked in six Theme Units (Climate, Soils, Chemical Tests, Electrical Circuits and Ecosystems), that is, in four years the attention to students increased about 248.34% and the number of participating schools increased 266%.

> *"The support of the state educational policy and the fact of being anchored to the institutional administrative structure both help in the planning of financial and human resources requirements, facilities for the organization and advance of all actions"*

## ASSESSMENT AND FOLLOW UP PROJECT

As a Program attached to the administrative structure, its progress in the achievement of goals is measured through the pertaining administrative units in the institutional structure. In addition the Program has an assessment and follow up project designed when it started operating with the purpose of: Identify progress made in the achievement of the purpose and goals established for the Program, through a follow up process and the permanent monitoring of all actions taken during the aplication of the program and the other indicators that show the improvement of competencies of students who participate in SEVIC in our entity.

This assessment and follow up project is mainly focused on learning and considers both evaluation and assessment of the project. These processes have the support of the Operational Coordination staff who visit schools with the purpose of seeing how activities are carried out, identifying problems and supporting teachers in relation to the units' contents.

The process evaluates the way the Program is implemented, obtaining information about what we call related factors such as the organization of school dynamics and the participation of the various players involved (school authorities, teachers, students, parents and another school staff) in carrying out the activities.

Also, evidence with respect to pedagogical activities that take place within the classroom are collected. We expect to have sufficient information to assess achievements with respect to science education.

What do we understand for assessment and follow up? Assessment is the continuous process of adjusting all


actions in the Program to verify that what is being done has some favorable bearing on the achievement of the objective or the timely identification of obstructions in order to strengthen or reorient all actions in the search for the achievement of the proposed objectives.

## EXPANSION OF THE ASSESSMENT PROCESS

Since this is a continuous process, there are some development and achievement stages in the short, medium and long term, with specific purposes according to the stage and the achievement of the success in question.

The first stage (short term) is the recognition of the initial situation when the Program was implemented. The second (medium term, four to five years) is where we are right now and consists in obtaining information about the implementation of the Program in schools and classrooms. The third stage (long term, five to six years, when the first generation of SEVIC students graduate from the elementary schools where the Program started) will consist in the assessment of the impact, that is, to assess the degree of progress in the students´ scientific competencies.

Based on the opinion of 563 teachers of the 705 that started the Program (same opinions that were obtained by answering a questionnaire) and with references to specialized bibliography, the outlook of the work on science matters in the schools was as follows: lack of educational material, very modest command of scientific knowledge, no interest in scientific topics and the perception of the slight possibility of using scientific knowledge.

On the other hand, the results of the National Assessment of Achievement in School Centers (known as ENLACE in Spanish) in the year 2008 indicated that the highest percentage of elementary school students was in the basic achievement level, that is, they had minimum knowledge on the subject.

With respect to the professional profile of the teachers with whom this Program started, we can say that 88% have a degree in Elementary Education, 11% have attended Graduate School and 1% is specialized in Science. They are teachers at fully organized urban schools with three groups per school year and an average of 35 students in each group. This situation changes considerably when we face the growth that we have experienced.

The second stage consists of a process through which we watch and document the application of SEVIC in participating schools, using resources and tools for collecting information that implies obtaining data, information and evidence on the work of the Theme Units in schools and in the classroom.

About 20% of participating schools are helped each school cycle using information collection tools as a guide for watching what happens in schools, classrooms and school products, and questionnaires are also applied to school authorities, teachers and students.

## INFORMATION ARCHIVES

All this process has helped us collect information with which we have created three types of archives: 1) documentary, 2) photographic and 3) videotapes.

1. **Documentary archives**

a) Documentary archives consist of reports drawn up for each school cycle. To this moment we have five. A quantitative report describing the activities carried out and the goals reached is submitted in each document and also a qualitative report of strengths and weaknesses and a section for suggestions and ideas to reorient what needs to be reoriented, everything with references to organizational and pedagogical aspects.

b) Other information that is part of the documentary archives is the information provided by ENLACE for 2008 and 2012 when the assessment included the Science subject. This information helped us identify that in 2008 only 17% of the 44 schools that started with SEVIC were above the State average. However, in 2012 the percentage of these schools reached 50%, that is, there was an increase of 11 percentage points in the number of schools above the State average.
The information provided by ENLACE helped us in the follow up of the performance of schools with respect to results and there was a favorable variation in the average point between 2008 and 2012 as well as positive and negative differences. Schools that improved with respect to the first assessment were identified and also those that showed a decrease with respect to grades.
Even though it cannot be stated that these schools achieved those results due to SEVIC, as there is a large amount of elements that

COMMENTS

*Because of its fun and experimental structure, the SEVIC program makes students find school attractive. I believe that the program should also be expanded to high school where we have a serious desertion problem. According to some of our analyses, one of the main causes for school desertion is economic and the other is that students do not find school appealing. We therefore need to produce a new profile for teachers and, of course, new approaches to teaching science that appeal to youngsters.*

*“ENLACE helped us in the follow up of the behavior of schools with respect to results. (...) Schools that improved with respect to the first assessment were identified.(...) it is important to recognize that the Program (SEVIC) is part of the elements in the school context that allowed such favorable variation.”*

influence such results, it is important to recognize that the Program is part of the elements in the school context that allowed such favorable variation.

c) In the documentary archives we also have other very valuable material and that is the direct testimony of major players: boys and girls who participated in the Program. This material was obtained from the questionnaires answered by the students in their own handwriting. In such questionnaires we found expressions that say much about the interest and enjoyment SEVIC is leaving in boys and girls in comparison with the slight interest shown in Science topics when we started the Program.

2. **Photographic Archives**
 This archive is an evidence of the work carried out in schools, of the teachers and of boys and girls, of the image left as a testimony of what has been done, of the ways for organizing the classroom, of work styles, of students´ skills and of the many elements that may be identified by an observer interested in the image as a work footprint. Images of children and teachers working with the materials provided, are a rich source of analysis. For that reason the integration of this archive has been promoted as a source of information for assessment.

3. **Videotape Archives**
 In the same manner as the Photographic Archives, these videotapes are a source of valuable information and testimonies that can be analyzed in order to assess through oral expression and the moving image, the management of concepts by teachers and students, the increase of skills, attitudes and interest.
 This material has a great advantage: it can be repeatedly used in various moments and forums.

The archives created up to this moment are a major source of information that serve as reference to prepare the third stage of the process which is the assessment of the impact of results, that is, on the increase of competencies and skills that contribute to scientific education of students that participate in SEVIC.

For that reason it is necessary to carry out an analysis and an assessment of the achievements with respect to significant life experiences and the increase of competencies, considering elements such as understanding and managing information, participation in the educational processes, skills for searching and using information, to question and reach conclusions as well as the expression of collaboration attitudes and the sensitivity when facing problems related to scientific knowledge and daily life.

Another challenge that we are already facing is the growth towards preschool and secondary education levels, because our intention is to strengthen scientific education through elementary school.

This stage is the major challenge for the Program assessment. However, we are convinced that we have the inputs as well as the human and material resources to face this task and with that we are sure that we will be able to speak in the medium term of results in the education and development of scientific competencies because we already have evidence of those achievements that surely will have an impact on the academic education of students.

The Government of the State of Mexico, abiding by the guidelines in the Educational Reform already enacted and initiated by the President of Mexico, Mr., Enrique Peña Nieto, uses these Programs to reach the Minimum Normality that contemplates the current project of a Nation. Thank you very much.*

*"Another challenge that we are already facing is the growth towards preschool and secondary education levels, because our intention is to strengthen scientific education through elementary school."*

* Document for the Presentation

## COMMENTS

*Canada as many countries is making changes in its educational system. These reforms are long-term changes. According to the new frameworks the competencies to be implemented are problem solving, critical thinking and communication. I am very concerned about how the classroom teachers will handle the complexity of these frameworks. In addition these goals for innovation result in a considerable complexity in terms of assessment.*

*Particularly the Primary teachers I think of those that I worked with. They don´t have a background in science; they have to work very hard with understanding science concepts moving away from the delivery of facts to developing ideas with children that will eventually progress in science concepts. Then there is the diversity of purposes for assessment. Why is assessment carried out?*

*There are also crosscutting concepts that are common to all science disciplines.*
*Primary teachers are not science specialists. They teach language, maths, art, history, etc. They teach all subjects. How are they going to cope with these many dimensions of assessment and particularly when they are inserted in IBSE.*

# CONCLUSIONS

After analyzing the data from various research studies, it can be stated that teaching science with an inquiry-based approach makes a positive contribution to students learning and developing skills. In this regard, formative assessment has proven to be fundamental to identify learning outcomes and reinforce what was taught using this methodology. Thus, assessment becomes a fundamental part of the teaching and learning processes.

Nevertheless, in order to effectively incorporate assessments that contribute to education, we must shift from this culture of excessive evaluation aiming to measure results towards a culture in which students are considered individuals capable of actively participating in the development of their learning abilities and skills.

# Panel 4 Discussion

## EDUCATIONAL POLICY AND INQUIRY-BASED SCIENCE EDUCATION LEARNING ASSESSMENT


**Speakers:**
- Armando Loera Varela
- Lee Yee Cheong

**Panelists:**
- Ubaldo Ávila Ávila

**Moderator:**
- Arturo M. Fernández Pérez

# The gap in science teaching: comparison of the teaching of sciences and mathematics between three Latinamerican countries and participants from the TIMSS 1999 study

Armando Loera Varela

The following presentation shows the results of a study made by Heurística Educativa, at the request of the Inter-American Development Bank (IDB). The purpose of the study was to understand what teachers do once they close their classroom doors and review the most effective learning practices in different Latin American countries. We analyzed variables such as: duration of lessons, time devoted to group discussion, the level of experience-based pedagogy in experiments and procedures, among others.

The study determined that encouraging reflection, analysis and debate, while devoting an equal share of time to orthodox literacy and procedures brings about better outcomes in student learning. .

Good morning teachers. Thank you for focusing so much on the learning needs of children. For about 30 years now we have been carrying out studies at schools, and in the classrooms of several countries, basically Mexico. We have realized that although we still need more data, there is a certain difference between genders in their pedagogical practices. Female teachers are more watchful of the learning needs of their students than male teachers. Thank you very much for that.

Male teachers thank you very much because those same studies indicate that you are more concerned about concluding programs at any cost, even if your students do not fully understand them, because it is also very relevant to move forward at a good pace and in a timely manner. Thank you national and international guests for talking with us about the status of the learning and teaching of a subject as relevant as science.

## THE QUESTIONS

The Inter-American Development Bank (IDB) became aware of the various studies we'd done on what male and female teachers do in their classrooms by recording their work on video, and requested we conduct a qualitative study in several sites of interest with a representative population sample. The question from the Education Division of the IDB in Washington— specifically from the area of the teaching of mathematic sciences was, "What do teachers do once they close their classroom door?"This was because reforms come and go, curricula change, study programs change and texts change, but we really don't know what goes on inside classrooms.

The little knowledge we have about what happens is highly biased by self-reporting, (when you ask teachers, "How do you teach this?") or is very biased due to other factors. Videos are not the perfect way to observe reality as they carry their own biases, but they do complement information that is otherwise forgotten.

In Washington the question was divided into two: How different is what a Mexican teacher does relative to what teachers do in other Latin American countries? And, how different is what our Mexican and other Latin American teachers do from what is done by the teachers from developed countries that participated in the very famous TIMMS Video Study? TIMMS Video (Third International Mathematics and Science Study), was a video-based study that in some of the participating countries lead to the debate of the relevance of touching upon pedagogical aspects. In education and educational policy we largely debate about how much should be spent on education, how much autonomy should be afforded to schools, and how schools and teachers should be assessed, but we devote very little time to the hardest thing to change in education and that is how to teach. In other words, how are subjects that society wants to be taught at school, actually taught and learned.

I'm going to give you a number to visualized the dimension of the challenge ahead of us. The panel before this one made a call to do things differently. Now I'm going to show you where about are we in regard to one of the most important groups in our country.

Before we had the famous OECD PISA evaluation, there was en equally famous international test called TIMSS (Third International Mathematics and Science Study). This test was done in 1995. In a group of 45 counties, Argentina and Mexico were the only Latin American participants. TIMSS involved tests and surveys of associated factors, which are traditional tools in these kinds of international assessments. It was one of the first instruments of its kind to generate an international debate on matters such as scaling and establishing hierarchies for countries according to how much their children learn. By 1999, the name of this study became so famous they replaced the meaning of the "T" in "Third", to "Trends". Now the name is Trends in the Teaching of Mathematics and Science.

These studies generate great speculation. When the children of a country (usually an Asian country) exceed the scores of US children, the question is why? If Asians score better than Latin Americans, nobody asks why. The usual remarks are, "Well of course, they are rich, we are poor, there's not much we can do about it."The big excuse. Then come a series of illogical, inconsistent or weak statements about why things are so. Typically, direct blame (and politicians love to blame them) goes to whom? To the teachers. Everyone is to blame except the system, except the people in charge of educational policy, but the ones to blame are the teachers, the people standing in front of the children trying to get things to move forward. Somebody concerned over in this issue once said, Could it really be true that teachers have something to do with this? Can we specifically attribute to teachers work these huge differences? These gaps in student learning? What makes possible to learn so much in one country and so little in others? Can it be so simple? What if we conduct a study to find out what goes on inside classrooms? Thus, in 1995 the first video-based study got underway. It basically looked at mathematics classes. It took 4 years to complete the translations and analyze

them, so some of the first results only started to appear around 1999.

One of the most interesting books to emerge from this study was, *The Teaching Gap*. It was very focused on proving why Japanese teachers taught mathematics much better than the teachers in the United States or Germany, for example. This was highly interesting because for the first time in many years pedagogical aspects took center stage in the debate, and took us beyond economic and general strategy. The debate now focused on what happened inside classrooms.

Then people said, "Let's create another study with a much broader sample and other criteria to follow this up much closer." Once again, the United States became involved. The US National Statistics Center together with Japan has been financing these studies that used the same 1995 videos. New in this effort were countries such as Australia, the Czech Republic, Hong Kong, Holland, and Switzerland.

Our mandate from the IDB was, what if we replicated this study? What would happen if we applied it in Latin American countries? For that purpose we asked ourselves if there was a Latin American study that would help us accomplish what the TIMSS Video did.

TIMSS Video was based on the sample from TIMSS. We needed a Latin American study that we could take their sample from, and that sample had to be random and represent the countries that were going to participate. Fortunately we were able to use a study done by the regional office of UNESCO for Latin American and the Caribbean, headquartered in Santiago, which performed a study called Second Regional Comparative and Explanatory Study (SERCE). The study was done between 2005 and 2006 and focused on reading, mathematics and sciences in 3rd and 6th grade. TIMSS Video centered on 2nd grade secondary, while SERCE, looked at 3rd and 6th grades of elementary school.

## THE SAMPLE

The school sample used by SERCE is exactly the same one we used. What did SERCE show? First of all, that Cubans do extremely well in mathematics and sciences, much more so than other Latin-American countries. As measured by PERCE and SERCE (respectively the first and second Latin American studies by UNESCO), Cuba very well exceeds one and a half standard deviations. And this is more than anybody can explain. Obviously we researchers would have loved to go to Cuba, look inside the classrooms and understand what goes on in them. The problem is that the researchers who wanted to go to Cuba to visit classrooms had IDB money, and as you know that posed geopolitical limitations. Even though one of the main Cuban participants in SERCE gave his support, we were unable to visit classrooms in Cuba. Who ranked second in sciences and mathematics in this study? Uruguay and Nuevo León who tied. When SERCE was conducted, Doctor Reyes Támez asked UNESCO to run a specific study for Nuevo León, and that was very fortunate indeed. Nuevo León participated in mathematics and sciences, whereas Mexico as a country only participated in math, not science.

Colombia scored the average for the region. We were very much interested in Colombia, but all of this was happening in 2010, a year of presidential elections in the country and therefore nobody from the Colombian Education Ministry wanted to head the effort. As so often happens with these kinds of international studies we had to go wherever we were able to, to countries where we were able to negotiate entry. We were able to go to Dominican Republic, which came out last in the Latin American SERCE, and to Paraguay, which is nowhere near Colombia in terms of education, but nonetheless gained an interesting position. Still pending, however, is the need to visit Cuba to understand what happens there.

This study was not based on examinations, or on teacher reports, but fundamentally on lessons recorded on videotape.

> *"Can we specifically attribute to the work of the teachers these huge differences? These gaps in student learnings? What makes possible to learn so much in one country and so little in others? Can it be so simple? "*

We focused on 6th grade elementary school in Paraguay, Dominican Republic and Nuevo León. Fieldwork in all three sites took place between September and November 2010. The sample, as I have mentioned, was based upon SERCE. It had been 5 years since SERCE when we did the study, so not all the schools that had participated were still operating, except for the schools in Dominican Republic. In that case we found 100% of the schools that participated in the SERCE study; in Paraguay we found 84%, and in Nuevo León, 73%. Despite all of this we were able to obtain a sample size comparable to the one in the TIMSS study.

The 1995 TIMSS Video study has a general sample of 231 math and science lessons. The 1999 TIMSS Video study included 638. The size of the sample per country for TIMSS Video was set at 100. Switzerland's sample was 140 because of its linguistic and cultural, diversity. 100 lessons, however was the ideal size. We had 101 lessons from Nuevo León, 100 from Paraguay, and 96 from Dominican Republic. For a qualitative study the size of the sample was pretty important. For sciences, we collected 234 hours of teaching in all, from all three sites. Based upon the analysis made by TIMMS Video, our study focused on 57 dimensions of teaching and we made 137 comparisons in sciences.

We used different analysis strategy software. People from the National Pedagogy University (UPN) of Nuevo León, participated in recording the lessons. We installed two cameras. One focused on the teachers and the other on the students. We reflected on what happened in the classroom together with the teachers, and used a questionnaire with principals, teachers and so on. Also participating in the analyses of this information were people

## COMMENTS

*What we've learned about how to improve education systems after analyzing international experiences is that we must trust in teachers or fail. This is a universal lesson: the key lies in teachers' actions.*

*It is relevant to recognize the state of affairs just as they seem to be. It would be biased to present an in-depth description of the status of science teaching in Nuevo León as representative of what goes on in Mexico. Nuevo León always ranks high in academic performance regardless of the assessment used: ENLACE, INEE tests, or even the tests that used to be provided by the General Evaluation Office. We know well that Mexican children will do well or poorly in tests according to the social-economic status of their families and not necessarily according to what happens at school or in the classroom.*

*What happens in the classroom accounts for more or less 10 % of variance in academic achievement. What happens at school, depending on the study you want to consider, accounts for 15 to 22%, but* from the National Pedagogy University, doctors in sciences from the University of Sonora, and from the University of Puebla in physics, chemistry and biology.

## VARIABLES AND RESULTS

Let's move on to results. For this presentation we'll share what specialized international literature considers some of the critical dimensions of teaching; i.e. what really matters for good teaching. Obviously there is a conceptual bias in this way of looking at things, because it assumes that there is a kind of good and universal pedagogy. I will try to avoid speaking about good or bad teaching, and simply say, "that's how it was measured in our countries and these are the results". During our discussions later we can determine whether or not this has pedagogical significance or not. I also think that by comparing specific variables it is possible to gain a lot of insight as to what could be happening, but at the cost of a more holistic view of what can be happening in the classroom. For that reason, we also analyzed flow pedagogy that I will explain on another occasion, as I will not delve into it today.

There is a variable that is typically considered to be very relevant: **the duration of the lesson**. In its lessons Nuevo León showed us that the Mexican teachers in this sample teach more time on average that the ones in Dominican Republic. The former average 50 minutes teaching, while Dominicans average 39 minutes. On the other hand, those 50 minutes are less than the duration of lessons in the United States but similar to the time lessons last in other countries that participated in TIMSS Videos.

But most important is how time is used, how much is actually **effective time**. In Nuevo León effective time is the same as in Dominican Republic, but less than in all other countries that participated in TIMSS.

The predominant way of teaching sciences in TIMSS is group discussion, and then the students work at their tables individually or in teams. The teacher presents the topic and then has the youngsters work in teams. In Nuevo León this manner of teaching was used for the smallest percentage of time, lower than in any other TIMSS participating country, even below Paraguay, but greater however, than in Dominican Republic. TIMSS people usually say, "pay attention to what Japan does" because at one point Japan had the highest learning outcomes in the TIMSS. Japan, however, did not score high in this particular dimension. The Czech Republic had the highest frequency in this manner of teaching.

**What is taught in science classes?** In Nuevo León, just as in Dominican Republic and Paraguay life sciences, such as biology, predominate. There is little teaching on physics, chemistry or earth sciences. Paraguay was the exception because it showed a little more balance. In United States, Holland, Japan, Czech Republic, and Australia there is much more balance in the subjects taught.

**What is being taught as something scientific?** A major part of this has been discussed here. The people who worked on TIMSS Videos referred to facts as canonic knowledge. We would probably name them memorization data. As you can see, in Nuevo León most of the lesson is devoted to memorization data, contrary to the countries in TIMSS. Perhaps only the Czech Republic also showed a high rate of time for memorization data. Nevertheless, Dominican Republic has the highest rate of memorization. Remember that Dominican Republic ranks last in SERCE.

**How much pedagogy is based on experience, experiments, and procedures?** In Nuevo León barely 3% with respect to the entire sample. Even Dominican Republic scored higher than Nuevo León. Meanwhile Japan, the champion at the time in science teaching devoted 25% of lesson time to a procedures-based pedagogy.

**How relevant is what students are taught in class to their daily lives?** Curiously enough, among TIMSS participants, only 6% of the lessons taught in Japan were relevant in this manner, which was very low compared to Holland where that figure was 17%. Our teachers in Nuevo León scored 3%.

**Are scientific laws or theories announced in these contents?** Yes, more or less the same as in Japan. Japanese do not mention scientific laws or theories too often in their lessons, whereas Paraguay does so more than Japan.

**Are procedures or concepts discussed?** TIMSS shows that procedures are discussed much more than concepts. The only exception is Holland, which took one of the lowest positions in TIMSS. In our group of countries, the opposite is true. In Nuevo León, Dominican Republic and Paraguay, discussions are not about procedures, but about concepts, basically. Students are instructed to follow procedures, and very little to explore questions. Just as in Dominican Republic, in Nuevo León there is very little opportunity to learn sciences. In those countries that participated in TIMMS, for every class event students are asked questions, their is interest awakened, and they are asked to follow procedures in order to do interesting things using the scientific method: observations and data interpretation. In every TIMSS country this is predominant, whereas in our countries the level of difficulty of science teaching is very mediocre.

**What is more often generated: discussion or demonstration?** We observed that in our countries, students are asked to "try to guess what would happen if…" even without much insight, evidence nor coherence to make predictions. This is why, in our classrooms there is a higher rate of these events, than in TIMSS countries. In other words, there is a lot more prediction although very little effective data interpretation.

COMMENTS

*and the vast majority of the children who attend a Mexican public school are subject to high levels of social marginalization. And so, Mexican pedagogy cannot afford to be based upon (the notions of) theorists who do not consider poverty and cultural marginalization as a starting point, which is the case of many pedagogical theorists in teachers' education and at the Mexican National Pedagogical University itself.*

*We need to criticize and analyze what we are learning about how to teach from the reality of our public schools. Are we being taught to teach poor children effectively? I think not. And the irrefutable proof of this lies in our highly replicable public school: the least poor learn more, while the poorest learn the least. There is a great deal of evidence of this. We need to have enough pedagogical imagination to completely redesign our pedagogy in a way that fits the reality of our classrooms.*

**Student interaction with teachers and classmates**.This refers to whether there is an appropriate emotional climate when sciences are taught. Are they made interesting? Is there motivation? Cooperation? According to our results, we are at about 3% in our lessons.

**Homework.** Japanese are not that inclined to homework. They don't care much for having people do things at home. They prefer students to do everything at school, so their homework score is low. We have an average rate of homework, around 46%. Homework is not always good, unless it is reviewed and used for feedback. Otherwise, it's not that useful.

**Is there motivation to learn sciences? Are enticing activities used to generate enthusiasm?** In Paraguay a great part of their teacher training has to do with motivation. The rate was very high: 65% of the lessons have this characteristic, very much above the 21% seen in our classrooms and in Dominican Republic, where the score is barely 15%

**Are textbooks used?** Mexico does stand out from the rest of Latin America in this regard. Textbooks are used in 92% of the lessons, in contrast to 75% in Dominican Republic, and 18% in Paraguay. The manner in which textbooks guide not only students but teachers on the ideal sequence of activities and so on, is key and critical. For this reason it is critical to renovate, restructure and rethink textbooks.

What would happen if there were textbooks as interesting as *A Short History of Nearly Everything* by Bill Brayson? A review of astronomy, physics, chemistry, biology, archeology and almost everything everyone should know about sciences in a little compendium attractively narrated, that would provide teachers a basic understanding of science. Our science books, however, do not go too far in terms of motivation.

There was something TIMSS did not do, but we did. As science class analysts it was important for us to review the number of conceptual errors committed in those lessons. In Nuevo León, teachers made conceptual mistakes in 49.5% of the lessons, some of them very serious. Occasionally the personal opinion of the teacher was expressed more often than a science-based statement. In Dominican Republic there were errors in 68.8% of the lessons, and in Paraguay in 82%. Students also made mistakes, but a lot fewer than the teachers because they participated less and therefore had less chances of making mistakes, even though we know that making mistakes is pedagogically productive.

It should be pointed out that all of the teachers that participated in our study were very much committed to the quality of their teaching. They let us in. We weren't rejected even once and the only gift we could give them was the video we had recorded. The teachers were very appreciative of the videos because they were able to review and reflect upon their own pedagogical practices. I should mention that TIMSS paid the teachers in their sample 500 dollars to let them participate. Our teachers in Mexico, Paraguay and Dominican Republic received no payment at all. They all let us into their classrooms to record. They only thing we clarified was that these materials would remain anonymous and that we would not release any data that could hurt anybody because the investigation was scientific in nature.

For Nuevo León we were able to do something that couldn't be done in Dominican Republic nor Paraguay: associate one of the actions identified as important in classroom activities with the results of the ENLACE test in 2010. ENLACE had been applied in various subjects for 6th grade, and that year in particular it was geography.

*"We need to criticize and analyze what we are learning about how to teach from the reality of our public schools. Are we being taught to teach poor children effectively? I think not. And the irrefutable proof of this lies in our highly replicable public school: the least poor learn more, while the poorest learn the least. There is a great deal of evidence of this. We need to have enough pedagogical imagination to completely redesign our pedagogy in a way that fits the reality of our classrooms."*

What kind of teaching measurements had a positive link to the ENLACE results in sciences?

- **Effective class time.** Apparently less time wasted, results in better student outcomes in ENLACE.
- **Time devoted to discussion.** Criticizing, reflecting, analyzing, debate– everything we know about the essential nature of the scientific method is very, very useful.
- **Time devoted to demonstrations.** Proving to children that science is perfectly human, and not something for weird people that use hugely expensive things. Showing them activities that can be done inside the classroom in an inexpensive and simple way with a lot of student participation.
- **Equal time dedicated to memorization and procedures or practices.** Since ENLACE also examines data, we must not forget to teach data and not just methods. Science has also produced results that make it necessary to generate a conceptual change in our students regarding what a scientific vision of the world is vs. common opinion and other cultural sources of information students have available to them.
- **Time devoted to assess students.**
- **High level of difficult content.**
- **Group discussion focusing on individual deskwork**.
- **Presentations of scientific laws or theories.**

All of the foregoing were associated to positive results in the ENLACE tests.

I hope this sheds some light on what may be of interest to you. The Revista Latinoamericana de Estudios Educativos just published a paper of mine on results by kinds of schools; i.e. urban, rural, public, private and so on. You can visit http://cee.edu.mx/nuevaversion/publications/r2011-2020/r_texto/t_2013_2_02.pdf
This is our webpage: http://heuristicaeducativa.net/

On the IDB education page (http://www.iadb.org/es/banco-interamericano-de-desarrollo,2837.html) we have made available 7 reports we have prepared since last year with hundreds of pages containing analyses, which I hope will soon allow us to hold more informed debates on the status of the teaching of sciences.

Thank you.*

* Transcription.

# Educational policy and inquiry-based science education learning assessment

LEE YEE CHEONG

The Inter-Academic Panel (IAP) seeks UNESCO support for inquiry-based science education (IBSE) as one of the drivers for education in the 21st century. This requires that the representatives of member countries recognize the contributions of this teaching approach with actions such as helping education policy makers in different countries understand the value of the approach, and facilitate inquiry methods in science education in their school curricula and programs.

This debate must also insist upon the importance of not neglecting, but rather recognizing, strengthening and training an indispensable player in education: the teacher.

Thank you very much for the very kind introduction. I've never been in education. I have been practicing and engineering. And as IBSE is concerned I'm basically a promote, not a practitioner. Now, I will tell you [about] my life engagement with IBSE, then maybe you'll understand why I'm here. Actually I am one of the founders of the Academy of Sciences in Malaysia, which was also established in 1995. We joined the international community of the national academy of sciences of the world, under the umbrella of the Inter-Academy Panel. I was in charge of international relations in our academy; so I had the very good fortune of working very closely with Dr. Bruce Albert and also Professor Yves Quéré two of the outstanding and committed proponents of IBSE. I became one of the advisors.

In the early nineties from 1991 to 1993, I was promoting *La Main à la Pâte* in Malaysia and also in Southeast Asia, but then in 1993 I was appointed by the UN Secretary General Kofi Annan to be part of the study team called United Nations Millenium Project for the MDG, UN Millenium Development Goals.

So during my first early part of my engagement of IBSE, the object was to interest school children in science and mathematics so they will take up science and technology engineering careers in universities. At that time it was already quite plain that the pipeline for scientists and engineers is actually narrowing. But after my engagement with the MDG especially about poverty eradication, you know, and how to uplift the living standards of the very poor counties, I began to look at IBSE in a wider dimension of the poor world: the world of poverty and the world of hunger, and the world of diseases. But my experience in the United Nations also enabled me to work within the labyrinth of the very vast bureaucracy of the United Nations agencies.

Then in 2008, I was fortunate to be appointed by my government, the government of Malaysia, to be the chairman of the governing board of this estate, the International Science and Technology Innovation Center for South-South Cooperation, now, under the auspices of UNESCO. And this Center actually is totally funded by Malaysia, as Malaysia's contribution as a donor for promoting science and technology and innovation in less developed countries and poorer countries.

Now in this process, in the last five years I was in a very close engagement with UNESCO. After I got to know the UNESCO, the top hierarchy, from the director general downward. In the IBSE program of IAP, IAP has been trying to interest UNESCO to adopt IBSE as one of the main action plans for UNESCO, especially under the UNESCO decade for education, for sustainable development, which started in 2005 and is going to end next year, 2014. But the IAP had never been able to get UNESCO to adopt IBSE, as one of the primary drivers of science education.

Now, the reason actually as far as I'm concerned is very, very plain. The IAP as the umbrella body of scientific academies, they relate to the directorate of natural sciences in UNESCO. The IBSE to me, is not science, it's education. Because IBSE is both pedagogy and assessment, embedded in national science policies, and national curriculum. So this whole process is actually education. Am I right?

And so the big resources in UNESCO are in education. So we need actually now if we want IBSE to be adopted as one of the drivers of education in the 21st century inside UNESCO, we need to engage not the natural sciences directorate, but the education directorate. And that's why when Pierre Léna, of the IAP Scientific Education Program, the chairman of the global council asked me to take over, I said "My only target is to get UNESCO to adopt IBSE." And what I have been doing is trying; first thing to convince the director general of UNESCO that IBSE is the way forward.

Then the second thing is to try to shift the emphasis of IBSE from being under the natural sciences directorate to the education directorate. But this is not easy to do, because each directorate, they don't talk to one another; they are in silos. Ok, so to shift something that is within the competence of one directorate to the other is very, very difficult. Especially when the assistant director general of the natural sciences directorate and the assistant director general of UNESCO education directorate are two powerful ladies.

So what I did now is to go to the ground. If you can get the national ambassadors of the member countries of UNESCO to support this evidence-based education, or education by the scientific way, then from the ground there will be a clamor to get the UNESCO to adopt this as a driver. And of course, if the Director General is already very convinced and with the support of the member countries, it is my hope that within the next two years IBSE will be adopted as a main driver of education by UNESCO in the 21st century. So that is why I'm here, to try to listen to you and to get feedback.

Now, before I leave the podium, I just want to say a few things about what I have observed in listening to the speakers and also in the questions and answers yesterday and today.

I think that when we talk about science and engineering and technology, we need to be very transparent. I heard about the wonders of science and technology yesterday and today. It is all not wonderful. Why? Because we never talk about the weapons of mass destruction or the weapons of destruction that have been actually propelled by science and technology.

I remember two years ago my granddaughter, she watched a video of a village school in Pakistan being attacked by a drone; a missile fired from the air by the drone and two schoolteachers and children were killed. And she asked me "Why is this that somebody from the US can press a button and shoot a satellite to pass a signal to the drone, the drone fires

## COMMENTS

*A few moments ago there was a talk about what could differentiate to some Asian countries in their teaching compared to other countries. The most respected personality in Chinese history is Confucius, the great master in China. I think that Confucius influenced half of Asian countries. What I pick from Confucius is that among any company of three you are bound to find a teacher.*

*(…) So I think for education for students, for teachers, for human beings, we must say that we must start with humility as we can always learn from the others.*

**Lee Yee Cheong**

missiles and kills children, and some schoolteachers?" She asked me, "Is this technology very good? And why the newscaster says: 'This is only collateral damage'?" To her it was just like killing. But we have invented a system, a technology system that kills without human intervention. And yet, it is not called murder or killing, but the term collateral damage.
And you look around you, here in the Gulf of Mexico. There was a BP disaster, by drilling down into the earth. Now there are engineering, science and technology that allow you to drill deeper and deeper into the earth. Yet after BP paid 4-billion or more in compensation and all this, the drilling is starting again. And we are drilling and drilling for shale oil and gas. Everywhere in Europe we are drilling for shale oil and gas called fracking, and then causing earth tremors.
So I want to ask you, can Mother Earth suffer all this drilling without reaction and revulsion? And if Mother Earth reacts violently, who are going to suffer? So my message to science teachers and those promoting science education is please tell the bad side of science as much as you tell the good side of science! Because in this present dangerous situation of climate change [in which there is] still hunger and poverty in this world, it is science and technology that can provide a solution. But it is also science and technology in the wrong hands that can destroy our earth.


So please! The young people are not easily fooled by us telling them that engineering, science and technology are wonderful. Let them, by the evidence-based method, the scientific method of inquiry; then they themselves can decide what is right and what is wrong, and what is true and what is false. Don't try to tell them the story that everything is very wonderful by science and technology.

Yesterday and today I've been seeing videos of children, smiling. You know, the school children in every image they are smiling, excited and eyes brightened by the wonders of science. But there are children who die every year under five years of age. And 60% of this is due to malnutrition. Children in Sub-Saharan Africa are 16 times more likely to die before reaching the age of five than those in developed countries.

Now how do the mothers and parents feel about science education for their children in primary school when they don't even have the chance to live to enjoy primary school? And those who survive, you know, these five year olds, their life expectancy is really very low. So, what I would like everybody to think about is that for modern science education, evidence-based education, please think about those poor children who never even have the chance to go to school, and how do we trust the priority to give them science education, or the priority to let them have some education? I think that you would all agree with me that the priority is to just let them have some education, right?

And before I finish, I just now think somebody was mentioning about Asia, because I was concentrating on the translation. I think your comment was that Asians compete better in science and mathematics in international competition. I can tell you I know Singapore, it is a neighbor of Malaysia. Why do they do well in mathematics and science competitions, by the students? Because in Singapore graduates from university, the highest paying job is primary school teacher. It is higher than a graduate engineer, higher than a graduate lawyer, higher than a graduate doctor. And because the primary schoolteacher is paid the highest, one percentile of the top of the graduates from universities apply for a job as primary schoolteachers, not just science schoolteachers, but primary schoolteacher of any discipline. You can imagine that under that school system, looking after the primary schoolteachers, they produce very bright students.

So I think that the message from science education should be, not about students, but about the teachers; that we need to look after the teachers.

And I hope that this conference, although it is an international conference, it's actually a Mexican organized conference, I would humbly recommend that you send a delegation of education administrators to Singapore. You may not reach the top that primary schoolteachers are paid, better than any others, but at least a trend, the trend of appreciating the value of the teacher. And in this, I don't mean that in Singapore they only look after the primary schoolteachers. They provide a careers development paths to secondary schoolteachers to be lecturers in the university and also to be researchers. It is a whole system.

So, I hope that Mexico will send a delegation to Singapore and learn from Singapore. Thank you very much!*

* Transcription.

## COMMENTS

*In this very young 21st century, schools must continue to be the articulating arm of community development. Schools are supported by what is experienced in the classroom, by a teacher's commitment to stimulate his students. Teachers who promote scientific literacy in every one of these classrooms are shaping individuals with critical attitudes capable of making proposals and who will become analytical, reflexive men and women.*

*In the month of December 2009 in Zacatecas, we established our first agreement so that science could become an experience in our classrooms. During the 2012-2013 school year, 57-thousand students shared this experience thanks to the commitment of the teachers to articulate this strategy, and the fact that 1250 of them are attentive to these kinds of education policies.*

*Follow-up and assessment are accomplished through tools that measure the efficacy and viability of the program in different contexts. In my state, SEVIC is used as a modality in pre-school and primary school, and in special education. We are implementing this modality in the Centros de Atention Múltiple*

# CONCLUSIONS

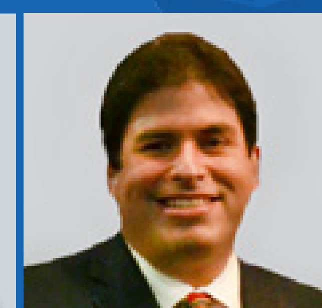
It is crucial to consider teachers' actions a key factor to improve national educational systems. It is necessary to insist on proper teacher training, as well as on providing the context that will allow them to do their work to satisfaction. Such a context includes schools with proper facilities, sufficient materials and dignified salaries.

However, we cannot forget that the effectiveness of an educational system is not entirely the responsibility of the teachers. It is fundamental to design educational policies that consider conceptual frameworks and pedagogical strategies that understand and address the social realities that exist in the classrooms.

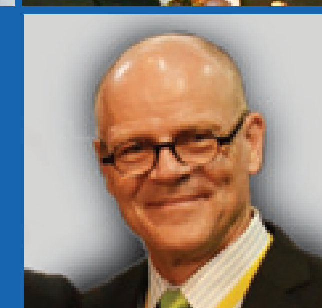
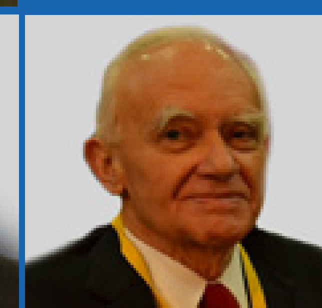
Different fronts, i.e. communities, states, countries, international bodies face the challenge of promoting elements that will make it possible to articulate and implement mechanisms at every level of educational systems in order to offer students scientific education that will develop their capacities for reflection, analysis, teamwork and problem solving. All of these skills are necessary in this 21st century.

# Panel 5
## Discussion

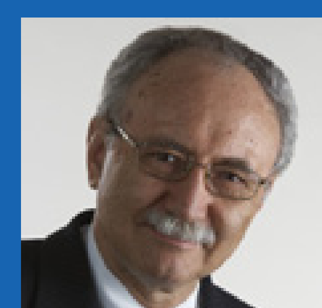
## TRENDS IN INQUIRY-BASED SCIENCE EDUCATION AND ITS ASSESSMENT MODELS

**Speakers:**
- Jon K. Price
- Daniel Alcázar Román

**Panelists:**
- Anders Hedberg
- Leopoldo Rodríguez



**Moderator:**
- Guillermo Fernández de la Garza

# Trends in inquiry-based science education and its assessment models.
## *Lessons Learned from Global Science Education Initiatives*

Ph. D. Jon K. Price

Intel Corporation is the world's largest semiconductor chip maker, developing advanced integrated digital technology, primarily integrated circuits, for industries such as computing and communications.

Because the Science, Technology, Engineering and Mathematics (STEM) skills are the skills we seek every day as we hire the best and the brightest engineers and scientists to operate our factories and conduct our research, we recognize the challenges facing global economies needing a knowledgeable workforce.

At Intel Corporation, we know that a chronic shortage of STEM students threatens each countries opportunity for innovation and economic development. We believe having a vibrant economy sustained by quality education, a skilled workforce, and innovation is key. Through our education initiatives and investments, Intel is helping communities build local capacity while preparing the next generation of innovators.

### INTEL'S EDUCATION STRATEGY

For more than four decades, Intel has made education the primary focus of our strategic philanthropic activity. We invest more than $100 million annually in programs that promote STEM education, encourage women and girls to seek careers in technology, foster and celebrate innovation and entrepreneurship among the best and brightest young students in the world and help teachers to incorporate best practices in math, science and the effective use of technology in their work.

We work in coalitions with other high-tech companies to support technology access, development and implementation of K-12 teacher professional development, mathematics and science content and curriculum, as well as assessments to support initiatives that develop 21st century skills, such as critical thinking, collaboration and creativity. These are skills students need to be the innovators of tomorrow.

Science competitions are at the center of the Intel education programs. Our goal in sponsoring the competitions is to identify and celebrate talented young scientists. Through them, we inspire younger students to follow in their footsteps and communities to invest in high quality science education. The Intel International Science and Engineering Fair, also a program of Society for Science & the Public, is the world's largest pre-college science fair competition. Each year, approximately 7 million high school students around the globe develop original research projects and present their work at local science fairs with the hope of winning, with winners who progress to regional, national, and ultimately the Intel's ISEF in the US.

To address the need for teacher professional development that moves beyond applications, Intel created a program designed to train classroom teachers to integrate technology into their lessons to promote problem solving, critical thinking and collaboration skills among their students into their existing curriculum. To date, the Intel Teach Program, has trained over ten million teachers in more than 70 countries worldwide.

In addition to program and infrastructure investments, Intel has also invested in exploratory research and rigorous program evaluation to establish and sustain continuous improvement of these educational products and activities. The research and evaluation compiled for this purpose has not only enabled the improvements of the program development efforts, but now also comprises a comprehensive body of evidence that demonstrates program impact (Price, Light, Michalchik, 2011). As a result of these efforts, critical evidence has emerged that may inform other evaluation activities designed to measure impact related to ICT in education in terms that extend beyond logistical measures and student assessment.

### ASSESSMENT FOR TEACHING AND LEARNING


To meet the demand in providing sufficient training for the teachers' capacity to use new educational technologies to support student learning, the Intel® Teach Program was developed as a professional development course that helps teachers integrate technology into their lessons and promoting students' problem-solving, critical thinking, and collaboration skills. With more than 10 million teachers trained in over 70 countries, Intel Teach is the largest, most successful program of its kind. These programs are designed to provide teachers with the knowledge and skills to develop 21st Century skills with their students, encourage project based, collaborative and personalized learning and effectively integrate information and communication technologies as critical tools into the classroom. Through a review of evaluation data and reports collected from studies of these successful professional development courses delivered over ten years and across multiple countries, the contextual factors regarding how formative assessment strategies can be effectively integrated into classrooms is provided .

The synthesis of the research and evaluation following the Intel Teach program and use of classroom assessments suggests tools and strategies share three important traits that in different degrees: 1) high quality teacher-designed assessments provide insight on what and how students are learning in time for teachers to modify or personalize instruction; 2) they allow teachers to assess a broader range of skills and abilities in addition to content recall; and 3) these assessments give students new roles in the assessment process that can make assessment itself a learning experience and deepen student engagement in content (Price, Pierson, & Light, 2011).

1) Provide Insight on Student Learning so Teachers Can Modify Instruction: Because many of these assessment tools and strategies are formative in nature, the information garnered from their implementation can be used to immediately inform teachers' instructional decisions. For example, information garnered from portfolios can help teachers evaluate the effectiveness of their

own instruction while helping them make informed decisions about future lessons. The implementation of portfolio assessments stimulates student self-reflection providing valuable feedback to both students and teachers, which in turn can be used to inform the teaching and learning processes. When employing the peer assessment strategy, if students and teachers assess a student differently it can open up productive dialogue to discuss student learning needs and goal creation (J. Ross, 2006). The teacher can then use that information to structure the following lesson around the needs and goals of those students. Whether taking a pre and post survey poll or asking multiple-choice questions to reveal student's subtle misunderstandings and misconceptions, a Student Response System (SRS) allows teachers to take a quick snapshot of where his or her teachers are on a learning continuum and devise the appropriate strategies to take them to the next level. As teachers become more aware of their students' interests, needs, strengths and weaknesses, they are better positioned to modify their instructional strategies and content focus to help maximize student learning.

2) Assess Broader Range of Skills and Abilities: Traditional forms of assessment like multiple-choice, fill in the blank, and true/false, privilege memorization and recall skills that demand only a low level of cognitive effort (Dikli, 2003; Shepard, et al., 1995). The assessment tools and strategies outlined in this paper provide more robust means to measure higher order thinking skills and complex problem solving abilities (Palm, 2008). Strategies such as performance based assessment (PBA) and portfolios, take into account multiple measures of achievement, and rely on multiple sources of evidence, moving beyond the standardized examinations most commonly used for school accountability (Shepard, et al., 1995; Wood, Darling-Hammond, Neill, & Roschewski, 2007). Self-and peer-assessment both teach and assess a broader range of life skills like self-reflection, collaboration, and communication. As a tool to measure student learning, rubrics allow teachers to measure multiple dimensions of learning rather than just content knowledge, and to provide a more detailed assessment of each student's abilities instead of just a number or percent correct.
3) Give Students New Roles in the Assessment Process that Make Assessment a Learning Experience: In contrast to the traditional teacher-designed, teacher-administered, teacher-graded tests, this cadre of assessments involves students throughout the assessing process. Involving students in the creation of assessment criteria, the diagnosis of their strengths and weaknesses, and the monitoring of their own learning, transfers the locus of instruction from the teacher to his or her students (Nunes, 2004). For example, the most successful rubrics involve students in the creation of the evaluation criteria. This creates buy-in, increases engagement, and fosters a deeper commitment to the learning process. In the assembly of a portfolio, students not only get to decide which work is graded, they have the opportunity reflect up and evaluate the quality of those submissions. This type of involvement fosters meta-cognition, active participation, and ultimately puts students at the center of the learning process (McMillan & Hearn, 2008). During peer-assessment students are asked to be the actual evaluator offering feedback and suggestions on how to improve their classmates' work. When created collaboratively, many of these assessments enable teachers and students to interact in a way that blurs the roles in the teaching and learning process (Barootchi & Keshavarz, 2002). When students are part of the assessment process they are more likely to "take charge" of their own learning process and products and will be more likely to want to make improvements on future work (Sweet, 1993).




Our observations show that classroom assessment strategies can work within the contextual challenges of developing countries– large class size, short lesson periods, and limited resources. Based on our studies, we recommend six classroom formative assessment strategies that are a good place to start: (Price, Pierson, & Light, 2011).

- Rubrics,
- Performance-based assessments (PBAs),
- Portfolios,
- Student self-assessment,
- Peer-assessment,
- Student response systems (SRS)

## IMPROVING HANDS-ON INQUIRY BASED LEARNING THROUGH SCIENCE COMPETITIONS

Each year, approximately 7 million high school students around the globe develop original research projects and present their work at local science competitions with the hope of making it to the Intel International Science and Engineering Fair, a program of Society for Science & the Public. The top projects —1,600 winners of local, regional, state, and national competitions—are invited to participate in a week-long celebration of science, technology, engineering, and math. At the event, the young innovators share ideas, showcase cutting-edge research, and compete for more than USD 4 million in awards and scholarships. At Intel ISEF, awards are based on students' abilities to tackle challenging scientific questions, use authentic research practices, and create solutions for the problems of tomorrow (Intel, 2013).

Intel's three objectives related to the Intel ISEF program are: to encourage and reward excellence in student-based research; to motivate students to pursue science, math and engineering careers; and to promote inquiry and project-based science teaching and learning in the schools. To understand the impact of the competition, data was sought to evaluate the achievement of these goals, and to improve the program. Online self-report surveys, focus groups and interview responses were obtained from four distinct groups of important participants in Intel ISEF: teachers, students, regional fair directors, and judges to explore the perspectives and experiences of students, teachers, and judges. Inclusion of the regional fair directors provided information for program improvement. The results of the study suggest that all three goals are being met, for a complete report on the findings see: Intel International Science and Engineering Fair 2005 Evaluation Report, Rillero, Zambo & Haas, 2005.

Moving beyond the stated goals of Intel ISEF, survey responses and interviews of student finalists suggest that participation in the science completions at preliminary levels helped them improve their science project, and therefore inquiry based learning in two ways. First, they were challenged by questions and received suggestions that enabled them to improve the project by helping narrow the focus and making them more concise, and second they had the opportunity to improve their presentation skills and the defense of their ideas. One important factor found in the means of receiving feedback may also be the result of the student's relationship with mentors, both in-school and out-of-school. The first relationship identified as "Very Important/Important" being Parent/Guardian influence (73%), Current Teachers (66.9%), and (external) Mentors (63.3%). These relationships were found to be a more significant factor to successful science fair completion than access to outside laboratories (Rillero, Zambo & Haas, 2005).

*"Intel's three objectives related to the Intel ISEF program are: to encourage and reward excellence in student-based research; to motivate students to pursue science, math and engineering careers; and to promote inquiry and project-based science teaching and learning in the schools."*

Considering this study as a way to assess inquiry based science education, the study also explored if Intel ISEF had an effect on teaching strategies and assess student learning. About two-thirds of the teacher respondents agreed or strongly agreed that their involvement with Intel ISEF had changed the way they teach and 89.1% agreed or strongly agreed that external competitions had a positive impact on their teaching. At a school wide level, however, the effects are not as strong. When faced with the possibility of removing Intel ISEF but not the affiliated fairs, 47.5% agreed or strongly agreed that it would affect science or mathematics programs at their school. When asked if all science fairs were gone, 63% agreed or strongly agreed that it would change their school's programs. When survey data was compared to teacher interview responses, the impact of Intel ISEF on teaching strategy results from application of project based learning in ways meaningful to the student most often in a formal research class, resources and support for the teachers to apply these new strategies, and non-traditional environments such as research courses, clubs, or informal after-school time to support student inquiry. In addition, when the teachers were asked to indicate how difficult a set of tasks were for the students as they completed their science fair projects, data related challenges were identified within the top problems students encounter: statistical analysis (35.8%), data analysis (15%), getting accurate measurements (9.6%) (Rillero, Zambo & Haas, 2005).

Lastly, perhaps the most impactful assessment strategy applied in the science competition is the analysis of the judges perceptions. Imagine if all exams were oral exams & the questioning strategy allowed the examiner to dig deeper in to the level of understanding. This is the benefit of the judge. This level of exploration provides true insight into the depth of inquiry exhibited by the student and illustrates the most successful science fair projects are those in which the participants demonstrate critical thinking skills; this is what makes scientists and engineers unique. For the purpose of this study, a set of 16 factors were ranked in order of importance, with (1) as Not Important – (4) as Very Important.

COMMENTS

*In the United States we have this somewhat complex document: The Next Generation Science Standards that addresses a specific need, because in different states we have different standards that seem more like dictionaries full of terms than a system that works better. We therefore still have a long way to go, but this book is a step forward in explaining that it is important not only to learn concepts, but also procedures and to develop skills and competencies.*

❖

*Innovation will get us ahead. We are all technology consumers, but we want to participate in the development of these technologies. We want people who not only consume, but also produce.*

**Daniel Alcázar**

Tables 1 -4 provide the judges results (Rillero, Zambo & Haas, 2005).

| Factor | Mean (SD) | Very Important | Very Important or Important |
|---|---|---|---|
| Methodology | 3.49 (.61) | 53.8% | 93.8% |
| Quality of Data | 3.43 (.63) | 49.7% | 91.5% |
| Data Analysis | 3.37 (.66) | 45.7% | 89.9% |
| Hypothesis | 3.17 (.74) | 35.4% | 81.8% |
| Problem Selected | 3.09 (.80) | 33.1% | 77.6% |
| Theoretical Framework | 2.94 (.72) | 20.8% | 73.7% |
| Literature Review | 2.73 (.77) | 15.2% | 60.2% |

Table 1. Seven of the 16 factors corresponded to aspects of the scientific method and the quality of the data collected.

| Factor | Mean (SD) | Very Important | Very Important or Important |
|---|---|---|---|
| Findings expanded scientific knowledge | 3.03 (.83) | 32.0% | 73.1% |
| Potential of results to be used by others. | 2.82 (.92) | 27.1% | 61.6% |

Table 2. Two of the 16 factors related to the use of the research beyond the project.

| Factor | Mean (SD) | Very Important | Very Important or Important |
|---|---|---|---|
| Oral Presentation | 3.18 (.76) | 37.6% | 80.3% |
| Visual Display | 2.65 (.77) | 12.9% | 56.7% |
| Written Report | 2.60 (.82) | 12.2% | 56.5% |
| English Language Skills | 2.14 (.91) | 7.0% | 34.4% |

Table 3. Four of the 16 factors related to the presentation of the project.

| Factor | Mean (SD) | Very Important | Very Important or Important |
|---|---|---|---|
| Access to outside mentors | 2.41 (.90) | 11.2% | 46.3% |
| Access to outside research labs | 2.30 (.94) | 11.3% | 40.7% |
| One of more parents working in scientific or technical fields. | 1.85 (.92) | 5.0% | 25.1% |

Table 4. Three of the 16 factors related to the availability of outside assistance.

When asked how most of the Intel ISEF finalists could make their projects better, the judges responses can best be categorized as:

- Improving the Methodology.
  By providing a larger sample size or increasing the number of trials conducted, the students could have a greater understanding of the impact and validate their claims.
- Increase Clarity of Presentation
  By providing their project data and findings in an easy to read format that includes the most relevant data the students can then describe their depth of knowledge through the project notebook and discussion. With a clearly defined problem that is the focus of their research, followed by describing carefully planned experiment and data analyses, the student can exhibit the critical thinking skills that are viewed as most important.
- Communication Skills
  Finally, the ability to answer questions and talk informally about the project is more important than English language skills or an impressive display. Students need to illustrate how their critical thinking goes beyond finding a solution; it is about finding ways to apply knowledge to similar situations, not being afraid of failure, exploration for deep understanding of how things work and a willingness to admit that learning is a continual process and not an end state. Even a failed experiment is not a failure but a great learning experience and a springboard for asking another question or looking at the problem from another angle.

## FINDINGS FROM SUCCESSFUL STEM SCHOOLS

As part of Intel's most recent exploration into STEM education, the research team, in association with SRI International, conducted case studies of five schools that provide rich, rigorous science and math learning experiences for students. A report was generated to describe the schools, the challenges each faced, and how they were able to build effective learning environments. The significance of these schools' accomplishments can best be understood within the current interest in STEM instruction and learning. The report, Creating New Opportunities for STEM Learning: Insights from Case Studies of 5 Schools, can be found at www.intel.com/education/evidenceofimpact.

The case study schools were selected for their diversity of location and context. They were all public schools, serving students representative of their area and selected due to their record of improvement.

Researchers from SRI and Intel visited each of the five schools for 2–3 days. During these visits, interviews were conducted with school leaders, district staff, teachers, and other school staff members working with technology or STEM curriculum. In addition, at each school a parent focus group, a student focus group, and at least three classroom observations also conducted. These activities provided insights into the context, challenges, and strategies implemented in the schools.

Across the successes documented in these very different schools were consistent underlying themes. First, all the schools broke from the norm in some way and created a new vision and culture of education. At George Hall Elementary, the break from business as usual included a new school leader, replacement of nearly all the teachers, a new curriculum, and drastically different school practices. In the restructuring, school staff even worked to clean up the school building. Change at this scale was difficult for the community at first, but community support was built over time.

*"(...) the ability to answer questions and talk informally about the project is more important than English language skills or an impressive display. Students need to illustrate how their critical thinking goes beyond finding a solution; it is about finding ways to apply knowledge to similar situations, not being afraid of failure(...)"*

Although the changes at George Hall are dramatic, all five schools made bold changes of some kind. Byron High School flipped its classroom and homework periods, redefining the approach to student learning with a full openness to the digital revolution in education. Preston adopted a growth mind-set as a teaching staff, made every class advanced, and then developed ways of scaffolding and differentiating learning for the newly empowered students. An important factor within each school was that teachers and administration broke free of their old habits and thoughts, charted a new course, and harnessed their courage to make bold changes.

Second, the schools all provided professional learning communities (PLCs) and professional development opportunities shaped and directed by teachers. Teachers are often expected to implement curricular or pedagogical changes designed by experts or someone outside the school, and professional development or PLCs are used to support teachers in making these changes. At these schools, the opposite was true. For example, at Byron teachers were provided the tools, time, and training to form innovative PLCs where members were encouraged and supported in attempting collaborative, measured, systematic implementation of radically new and ultimately quite

## COMMENTS

*The complexity of evaluation has been illustrated throughout this conference. It involves many more aspects than those we usually approach when we discuss the design of SEVIC because evaluation necessarily involves students. Along the way we have very important decisions to make. For example: what do we exactly want to assess in a student? In the SEVIC program? At school? Do we want o assess learning outcomes or competencies?*

*This poses a very significant dilemma.What do we want to emphasize? Do we want to trigger more engineering or scientific vocations and eventually produce better engineers and better scientists? Or do we want a broad part of populations, even the whole student population of a country to enjoy a level playing ground of capacities that will allow them to be more competitive in the 21st century?*

*If I were in the United States I'd perhaps seek to have more and better engineers and scientists. But, would that be my priority for Mexico? Very likely not, perhaps because of my involvement in the challenge that innovation represents for Mexico."*

effective instructional methods. Preston math teachers had a similar story. Their principal pushed for a culture where it was expected that all students could learn at the highest levels, but he then empowered teams to explore and craft their own innovative means of achieving that goal. Professional development and PLCs, then, were not a method of integrating someone else's' school improvement but a way to equip teachers with the skills and opportunity to envision and implement their own classroom improvements.

Third, all the schools provided creative and thoughtful out-of-school time learning opportunities for students. Preston offered plentiful opportunities to engage in hands-on science activities such as wildlife habitat restoration and created an elective system, with the final period of the day dedicated to help students discover and pursue their passions. Taking advantage of its New York location, MS223 provides students with the cultural opportunities routinely enjoyed by wealthier families, such as outings to museums and Broadway shows, in order to provide students a broader and more enticing worldview. Byron's use of the flipped classroom, where students are introduced to content outside class and work on problems during class time, is a creative redefining of in-school and out-of-school time. In all five schools, teachers and administrators took a broad view of learning, beyond classrooms and standards, to encompass the motivations, interests, and passions of their students.

Fourth, the schools devised sensitive methods to meet each student's unique learning needs. MS 223 hired a full-time math coach to support the use, interpretation, and teaching responses to formative assessments. In addition, it developed a mutually beneficial relationship with a teacher training program that now puts additional teaching professionals in high-need classrooms to maximize opportunities for small-group learning at a pace and with methods more customized to each students' needs. George Hall also used a data-driven process, using a technology system that provides instant analysis of student skills, ensuring that what students learn is appropriately leveled and based on need. After pushing all students into an accelerated math program, Preston needed to develop a robust system for ensuring that all students learn. Its differentiation of instruction now comes through flexible student groupings where, each week, students across grades and classrooms address different learners' needs lesson by lesson. In each case, the schools are finding ways to meet students where they are and support them in their personal learning trajectory.

Finally, all the schools sought out and leveraged nearby resources. MS223 partnered with Teach For America to bring more adults into the math classes, Yale University to provide arts education, and local college fraternities, sororities, and sports teams to glamorize college for their students. Farmington View conducted an asset-mapping project and used it to identify local resources such as the Jackson Bottom Wetlands where students participate in wildlife preservation research. In addition, the mapping of resources extended to the school's parent population, where any adult with valuable expertise was tapped to lead an afterschool club or activity. Byron High school has used technology to gather free web resources and tools to build and deliver learning content, first with math and now with many subjects. Each school has found creative ways to forge external partnerships to expand the learning resources available to students.

The five common themes among these case studies demonstrate that the schools in this study were recognized for and built their success on a large view of their educational mission. They took bold and brazen approaches, they supported ground-up changes from teachers, worked to maximize both in-school and out-of-school learning opportunities, and brought in external resources to support their efforts. Each of the five schools provides a story of success within its unique setting and context, yet this bold, broad, and resourceful perspective runs through each case.

## CHALLENGES

Classroom instruction is a complex enterprise that occurs at the intersection of teachers, students, and texts within the surrounding classroom, school, and community environments. Effective education reform and sustained policy-based (macro) initiatives to enhance equity and excellence must be designed and understood at the classroom (micro) level and secondarily at the school (meso) level, (Scheuermann, et al., 2009 & Price & Roth, 2010). In each of the cases discussed, the assessment strategies applied and the procedures developed were designed and implemented with the individual student in mind. As such, audience and magnitude of scale becomes the prominent barrier to effective dissemination of findings (CDC, 1999).

When primary stakeholders are immediate, such as local content teams, standards for data collection may favor qualitative observations and interactions with participants. Whereas, when governments and expenses are involved, the standard for credible data often requires more rigorous experimental designs. Regardless, either of the approaches can be improved by using multiple procedures for collecting data. The most common source of data is through the program participants themselves via surveys, interviews, focus groups or observations. However, data is also available through document review. In educational settings, such artifacts may include administrative records and teacher or student products or portfolios (Price, Roth, McAllister, 2011).

If we were to revisit the idea that meaningful assessment of student knowledge and understanding through the strategies presented here, and if we acknowledge that a change in the roles in the assessment process that can make assessment itself a learning experience and deepen student engagement, then we must also acknowledge a need for a change in the practices and teacher/learner relationships to enable new measurement strategies.

COMMENTS

*PISA tells us that this competency (it clearly qualifies this notion as a competency) is an individual's ability to become effectively involved in a process in which two or more agents are attempting to solve a problem by sharing the understanding and effort required to reach a solution.*

*If this continues, by 2017 we will be assessing these competencies actively and routinely. The challenge is great, so we must prepare ourselves.*

Current research illustrates the value of micro level assessment strategies. Fortunately, large scale, macro level international collaborative projects such as the Assessment and Teaching of 21st Century Skills, (ATC21s.org), and the New Pedagogies for Deeper Learning, (newpedagogies.org), initiatives are attempting to understand, define and initiate such changes in how relationships between students and teachers are structured, in how teaching and learning is practiced, and in how learning is measured. In each, the opportunities available through technology can be seen.

## CONCLUSION

Education has been Intel's primary philanthropic focus for decades. Intel has over 200 programs in more than 70 countries that provide professional development for teachers, support and celebrate student achievement in science, technology, engineering, and math, and bridge the digital divide with relevant, local online content for educators, students and parents. Our experience in education worldwide has informed our understanding of the need for higher standards and more rigorous assessments for our students. Often, we hear that our education systems are broken. The issue is not so much that they are broken as it is that they were built at a time and for functions that are no longer critical, and measured in ways that are no longer meaningful. Today, we need far more people with analytical skills to pursue innovation in academia, in industry, in government. Today, our young people need a far better grasp of technology and science simply to live in this increasingly complex and rapidly changing world inundated with data, climate changes, and revolutionary advances in medical science. Today, we need systems of measurement and assessment that move away from an emphasis on subject knowledge, and move towards understanding skills and attitudes and the increasingly important competencies of critical thinking. As a result, as technology becomes ever more complicated, there is a better understanding to approaches to student learning through a process of constructing and developing knowledge and the meaning of learning in their own lives and assessment strategies must reflect this.*

* Document for the Presentation.



# REFERENCES


Barootchi, N., & Keshavarz, M. H. (2002). Assessment of achievement through portfolios and teacher-made tests. Educational Researcher, 44(3), 279-288.

Centers for Disease Control and Prevention. (1999). Framework for program evaluation in public health. MMWR 1999;48(No. RR-11)

Dikli, S. (2003). Assessment at a distance: Traditional vs. Alternative Assessments. The Turkish Online Journal of Educational Technology, 2(3).

McMillan, J. H., & Hearn, J. (2008). Student Self-Assessment: The Key to Stronger Student Motivation and Higher Achievement. Educational Horizons, 87(1), 40-49.

Nunes, A. (2004). Portfolios in the EFL classroom: disclosing an informed practice. English Language Teachers Journal, 58(4).

Palm, T. (2008). Performance Assessment and Authentic Assessment: A Conceptual Analy sis of the Literature. Practical Assessment, Research & Evaluation, 13(4). Retrieved http://pareonline.net/getvn.asp?v=13&n=4

Price, J., Light, D., Michalchik, V. (2011). Ten Years of Evaluation within Intel Education Initiatives. Retrieved from: https://www-ssl.intel.com/content/www/us/en/education/evaluations/ten-years-of-evaluation-within-intel-education-initiatives.html .

Price, J., Pierson, E., & Light, D. (2011). Using Classroom Assessment to Promote 21st Century Learning in Emerging Market Countries. Proceedings of The Associa tion for the Advancement of Computing in Education: Global Learn Asia Pacific, Melbourne, Australia.

Price, J, & Roth, M. (2010). Evaluating Effective Teaching and Learning within Complex Levels of Interaction. Proceedings of the Global Learn Asia Pacific 2010 - Global Conference on Learning and Technology. Penang, Malaysia.

Price, J., Roth, M., & McAllister, B. (2011). The Framework for Program Evaluation: Implementation Strategy of the Intel Education Initiatives. Unpublished Manuscript. Intel Corporation.

Rillero, P., Zambo, R., & Haas, N., ( 2005). Intel International Science and Engineering Fair 2005 Evaluation Report. Retrieved from: http://www.intel.com/content /www/us/en/education/evaluations/intel-isef.html

Ross, J. (2006). The Reliability, Validity, and Utility of Self-Assessment. Practical Assess- ment, Research & Evaluation, 11, 1-13. Retrieved from http://pareonline.net/pdf/v11n10.pdf

Rossier, S., House, A., Price, J. & Michalchik, V. (2013). Creating New Opportunities for STEM Learning: Insights from Case Studies of 5 Schools. Retrieved from: http://www.intel.com/content/www/us/en/education/evaluations/other-programs.html

Scheuermann, F., Kikis, K., & Villalba, E. (2009). A Framework for Understanding and Evaluating the Impact of Information and Communication Technologies In Education. http://crell.jrc.ec.europa.eu/ICTeducation/090428%20-%20CRELL%20-%20Frameworkconsiderations.doc

Shepard, L. A., Flexer, R. J., Hiebert, E. H., Mario, S. F., Mayfield, V., & Weston, T. J. (1995). Effects of Introducing Classroom Performance Assessments on Student Learning. Boulder, CO: National Center for Research on Evaluation, Standards and Student Testing, University of Colorado at Boulder.

Sweet, D. (1993). Student Portfolios: Classroom Uses. Retrieved http://www2.ed.gov/pubs/OR/ConsumerGuides/ .

Wood, G. H., Darling-Hammond, L., Neill, M., & Roschewski, P. (2007). Refocusing Account ability: Using Local Performance Assessments to Enhance Teaching and Learning for Higher Order Skills.

# Teaching and assessing inquiry-based science: challenges and opportunities from the United States perspective

M. Ed. Daniel Alcázar-Roman

Assessing the transfer of Scientific and Engineering Practices may require instruments that engage students in performance tasks that reflect how science and engineering is done in the real world. The evidence that is gathered and tracked by performance tasks and over the student's scientific education may provide a better picture of student progress and learning gaps.

Today the United States government and the private sector are making investments in the improvement of Science, Technology, Engineering, and Mathematics (STEM) education. Some early outcomes of this focus are the development of new standards, and large scale studies sponsored by the Investing in Innovation (i3) Fund. Both of these are making contributions to a better understanding of how to implement science reform across the nation. High quality inquiry-based science standards are triggerring changes in the way formative and summative assessment is implemented in schools.

## SCIENCE STANDARDS

Traditionally, in the United States, science learning has been guided by education standards that separate content from process skills. In most cases only the content is assessed using selective response assessment tools that are given once a year in selected grades (5th, 8th, and some high school courses). The need for science education reform has led to the development of the Next Generation Science Standards, NGSS (NRC, 2012). These new standards present performance expectations that integrate core ideas, and inquiry skills for instruction and assessment. The NGSS seeks to help students understand the nature of scientific knowledge or inquiry by

### Scientific and Engineering Practices (NRC, 2013)

- Asking questions (for science) and defining problems (for engineering)
- Developing and using models
- Planning and carrying out investigations
- Analyzing and interpreting data
- Using mathematics and computational thinking
- Constructing explanations (for science) and designing solutions (for engineering)
- Engaging in argument from evidence
- Obtaining, evaluating, and communicating information

### Crosscutting Concepts (NRC, 2013)

- Patterns. Observed patterns of forms and events guide organization and classification, and they prompt questions about relationships and the factors that influence them.
- Cause and effect: Mechanism and explanation. Events have causes, sometimes simple, sometimes multifaceted. A major activity of science is investigating and explaining causal relationships and the mechanisms by which they are mediated. Such mechanisms can then be tested across given contexts and used to predict and explain events in new contexts.
- Scale, proportion, and quantity. In considering phenomena, it is critical to recognize what is relevant at different measures of size, time, and energy and to recognize how changes in scale, proportion, or quantity affect a system's structure or performance.
- Systems and system models. Defining the system under study—specifying its boundaries and making explicit a model of that system—provides tools for understanding and testing ideas that are applicable throughout science and engineering.
- Energy and matter: Flows, cycles, and conservation. Tracking fluxes of energy and matter into, out of, and within systems helps one understand the systems' possibilities and limitations.
- Structure and function. The way in which an object or living thing is shaped and its substructure determine many of its properties and functions.
- Stability and change. For natural and built systems alike, conditions of stability and determinants of rates of change or evolution of a system are critical elements of study.

### Student Expectations

These are two examples of how the student expectations in the NGSS seek to integrate content and the nature of science into statements that guide instruction and set clear assessment guidelines:

Grade 5 Physical Science Student Expectation (5-PS1-2): Students who demonstrate understanding can: Measure and graph quantities to provide evidence that regardless of the type of change that occurs when heating, cooling, or mixing substances, the total weight of matter is conserved. [Clarification Statement: Examples of reactions or changes could include phase changes, dissolving, and mixing that form new substances.] [Assessment Boundary: Assessment does not include distinguishing mass and weight.]

This student expectation combines the core ideas of structure and properties of matter and chemical reactions, the Practice of using mathematics and computational thinking, and the Crosscutting Concept of scale, proportion, and quantity.

Middle School Life Science Student Expectation (MS-LS3-2): Students who demonstrate understanding can: Develop and use a model to describe why asexual reproduction results in offspring with identical genetic information and sexual reproduction results in offspring with genetic variation. (Clarification Statement: Emphasis is on using models such as Punnett squares, diagrams, and simulations to describe the cause and effect relationship of gene transmission from parent(s) to offspring and resulting genetic variation).

This student expectation combines the core ideas of growth and development of organisms, inheritance of traits, and variation of traits, the Practice of developing and using models, and the cause and effect.

## DEVELOPING HIGH QUALITY ASSESSMENTS FOR INQUIRY BASED LEARNING

In order to obtain an accurate understanding of a student science performance, science assessment instruments focus on science learning in all domains: core ideas (concepts), Practices, and Crosscutting Concepts. Furthermore, this evidence should be gathered in the context of real-life scientific tasks. In the past year, Alexandria City Public Schools (ACPS) has rolled out a system to monitor the science achievement and progress of students through transfer tasks. These assessments are vertically aligned and allow the school system to monitor students' performance in specific measurement topics from Kindergarten to grade 12. Additionally, the National Assessment of Educational Progress (NAEP) has developed tools to assess scientific inquiry and provides good examples of alternative ways of assessing students.

### Transfer Tasks

Performance tasks or transfer tasks allow educators to capture the kind of information that would best serve as evidence of scientific transfer. In these assessments tools, students are asked to demonstrate what they learned in similar yet different contexts. For example, if students have been learning about interdependent relationships in ecosystems by exploring the school pond, the transfer task would ask them to create a model of a terrestrial ecosystem. Crosscutting Concepts are useful while creating transfer tasks because they allow the assessment of concepts using different contexts. For instance, if an instructional unit focused on investigating a core idea through the lens of systems and system models, a transfer task might ask a student to explore the structure and function of the parts of the same system.

Transfer tasks also ask students to demonstrate a particular scientific or engineering Practice. For example a transfer task that requires students to explore the Crosscutting Concept of scale, proportion, and quantity, would also assess the Practice of analyzing and interpreting data.

### Sample Grade 5 Transfer Task:

You will plan investigations to help Caroline better understand changes in matter.

Lisa found her sister Caroline stirring her iced tea madly. Caroline said that she was trying to make sugar melt. Lisa told Caroline that what she was actually doing was dissolving the sugar, not melting it. Caroline was confused. She said that she knew that lots of things melt such as ice cream, ice cubes, candy in her mouth, sugar in hot tea, and chocolate. After thinking about it some more Caroline, concluded that melting and dissolving were the same thing.

Your job is to design a series of investigations that Caroline could follow so that she can learn the difference between melting and dissolving.

Be sure to include:

- Clear directions for the activities (include diagrams and pictures).
- A list of materials and tools that are needed.
- Sample graphic organizers that Lisa can use to record her observations.
- Teacher notes with the expected results for each activity and the patterns that would be observed during the melting or dissolving.

This Transfer Task targets the Practice of planning and carrying out investigations and the Crosscutting concept of analyzing patterns.

### Measurement Topics and Rubrics

At ACPS we determined that our K-12 science measurement topics would match the NGSS Practices and Crosscutting Concepts.

For each of those measurement topics we developed rubrics in grade-bands (Kindergarten-2nd, 3rd-5th, 6th-8th, and 9th-12th). Each transfer task targets two measurement topics (one Practice and one Crosscutting Concept). The corresponding rubrics provide performance descriptors with a four-point scale. It also provides teachers with language to use as feedback for their students. The scores for each measurement topic are tracked from year to year. Transfer Task scores are also used as part of a larger comprehensive assessments plan to monitor students' progress throughout the school-year.

Table 1-Tracking Student Performance on Science Transfer Tasks by Measurement Topic

| | Measurement Topics | Grade Level | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | K | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| Scientific and Engineering Practices | Asking Scientific Questions and Framing Engineering Problems | | | | | | | | | | | | | |
| | Developing and Using Models and Visual Representations | | | | | | | | | | | | | |
| | Planning and Carrying Out Investigations | | | | | | | | | | | | | |
| | Analyzing and Interpreting Data | | | | | | | | | | | | | |
| | Using Computational Thinking and Mathematical Reasoning | | | | | | | | | | | | | |
| | Constructing Scientific Explanations and Designing Engineering Solutions | | | | | | | | | | | | | |
| | Using Evidence to Engage in and Support Arguments | | | | | | | | | | | | | |
| | Analyzing, Evaluating, and Communicating Information | | | | | | | | | | | | | |
| Crosscutting Concepts | Analyzing Patterns | | | | | | | | | | | | | |
| | Analyzing and Explaining Causal Relationships | | | | | | | | | | | | | |
| | Assessing the Impact of Scale, Proportion, and Quantity | | | | | | | | | | | | | |
| | Investigating Systems and System Models | | | | | | | | | | | | | |
| | Analyzing Flows, Cycles, and Conservation of Energy and Matter | | | | | | | | | | | | | |
| | Exploring Structure and Function | | | | | | | | | | | | | |
| | Investigating Stability and Change | | | | | | | | | | | | | |

Table 2- Sample Rubric

**Measurement Topic: Developing and Using Models and Visual Representations**

Modeling in K–2 builds on prior experiences and progresses to include using and developing models (i.e., diagram, drawing, physical replica, diorama, dramatization, or storyboard) that represent concrete events or design solutions.

| Score | Performance Indicators |
|---|---|
| **4. Advanced** | • Your response shows that you are highly effective at distinguishing between a model and the actual object, process, and/or events the model represents.<br>• Your response shows that you are completely accurate when comparing models to identify common features and differences.<br>• Your response shows that you are completely accurate when developing and/or using a model to represent amounts, relationships, relative scales (bigger, smaller), and/or patterns in the natural and designed world(s).<br>• Your response shows that you consistently develop a simple model based on evidence to represent a proposed object or tool. |
| **3. Proficient** | • Your response shows that you are somewhat effective at distinguishing between a model and the actual object, process, and/or events the model represents.<br>• Your response shows that you are generally accurate when comparing models to identify common features and differences.<br>• Your response shows that you are generally accurate when developing and/or using a model to represent amounts, relationships, relative scales (bigger, smaller), and/or patterns in the natural and designed world(s).<br>• Your response shows that you generally develop a simple model based on evidence to represent a proposed object or tool. |
| **2. Developing** | • Your response shows that you are highly effective at distinguishing between a model and the actual object, process, and/or events the model represents.<br>• Your response shows that you are inaccurate when comparing models to identify common features and differences.<br>• Your response shows that you are inaccurate when developing and/or using a model to represent amounts, relationships, relative scales (bigger, smaller), and/or patterns in the natural and designed world(s).<br>• Your response shows that you occasionally develop a simple model based on evidence to represent a proposed object or tool. |
| **1. Basic** | • Your response would benefit from showing the similarities and differences between a model and the actual object, process, and/or events the model represents.<br>• Your response would benefit from developing and/or using a model to represent amounts, relationships, relative scales (bigger, smaller), and/or patterns in the natural and designed world(s).<br>• Your response would benefit from developing a simple model based on evidence to represent a proposed object or tool. |



## COMMENTS

*Innovation drives investment in Research and Development, which in turn encourages a STEM-based economy, and in turn that leads to an increased national interest in STEM education. Continuing from there, effective STEM education, in turn, now breeds innovation, so now we have a positive cycle. One thing leads to another and basically we are spiraling upwards here. So, innovation is part of the key themes of what I would like to talk about.*

*Innovation is hampered by a widening workforce skills gap. Here I'll talk about a few data points in the US. There are international data available as well. Today there are 2.7 million jobs open due to the lack of qualified applicants. There is only a little bit less in terms of the number of unemployed individuals who often cannot get a job because they are not sufficiently qualified. That's the skills gap. It's a mismatch between the needs and the access to the talent to meet those needs.*

*Looking into the future, it is a little scary actually, because the number of jobs that will require a STEM college degree (and in fact not just a STEM college degree, but a STEM understanding, STEM literacy if you wish) will continue to grow. The prediction is that within the next ten years the STEM-requiring jobs will exceed those that do not require STEM jobs and probably several-fold. In 2018, the prediction is that the number of STEM jobs will exceed 8-million in the US alone.*

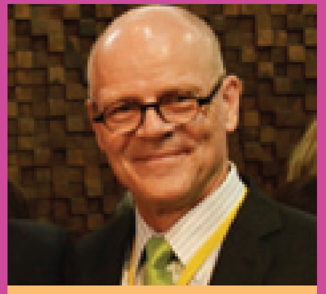
*When we talk about skills and competencies, obviously we are all aware and comfortable with problem-solving skills, critical thinking skills, and other things that are part of the 21st century curricula. But there is a whole host of skills and competencies that are now growing out, becoming important and rare to find because today's workplace lives by these.*

*There is a group in Washington called "The STEM-Connectors". We have an innovation task force that I am on and we are now working. We are trying to understand for ourselves, because many of us are from the applied sciences STEM sector, what is STEM 1.0?*
*If we define 1.0 of STEM as what we have today, it's obviously not sufficient. We can't live with 1.0. We have to begin to develop a STEM 2.0 concept. The only way you can do that is to understand what the future job tasks are going to demand in terms of what those skills are. And I can't rank them to you or give them to you. They're somewhere on the list we just studied. What can and what must the education sector provide? Then we have to imagine that to get moving from where we are today we have to step up on competencies platforms, capability platforms; one at a time up to the point where we have achieved an armamentarium or a portfolio for every person that goes through this, that gives him or her the skills that they need to be competitive in today's environment of the workplace.*

*Given the known outcomes of inquiry based science education, we understand and know that it leads to better critical thinking and problem solving, how can we now explore how young students can be encouraged to innovate in the classroom?*

❖

*If we understand what innovation is, if students do it already, it's not going to be heavy lifting. The heavy lifting is going to be to define to ourselves what is it that tells us that they are now innovating?*
*We need to develop then the practices and the tools to evaluate innovative or creative skills, and you know how to do this better than I do. We don't talk enough to one another. Intel's (...) represented here, engineer Rodríguez represents the business community and many others of you do as well, but we've all got to do this better.*



**Anders Hedberg**


### LARGE SCALE PERFORMANCE ASSESSMENTS

The NAEP assessment includes interactive computer and hands-on tasks designed to assess how well students can perform scientific investigations, draw valid conclusions, and explain their results. As a part of the 2009 science assessment, a new generation of hands-on tasks was administered during which students worked with lab materials and other equipment to perform experiments. While hands-on tasks have been used in NAEP since the 1990s, these new tasks present students with more open-ended scenarios that require a deeper level of planning, analysis, and synthesis. For the first time, the NAEP science assessment also included interactive computer tasks in science. While performing the interactive computer and hands-on tasks, students manipulate objects and perform actual experiments, offering us richer data on how students respond to scientific challenges. Several key discoveries were observed (NCES, 2011):

- Students were successful on parts of investigations that involved limited sets of data and making straightforward observations of that data.
- Students were challenged by parts of investigations that contained more variables to manipulate or involved strategic decision making to collect appropriate data.
- The percentage of students who could select correct conclusions from an investigation was higher than for those students who could select correct conclusions and also explain their results.

Sample questions, scoring criteria, and student responses can be found on the The Nation's Report Card website: http://goo.gl/BdtqOt *

* Document for the Presentation.

## COMMENTS

*We need to communicate more between us. When I hear the comments that have been made in this Conference, the idea of a Symphony Orchestra comes to my head: we can be wonderful instruments, but if each one plays its own melody, all is a disaster. Normally we see individual offices, individual institutions, individual organizations and individual members inside the organizations working each person by his side. There must be much more collaboration between all.*

❖

*Policies should support curricula, the use of technologies applied to education, the development of assessments that take into account these new teaching approaches and its implications. All of this should be coordinated. There are many organizations who are interested in walking in this sense. We must find a way to articulate them, to organize them. In addition, not only asking them to commit, we must commit ourselves as individuals, as members of our organizations, as parents.*

**John K. Price**

# REFERENCES


National Research Council. A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas. Washington, DC: The National Academies Press, 2012.

National Research Council. Next Generation Science Standards: For States, By States. Washington, DC: The National Academies Press, 2013, Appendix F http://goo.gl/wLebFf

National Research Council. Next Generation Science Standards: For States, By States. Washington, DC: The National Academies Press, 2013, Appendix G http://goo.gl/flECjc

Wiggins, G. & McTighe, J.(2005). Understanding by Design (Expanded 2nd Ed. USA). Alexandria, Va.: Association for Supervision and Curriculum Development.

National Center for Education Statistics. Science in Action: Hands-On and Interactive Computer Tasks From the 2009 Science Assessment. U.S. Department of Education, 2011, http://goo.gl/Q0miba


# CONCLUSIONS

**High-tech companies, those that base their growth on new product research and development require a growing number of human resources well trained in science, engineering, technology and mathematics. This challenges nations to promote the right education policies and curricula to address their social and economic realities as well as the demands these companies have for labor markets.**

**This is fundamental to detonate and maintain innovation processes whose repercussions will translate into dynamic national economies. Achieving this requires seamless communication among sectors and institutions as well as coordinated efforts targeting common goals.**

**This is urgent, because the 21st century needs individuals whose competencies go beyond critical thinking and teamwork. It requires a "Generation 2.0," so to speak, with competencies formed by the development of systemic thinking, negotiation skills, openness to multiple cultures and remote collaboration to name a few. The challenge is a great one.**

# OTHER ACTIVITIES

## Welcome Dinner

November 13th, 2013.
Interactive Museum of Economy, Mexico City.


INNOVEC thanked the support and participation of speakers, panelists, and non profit organizations, enterprises and partners, that made possible the "Seventh International Conference on Inquiry Based Science Education in Elementary Schools, Science Learning Assessment: Trends and Challenges"

During the ceremony, INNOVEC honored PhD. Rosa Devés and PhD. Jorge Allende, of the University of Chile and PhD. León Olivé, of the National Autonomous University of Mexico for their valuable contributions to Inquiry Based Science Education (IBSE).

## Workshops

November 13th, 2013.

Centro Cultural Tlatelolco, UNAM, Mexico City.

Trilogy of workshops for members of the State teams involved in the implementation of the SEVIC Program in Mexico.

**Workshop 1. Formative assessment in the classroom: instruments and supporting tools for teachers**

Instructor: **María Figueroa**

This workshop was designed for members of the State teams involved in the implementation of the SEVIC Program in Mexico. Purpose: To develop skills to design and implement strategies for formative assessment.


**Workshop 2. Importance of IBSE evaluation system. Primary Connections the Australian case.**

Instructor: **Shelley Peers**

The workshop was directed mainly to members of the State teams involved in the implementation of the SEVIC Program in Mexico. It was intended that the participants reflect on the need to implement evaluation systems of the SEVIC Program in their States, taking as a reference the experience carried out in Australia with the Primary Connections Program.

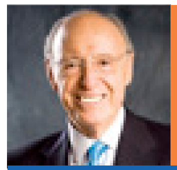
**Workshop 3. Use of additional resources to support secondary teachers on inquiry -based science teaching programs**

Instructor: **Anne Goube**

The use of the DVD "Learning Science and Technology in Secondary School" was promoted during the workshop as an additional resource for IBSE programs. The production of the international version of the DVD (Spanish and English) was in charge of INNOVEC and is an example of the collaborative work that INNOVEC has done with the French Programme *La main à the pate.*

# Biographies

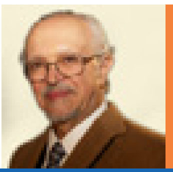
## Inauguration and Keynote Conference



### Jaime Lomelín Guillén

President of the Board of INNOVEC. Mexico.

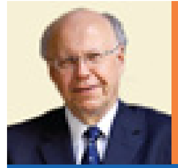
Lomelín holds a Bachelor of Science degree in Chemical Engineering from the National Autonomous University of Mexico (UNAM) and undertook business administration studies at the University of Wisconsin (1958-1959) as well as the AD2 Program at the Instituto Panamericano de Alta Dirección de Empresas (IPADE) in 1975 and also the Stanford Executive Program at Stanford University (1984).
Mr Lomelín is a member of the Board of Trustees of ITAM and an alternate director of Palacio de Hierro, Grupo Nacional Provincial, Valores Mexicanos Casa de Bolsa and is also Chairman of the Board of Trustees of the Chemistry School at the UNAM as well as a member of the UNAM Foundation. Mr Lomelín is President of the Councils of the non-profit organisations: INNOVEC, the Mexican Mining Chamber, and the Council of Economic Development of the State of Zacatecas.



### Mario Molina

Vice President of the Board of INNOVEC. Mexico.

Molina holds a Chemical Engineer degree (1965) from the National Autonomous University of Mexico (UNAM), a Postgraduate degree (1967) from the University of Freiburg, Germany, and a Ph.D. in Physical Chemistry (1972) from the University of California, Berkeley.
He is a pioneer and one of the main scientists in the world dedicated to atmospheric chemistry. He was co-author with Frank Sherwood Rowland, of the 1974 original article predicting the depletion of the ozone layer as a direct consequence of the emissions of certain industrial gases, chlorofluorocarbons (CFCs), earning them the 1995 Nobel Prize in Chemistry.
He was Institute Professor at the Massachusetts Institute of Technology (MIT) between 1989 and 2004; held research and teaching positions at the UNAM between 1967 and 1968; also at the University of California, Irvine between 1975 and 1979 and at the Jet Propulsion Laboratory of the California Institute of Technology (Caltech) between 1982 and 1989. Molina is a member of the National Academy of Sciences and the Institute of Medicine in the United States and since April 2011, he is one of the 21 scientists that serve on President Barack Obama´s Council of Advisors on Science and Technology (PCAST).
For his contribution to science he has received numerous awards including over 30 honorary degrees, the Tyler Prize for Environmental Achievement in 1983, the UNEP-Sasakawa Environment Prize in 1995, the 1995 Nobel Prize in Chemistry and the Presidential Medal of Freedom. Currently, he is professor at the University of California, San Diego. In Mexico, he is President of the Mario Molina Center. He is Vice President of the Board of INNOVEC.



### José Narro Robles

Rector of the National Autonomous University of Mexico. Mexico.

Jose Narro Robles was certified as a Medical Surgeon at the Faculty of Medicine at the National Autonomous University of Mexico (UNAM) where he received honorary mention for his thesis exam. He joined the staff at the Faculty of Medicine and carried out his postgraduate studies in Community Medicine at Birmingham University, England between 1976 and 1978. In the National University he lectured on preventative medicine, family medicine, public health, and served as titular for various graduate courses. He is currently a fulltime professor with 33 years of service.
At the university itself he worked as General Secretary, Dean of the Faculty of Medicine, among others. In November 2007 he was designated Rector of the UNAM for the 2007-2011 period. In November of 2011 he was appointed for a second four-year term.
He has held various posts in the federal public administration including General Director of Public Health in the Federal District, General Secretary of the Mexican Social Security Institute, Government Undersecretary in the Interior Department, and Undersecretary of Health Services in the Health Department.
He is the author of more than 220 academic publications and divulgation articles. Among his career he has received multiple awards.
Since 1992 he has been a member of the National Academy of Medicine and a member of the Mexican Academy of Sciences since 2004. He was also admitted to the National Royal Academy of Medicine in Spain as foreign academic correspondent.

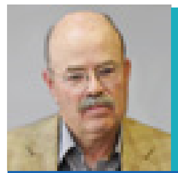
### Sylvia Schmelkes

President of the Board of Governors, National Institute for the Assessment of Education. Mexico.

Sylvia was born in Mexico City. She studied Sociology and holds a Master degree in Educational Research by the Iberoamerican University in Mexico City, She is an educational researcher since 1970.
She has held several institutional positions such as researcher-professor, Head of the Department of Educational Research from the Center for Research and Advanced Studies (CINVESTAV) of the National Polytechnic Institute (IPN) from 1994 to 2001. Advisor to the Secretary of Public Education from 1996 to 2000.
She chaired the governing board of the Center for Educational Research and Innovation of the OECD from March 2002 to May 2004.
She led the Research Institute for the Development of Education of the Iberoamerican University, Mexico City. She is the President of the Board of Governors of the National Institute for the Assessment of Education (INEE). She has carried out research in the fields of adult education, quality of basic education, intercultural education, and values formation. She is a Level Three National Researcher.

# Biographies

## Speakers, panelists and moderators



### Eduardo Backhoff Escudero

Member of the Board of Governors, National Institute for the Assessment of Education (INEE). Mexico.

Backhoff holds a degree in Psychology by the National Autonomous University of Mexico (UNAM). He also holds a Master degree in Education by the Washington University, and a Ph.D. in Educational Sciences by the Autonomous University of Aguascalientes. The evaluation of learning has been his line of research. He is a Level Two National Researcher.

He has collaborated with several dependencies engaged to the educational research. Interim Director of the Institute of Research and Educational Development of the Autonomous University of Baja California from 1993 to 1995; Director of the Institute of Research and Educational Development of the Autonomous University of Baja California from 1995 to 1999; Director of testing and measurement of the National Institute for Educational Evaluation (INEE) from 2004 to 2008; Scientific editor of the Electronic Magazine of Educational Research (REDIE) of the Autonomous University of Baja California from 2009 to 2011.

Member of the group of experts in questionnaires of context for PISA 2012, from 2010 to 2012; consultant of the Educational Testing Service (ETS) to lead cognitive laboratories of questionnaires of the context translated into Spanish for PISA 2012, in 2010. He is a Member of the Board of Governors of the National Institute for the Assessment of Education (INEE) since April 30, 2013.

### Shelley Peers

Director of the Primary Connections Project. Australia.

Shelley Peers is the Director of the Project 'Primary Conections: linking science with literacy' of the Australian Sciences Academy. She is medical biochemist and a professor of an elementary school.

She has held positions as a designer of programming and developing of curricular programs. She is Churchill member 2008-2009; member of the International Society for Design and Development in Education, and in 2010 she was awarded as an outstanding ex-student by the Faculty of Education of the Technology Queensland University, Australia. She has undertaken presentations about the program she leads on the IAP, so as to other events in South Africa, France, United Kingdom, United States of North America and Chile.



### Louise Hayward

Glasgow University. United Kingdom.

She is member of the Assessment Reform Group from United Kingdom. She has been working on assessment nationally and internationally with a focus on social justice.

She has a particular interest in the research and practice of public policy to improve education; specifically, how the assessment results can be based to take decisions on education and how teachers, learners, researchers and policy makers can improve them to work collaboratively.

She collaborates on the research work of the Assessment Reform Group, and has been part of the International Assessment for Learning Consortium. She has participated on the Scottish Policy Innovation in Research and Learning Initiative (SPIRAL) of Scotland.

### Ulrika Johansson

Director of the Science and Technology for Everybody (NTA) Program. Sweden.

Ulrika Johanson is a Mathematics and Science teacher since 1997.

From 2005 she leads the NTA (Naturvetenskap och Teknik för Alla) in Linköping, apart from being the responsible of the improvement program of Teaching and Learning Science and Technology in Elementary Schools.

On 2008 and 2009 Ulrika was the Development Director of the project called 'Science and Technology in and out of school.'

This project involved pre-school, elementary schools, and highschools in Linköping, developing a strategy to students from one year up to sixteen. The main approach of the initiative was the development of a strategic working plan of science and sustainable development collaborating with enterprises, museums, science centres and Linköping University. In 2013 Ulrika Johansson became a member of the National Swedish Agency of Educational Programs for the Science and Technology Development.

### María Alejandra González Dávila

Science at your School, Program of the Mexican Academy of Sciences. Mexico.

She holds a Master in Sciences with a specialty on Educational Research on the Analysis of Discourse and Science Teaching by the Educational Researches Department of the CINVESTAV.

She holds a degree in Pedagogy by the UNAM. She also studied Biomedical Engineering with a specialty in Electronic Medical Instrumentation and worked on the Nuclear Medicine area of the National Institute of Cardiology.

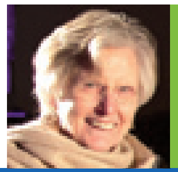
She collaborates on the magazine Mail of the Teacher and has been responsible of the teachers training of the Innovation Laboratory on Educational Technologies (LITE). Author of science text books, and interactive resources for distance education (*telesecundaria*) at the Latin American Institute for Educational Communication (ILCE). She has published texts for high school.

She is assessment coordinator and member of the teaching group of the Program 'The Science at your School' of the Mexican Academy of Science. (AMC).

## Carlos Mancera Corcuera

*Valora Consultores, S.C.* Mexico.

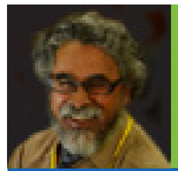
He is an economist from the Autonomous Technologic Institute of Mexico (ITAM). Since 1988, he has taken several positions on public administration dependencies, such as Advisors' coordinator of Programming and Budget in the Mexican Ministry of Public Education (SEP), and Assistant Director for Scientific and Technological Policy in the National Council of Science and Technology (CONACYT).
He was Assistant Secretary of Planning at SEP, since 1994, for a six-year period. He had an active role in multiple projects, wich conformed the Educational Policy between 1992 and 2000. Since 2001, he is partner of *Valora Consultoria S.C.*, an enterprise dedicated to the advise on cultural and educational topics. He has also done several works for the Interamerican Development Bank, the Global Bank and the OECD.



## Wynne Harlen

University of Bristol. United Kingdom.

Wynne Harlen is a graduated from Oxford University with honorary degree in Physics and a Ph. D. in the University of Bristol. Among other positions, she has been head of the Department of Teaching at the University of Liverpool and Director of the Scottish Council for Research in Education, in Edinburgh. She has been a consultant and Co-Director of Research Projects of the National Foundation of Science (NSF), TERC and Cambridge, among others.
She was President of the Scientific Group of experts to the OECD PISA Project during its first six years. She chaired the International Committee of Oversight of the Program of Science Teaching of the InterAcademy Panel (IAP) for the development of the inquiry based science teaching in pre-secondary schools. She is a founding member of the British Association of Educational Research.
Wynne Harlen was decorated by the Queen with the order of the British Empire in 1991, and in 2001 received a special recognition for her distinguished services in science education by the Association of Science Education (ASE). She has participated on the editorial boards of several international journals. Her publications include 25 research reports, more than 160 articles on specialised journals, contributions in 38 books, and 30 books of her authorship or co-authorship. She currently directs the Project 'Evaluation Systems for the Future,' sponsored by Nuffield Foundation in Cambridge University.

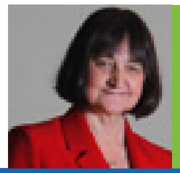


## Guillermo Solano

University of Colorado, Boulder. United States.

Postdoctorate in Measurement and Assessment Development by the University of California, Santa Barbara. Ph.D., in Education and a specialty in methodology and measurement, by the same University.
He holds a Master degree in Educational Psychology, by the National Autonomous University of Mexico, he is a specialist in the measurement of the education, assessment development, and relevant linguistic and cultural affairs for the assessment of the linguistic minorities.
He is professor of bilingual education and English as a second language in the Education School of the University of Colorado, in Boulder. He is the author of the theory: Test Translation Error, pointing to the reflection of the cross-cultural evaluation. He is currently, working on a research on several forms of measurement in maths testing, forms of assessment for students in the science classroom, and the design of illustrations such as assessment of English language learners by combining cognitive sociolinguistic, semiotic and scientific approaches.





## Rosa Devés

Vice President of the University of Chile in Santiago. Chile.

Rosa Devés got a degree in Biochemistry at the University of Chile in 1974 and a PhD in Biochemistry at the University of Western Ontario, Canada in 1978. She subsequently held a postdoctoral degree at the University of Southern California, in Los Angeles.
In 1980 she joined the Department of Physiology and Biophysics in the Medicine Faculty of the University of Chile. She has actively participated in the development of graduate education, including the creation of the Doctorate Program in Biomedical Sciences led by two periods of 5 years.
She also participated in the creation of the Institute of Biomedical Sciences. Between 2006 and 2010 was Director of Postgraduate degree from the University of Chile and since 2010 she has been VICE PRESIDENT. Since 2003, she has been a corresponding member of the Chilean Academy of Sciences.
Along with her scientific and academic career she has been involved in the improvement of science education at the school, collaborating between 1999 and 2002 with the unit of Curriculum and Evaluation of the Education Ministry as coordinator of the science teams in the curriculum development. In collaboration with Jorge Allende, in 2003 created the establishment of IBSE Program (inquiry-based science education) in a partnership between the Ministry of education, universities, the Chilean Academy of Sciences, the municipalities and the schools, in order to bring a quality science education to all children.

## Petra Skiebe-Correte

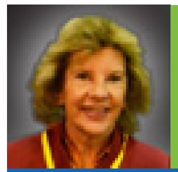
Director of Pollen Program. Germany.

PhD in neurobiology. Director of NatLab at the Freie University of Berlin where she offers both, elementary and higher education students the opportunity to perform updated experiments based on research in Biology and Chemistry, developed by scientists. She started a Laboratory Network of Informal Science in the States of Berlin and Brandenburg (GenaU).
To support science teaching based on long-term researching within elementary schools, she founded TuWaS! (Technik und Naturwissen¬schaften an Schulen, Technology and Science in Schools) in 2007 as a cooperation between the Freie University of Berlin and the Berlin Brandenburg Academy of Sciences and Humanities.
Currently, TuWaS! is serving over 140 schools in four different German states. Within the "Fibonacci" Project, TuWaS! cooperates with three countries: Luxembourg, Austria and Turkey.
Skiebe-Corrette also participates in the Advisory Council of the National Center of scientific resources in Washington, D.C., and in the LernortLabor Bundesverband der Schulerlabore e.V., a support association for laboratories in informal science within Germany.

## Anne Goube

La main à la Pâte Program. France.

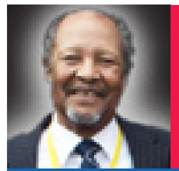
Master degree in Physics and Chemistry in Paris in 1974. Between 1982 and 1983, at the University of San Jose in California, in the United States, she was certified in Behaviour Management in the Classroom and in the Analysis of Teaching Styles.
She has a long professional career in which she has performed several positions: full-time teacher of secondary school from 1976 to 2001; instructor of teachers in service, from 1984 to 1990; instructor of teachers in training, 1990-2001; teacher of Chemistry in the Communal University of Austin, Texas, from 2001 to 2002.
Upon his return to France and until 2011, she served as an instructor of teachers in the Joseph Fourier of Grenoble University, and as an instructor for primary teachers on inquiry-based science teaching.
From 2011until today, she works as a national and international volunteer at the French program *La main à la pâte*, which promotes inquiry based science education at schools.



## Norma Sbarbati Nudelman

President of the Program of Education in Sciences of the Inter-American Network of Academies of Sciences. Argentina.

PhD in Chemistry at the University of Buenos Aires, Postdoctorate at MIT (Cambridge, United States); she was an associate researcher at the University of California (Santa Cruz, United States); visiting professor at the University of East Anglia, (Norwich, United Kingdom); professor at Kyushu University (Fukuoka, Japan); visiting professor at the University of Marseille (Marseille, France); visiting professor at the University of Valencia (Valencia, Spain), and visiting lecturer in several universities in Europe, Asia and America.
Currently works as a plenary professor in the University of Buenos Aires, as a senior researcher at the National Council of Research (CONICET). She is a holder member of the National Academy of Exact, Physical and Natural Sciences, and is President of the Program of Education in Sciences of the Inter-American Network of Academies of Sciences.
She has developed important research work in the field of organometallic chemistry; green chemistry; physical organic chemistry; environmental chemistry, applied on stability of drugs as well as pharmaceutical products.
She has received numerous awards including: the award for distinguished research by the Ministry of Science and Technology. The award for the trajectory in Chemical Education by the Chemical Argentinian Society; and Chemist of the year by the Argentina Society of Chemistry.



## Hubert Dyasi

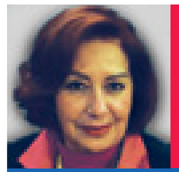
Emeritus Professor of the City College, City University of New York. United States.

Professor of Science Education, specialised in the professional development of science teachers. He obtained his Ph.D. in Science Education at the University of Illinois at Urbana-Champaign.
He has taught courses in science education at pre and postgraduate education level, and supervised students, of nationally and internationally school-based teachers, apart from leading the educational science programs for scholar districts of New York City. He also worked for the City College of the City University of New York.
He has collaborated with the Department of Education of New York and nationally in the US with schools and scholar districts to develop its programs of education in science and to implement inquiry based science education in the classroom.
He is a member of numerous councils and consultant of teaching and learning in science. He has participated in several panels as well as visit teams of the National Science Foundation.

## Jorge Alejandro Neyra González

Subsecretary of Basic and Normal Education of the State of Mexico. Mexico.

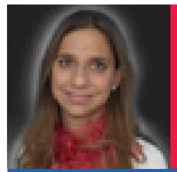
He holds a degree in Political Science and Public Administration at the Autonomous University of the State of Mexico (UAEM), with specialized courses in Education, Administration, and Leadership by the New Mexico University, in electoral studies by the Metropolitan Autonomous University and the UAEM. He is currently studying a Master in Administration with a specialty in Senior Management in the in the College of Postgraduate Studies of Mexico City.
From February 7, 2013, he works as Assistant Secretary of Basic and Normal Education of the State of Mexico.
He was the General Director of High school in the State of Mexico from 2008 to 2013; General Director of the Electoral Institute of the State of Mexico 2004-2005; he was Executive Member of the District Boards and President of the 34 and 04 District Councils with a seat in Toluca and Villa Nicolas Romero of the Federal Electoral Institute in the period of 1994-2004. He has been awarded for conferences given in academic and political institutions about education, electoral-political topics, oratory and political debate.



## Cristina Aguilar Ibarra

Subdirector of Mathematics and Natural Sciences Tests of the National Institute for the Assessment of Education (INEE). Mexico.

Teacher of elementary school, graduated from the National School of Teachers; she is also a High school teacher in Biology. She holds a Master degree in Education at the Latin American Institute of the Educational Communication (ILCE) and Doctoral studies in Educational Assessment at Anahuac University.
Member of the founding team of the General Board of Assessment of Education at the Mexican Ministry of Pulic Education (SEP), institution in which she has worked as Coordinator of the Natural Sciences area in assessment projects for basic, secondary and normal education. She was also head of the Department of Contents and Methods at SEP and Director of Education in the National Population Council. She is the author of several educational materials about the didactic of natural sciences, education in population and sexual education. Since 2003, she works at the National Institute for the Assessment of Education (INEE), in which she has performed, first as Coordinator for the elaboration of the EXCALE or examinations of the quality and the educational achievements of Natural Sciences and, currently, as an Subdirector of Maths, and Natural Science tests.



## María Figueroa

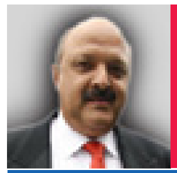
Dean of Education of Externado University of Colombia. Colombia.

Biologist by the University of the Andes with a Master degree in Education in Sciences by the Teacher's College of Columbia University. She holds a Doctorate in Education with an emphasis on assessment, by Stanford University. She did a research to compare learning in students who participate in inquiry based science teaching programs with students who participate in other programs.
Since 2001 and until 2006, she served as Coordinator of teachers' training at Little Scientists Program in Colombia. He is currently the Dean of Education in the Externado University of Colombia, and advisor in the Institute for Assessment of Colombia, in several projects.

## Patricia Rowell

Emeritus Professor in Science Education, University of Alberta. Canada.

Patricia Rowell is an emeritus Professor in Science Education at the University of Alberta in Canada.
She holds a Bachelor degree and a Master degree in Biochemistry of the University college, London and Oxford University. In addition, she holds a Ph. D in Science Education at the University of Alberta.
With USAID, was designated Senior Technical Advisor of Namibia's Government for two years, with the responsibility to develop the science curriculum for primary schools.
As a member of the Working Group of the InterAcademy Panel Working Group on Science Education, has collaborated with an international group of science educators and academics to support Inquiry Based Science Education (IBSE) in developing countries. Through an invitation from Chile Government, she was a member of the International Assesment Team of the Science Program based on Research on that country, together with professors Harlen and Lena, she also collaborates with the Centre for Research in Learning Science at Southeast University in Nanjing, which promotes a reform in primary science education.

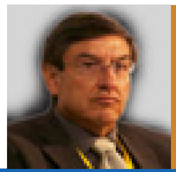


## Reyes Tamez Guerra

Autonomous University of Nuevo León. Mexico.

Reyes Tamez Guerra graduated as Chemical Bacteriologist Parasitologist from the Autonomous University of Nuevo León (UANL) and held a Master's degree and a Doctorate in Science with specialty in Immunology at the National School of Biological Sciences of the National Polytechnic Institute (IPN). At the UANL he had different positions, such as: head of the Immunology and Microbiology Department, Director of the Faculty of biological sciences, General Secretary, and Rector of the University.
In relation to education, he worked as a member of study and cooperation from the Latin American Union of universities (UDUAL), President of the Northeast Regional Council and member of the National Council of ANUIES (1996-2000). Among other public positions, he was head of the Mexican Ministry of Public Education (2000-2006), and on 2007 was head of the Ministry of Public Education of Nuevo León.
He has also been a member of the Evaluating Committees of Scholarships and Research Projects of the Direction of Scientific Development of CONACYT and member of the Board of the same institution.
He was a counsellor at the Commission of Planning of Senior Education (COEPES) of Nuevo León and alternate Vice-president of the Inter-American Organization in Canada.
He is currently a full time Professor of the Autonomous University of Nuevo León.





## Armando Loera Varela

*Heurística Educativa*. Mexico.

Bachelor in Philosophy by the Autonomous University of Chihuahua and Master in Education by Harvard University.
He has been Professor of the Institute for the Social Development of the Inter-American Development Bank in Washington D.C., of the Alberto Hurtado University of Chile and the Technological and Graduate Studies Institute (ITESM) - Campus Chihuahua.
As a researcher, he was part of the Institute for International Development of Harvard University; he was responsible of the coordination of the Research and Academic Development of the Education Department of Chihuahua Government.
Since 2001, he directs the *Heurística Educativa* team, conducting studies for the Mexican Ministry of Public Education, BID and UNESCO.
He has been Advisor of the National Pedagogical University and consultant for UNICEF, Global Bank, BID, among other institutions. He is currently doing a comparative study about teaching maths and science in the Dominican Republic, Paraguay and Nuevo León. He also supports as a consultant to the Enterpreneurs Foundation for Basic Education in Mexico (ExEb), where he has published (along with Esteban García Hernández and Oscar Cazares Delgado) books about his model of schools accompaniment: "The learning-centred teaching" and "Learning-centered school management".

## Lee Yee Cheong

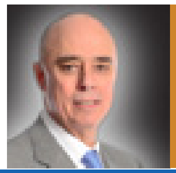
President of the Council of the Science Education Program, IAP. Malasia.

Electrical Engineer at the University of Adelaide. He was part of the National Board of Malaysia's electricity until 1980. Subsequently, he was Director and Executive Director of Tenaga Ewbank Preece (M) Sdn Bhd (TEP) until 2002.
He is Director of UMW Holdings Berhad.
Member of the Energy Commission of Malaysia.
He is an assessor of the Ministry of Science Technology and Innovation of Malaysia, founder of the General Secretariat, Vice President and Treasurer of the Science Academy of Malaysia (ASM), founding member of the inter-academic board of sciences academies of the world and a member of the Academic Council of the World Economic Forum (2001-05).
He is the General Secretary of the Federation of the academies and science associations of Asia (FASAS) and the founding President of the engineering and technology Academy in Asia (AAE T).
He is a member of the Australian Academy of technological Sciences and Engineering, and a member of the engineering Pan-American Academy. Dr. Lee was recognized by the Malaysia Government with DPMP and KMN awards for his engineering work.
He was appointed honorary official from de Australian Order (AO) for his contributions in the strengthening of relations between Malaysia and Australia.

## Ubaldo Ávila Ávila

Subsecretary of Basic and Normal Education of Zacatecas. Mexico.

Normalist Proffesor since 1981, working as rural multigrade teacher until 1988.
He studied at the Normal Senior School in Durango with a specialty in Natural Sciences. Teacher of Biology, Physics and Chemistry in technical secondaries from 1988 to 2004.
Speaker of Experiences in rural school and multigrade groups in forums in 1990 and 1991. He studied a master degree in Education.
He has been working in the public sector in the Secretariat of Culture and Education at the regional level with different responsibilities.
He has occupied several spaces of popular choice, among which stood out as Local Deputy of the LVI and LIX legislatures of Zacatecas State. Since 2010 he is the Subsecretary for Basic and Normal Education in Zacatecas, taking responsibility to coordinate and follow-up educational actions in education initial, preschool, primary, secondary and normal schools.



## Arturo M. Fernández Pérez

Rector of the Autonomous Technological Institute of Mexico (ITAM). Mexico.

Arturo Manuel Fernández Pérez holds a degree in Economics from ITAM; a Master degree and Ph.D. in Economics from the University of Chicago, he got second place in the national Prize of Economy on a level-research, BANAMEX in 1987. Director of the Technological Autonomous Institute of Mexico from 1992 up to date.
He has been coordinator of advisors of the Secretariat and Coordinator of the Economic deregulation program, Secretariat of Commerce and Industrial Development (1989-1991); General Director of Academic Division in Economics, Rights and Social Sciences of the ITAM (1987-1989); Head of the Academic Department of Economics, ITAM (1983-1986); Advisor at SHCP, (1983); Member of the Administration Council of Peñoles Industries, S.A.B. de C.V.; National Provincial Group, S.A.; El Palacio de Hierro, S.A.B.; Mexican Values,Brokerage House; Financial group BBVA Bancomer, S.A.; FEMSA; Bimbo, S.A.B. de C.V., and Fresnillo, PLC.

## Jon K.Price

Director of INTEL's Research and Assessment Program. United States.

He is a graduate of the University of New Mexico, the Harvard Graduate School of Education and received his PhD in Education from the Texas A&M University College of Education.
He has been managing the education technology program evaluation efforts for Intel global education K-12 education initiatives since 2003. In 2008 his responsibilities expanded to include additional research and evaluation in to how effective integration of technology in to multiple levels of education can impact teaching, learning, education reform and economic growth.
Jon Price manages series a rigorous program evaluations in order to ensure continuous, targeted improvement of all the Intel's educational products and activities. These evaluations take place as a result of research grands from the Intel Foundation and Intel Corporation for multiple Intel Education initiatives such as Intel Teach, Intel Learn and Intel International Science and Engineering Fair.



## Daniel Alcázar Román

Smithsonian Science Education Center. United States.

He obtained the Science Bachelor's degree, in the field of Engineering Sciences at the Lipscomb University, Nashville in 1999 and the Master degree of leadership in education at the South University of Texas, Houston. Since July 2012 he is Supervisor of development Curricular in Science and technology of Alexandrias's public schools
From 2007 until January 2012, he was a specialist in Curriculum and Evaluation of the Independent Scholar Districts of Houston; from 1999 to 2007 he was a Science Professor and instruction Coordinator at Houston independent school districts. He has been educational consultant in different places like the Smithsonian and the Northwest Evaluation Association (NKWEA).
Among the functions related to the management of STEM programs, he has been leader of district educational reforms, has linked the curriculum and evaluation divisions in State agencies of education, the National Center for Statistics Education and the Department of Education of the United States of North America.
Among the awards that have been granted are: expert STEM Instructor from the Smithsonian for LASER 2009, Award for Excellence in Environmental Education in 2007, scholarship for Excellence in Science Teaching by the Environmental Institute of Houston and the University of Houston, the Nothrop Grumman Foundation in 2006 Award and the Scholarship of Leadership in Science Teaching from the College of Baylor medicine.

## Anders Hedberg

President of Hedberg Consulting, LLC. United States.

PhD in Pharmacology form the Medicine School of Goteborg University, Sweden. He has taught and conducted research in the departments of Physiology and Pharmacology of the Medicine School of the University of Göteborg; AstraZeneca, Gothenburg, Sweden; Universitätsklinikum Johann Wolfgang Goethe, Frankfurt a/m, Germany; Medicine School of the University of Colorado in Denver, and the Medicine school of the University of Pennsylvania, where he completed a three-year Post-Doctoral research scholarship.
For 15 years, Dr. Hedberg was devoted to research and published several papers on the mechanisms of pharmacological intervention on hypertension, heart failure, thrombosis and myocardial ischemia, holding positions as: scientist in Chief, leader of the research group and head of section in cardiovascular Pharmacology and drug discovery in Astra-Zeneca and Bristol-Myers Squibb.
Dr. Anders Hedberg is a former business executive with 30 years of experience in pharmaceutical, Corporate Affairs and the promotion of science, health and education. In his role as Director of philanthropy area corporate of Bristol-Myers Squibb, Dr. Hedberg directed the world programme of education of science of Bristol-Myers Squibb (BMS).
From his position Dr. Hedberg established strong alliances between BMS and major national and international agencies for the teaching of science to government level, private enterprises and non-profit organizations.

## Leopoldo Rodríguez

Member of the Board of INNOVEC. Mexico.

Leopoldo Rodríguez is consultant and Associate to various enterprises. He has been awarded with the Andrés Manuel del Rio National Chemical Prize in 1988 and the Ernesto Rios Prize in 1997. Among other activities and posts, he was UNAM professor for 17 years, professor at the Ibero American University and the National Polytechnic Institute.

He is Immediate Past Chair of the Mexican Directives Association of Applied Research and Technological Development (ADIAT); Member of the Chemistry Engineering Commission of the Academy of Engineering of Mexico, and Member of the Innovation for Science Teaching Council (INNOVEC).

He is also Member of the Assistant Council of the National Sciences Resources Center-Washington, D.C.; Member of the Directive Council of the Mario Molina Center; Member of the Governing Board of CONACYT, Member of the Development and Technology Energy Sector Commission of SENER and Member of the CENEVAL Chemical Engineering Technical Council.

## Guillermo Fernández de la Garza

Chief Executive Officer of FUMEC, Miembro del Consejo Directivo de INNOVEC. México.

Electrical, Mechanical Engineer Bachelor, Physics Bachelor from the National University of Mexico (UNAM) and Engineering Economics Systems M.Sc. from Stanford University. He also coursed the IPADE's Program for CEOs.

As Deputy Director of the National Council for Science and Technology (CONACYT) he created different scientific and technological information services for industry, one of those is INFOTEC. He was Executive Director of the Electrical Research Institute (IIE). Fernandez has been a leading technical advisor to projects at the UNDP-GEF and has actively participated in several Mexican and international professional organizations such as the Mexican Association of University Mechanical and Electrical Engineers (AIUME), Mexican Association of Mechanical and Electrical Engineers (AMIME) and Institute of Electrical and Electronics Engineers (IEEE). He is Member of the Cross-border Institute for Regional Development (CBIRD), and Member of the Advisory Board of the National Science Resources Center (NSRC). He has participated as Advisor for the United Nations Educational, Scientific and Cultural Organization (UNESCO), United Nations Industrial Development Organization (UNIDO) and International Atomic Energy Agency (IAEA).

The Academy of Science of France has awarded him with the PurKwa Prize 2008, because his enthusiasm in fostering programs about scientific education for children in Mexico.

Innovation in Science Education
INNOVEC

San Francisco 1626-205,
Colonia del Valle. C.P. 03100, Mexico City.
www.innovec.org.mx