

---

# Assessing Learning Outcomes: What Works?

Isak Froumin

Invited Paper

31<sup>st</sup> Annual Conference of the Pacific Circle Consortium  
Honolulu, Hawai'i  
June 25-29, 2007



---

## **Sub Theme III**

What works? What are promising practices?

---

# **ASSESSING LEARNING OUTCOMES: WHAT WORKS?**

**Isak Froumin**



**WORLD BANK**

**Moscow,  
April 2007**

## I. Introduction

In the aftermath of the 2<sup>nd</sup> World War, educational policies became oriented towards a new target of increasing student participation rates in education. This goal was part of a wider education strategy aimed at increasing participation in education. As governments came to the realization of the value of education in economic development and direct and indirect benefits it provides for individuals and the society at large, they increased educational funding manifold (UNESCO, 1990). Political and educational leaders at the time assumed that channeling more money into education and providing more input in the form of teacher preparation, student participation rates, physical facilities, and curriculum materials would result in stronger and better education systems and thus make their economies more competitive. The focus on learning outcomes was almost missing in educational policy.

The first solid foundation for outcome-based education was laid by Benjamin Bloom's influential work "Taxonomy of Educational Objectives: the Classification of Educational Goals" published in 1956. Although the work had had little impact on curriculum, it had a significant effect on learner assessment in general. Bloom's taxonomy proved to be a powerful tool for turning the focus of attention towards outcome-oriented education, as it attempted to state clearly educational goals. It also gave rise to other literature on assessments such as Ralph Tyler's work on a model of evaluation and many others, which, in their turn, contributed, to a large extent, to the materialization of statewide assessments in the US in 1970 that grew to include 35 states by 1985. Practically all of them drew heavily on Bloom's taxonomy (Marzano, 2001). Soon after the publication of the Bloom's work, learner assessments went international in 1961, when a pilot study launched by the International Association for the Evaluation of Educational Achievement, got about 60 countries involved in the study of student achievement in various subject areas (Kellaghan & Greany, 2001).

Today, the interest in learning outcomes, being driven primarily by economic concerns, has come to take center stage in educational policy making and planning amid calls for greater accountability, efficiency and pressure to improve teaching and learning. One of the most important documents that brought outcome-oriented education to the fore for the international community had been adopted by the World Conference on Education for All in Jomtien in 1990. It indicated that provision of basic education for all made sense only if children acquired needed knowledge, skills and values. Almost 10 years after the Jomtien conference, the Dakar Framework for Action (2000) emphasized the significance of "having a clear definition and accurate assessment of learning outcomes (including knowledge, skills, attitudes and values)" (UNESCO, 2000). The number of assessments has since been growing steadily across the world with generous assistance from governments, especially of the developed nations and international agencies including the World Bank, UNESCO/UNICEF, and the International Institute for Educational Planning (IIEP). Techniques for assessing educational achievement, including performance across nations have become more sophisticated and advanced. Major research projects have been initiated to study patterns of educational achievement and to explore both in-school and external factors that contribute to learning. Considerable efforts have been made to inform developing countries of possible usage of data on student performance to make the most efficient and effective use of their educational resources (UNESCO, 2000).

At present, countries across the world are employing four major measurement devices, which could be grouped into two distinct categories in accordance with their functions – national and



international assessments<sup>1</sup> have best served as monitoring tools for school systems through collection of sample-based statistical information about students' learning progress, based on which school system performance is judged; public (national) examinations and school-based assessments have been used for selection and certification purposes which usually involve high stakes for individuals. High-stakes testing is also used in part for monitoring standards on the individual level, or as is the case with school-based assessments, for intermediate individual monitoring to improve immediate learning results, though monitoring is not their main goal. The dual nature of existing forms of assessments has stirred a lot of heated debates regarding the benefits of both low- and high-stakes testing for measuring the outcomes of education and their impacts on teaching, learning, policy setting, resource allocation, etc.

With all kinds of assessments mushrooming worldwide and heavy investment in education (around 5-6% of GDP in OECD countries), the interest in impacts of assessments is all but surprising especially in the light of persistent calls for evidence-based policy making which is possible only when outcomes of education are closely monitored and scrutinized and are then used to make informed policy decisions based on concrete data. Policy makers have been raising questions as to whether learning assessments have really made any significant impact on the education systems which they target. Has their use by governments and international organizations been changing teaching and learning in schools? How have assessments been changing existing school practice and influencing policy decisions? Have assessments been the driving force behind the education sector development precipitating major changes in education systems? Indeed, what is the use of organizing any type of assessment but to use the results thereof to improve learning and teaching and to effect necessary improvement and in some cases sweeping institutional reforms? The real value of spending a great deal of time, money and effort on preparation, implementation and secondary analysis of assessment output lies in scrutinizing the data gathered and feeding it back into curriculum, teacher training, learning materials and methods to introduce necessary adjustments.

This current paper attempts to provide answers to the foregoing questions with a view to inform educational planners, administrators, and managers as to whether the allocation of funds within education budgets towards carrying out learning assessments has brought about any meaningful changes in teaching and learning in the classroom and policymaking. The first part of the analytical note looks at current trends in learning assessments. The second part considers evidence of the impact made by different kinds of assessments on education systems and educational practice. Concluding thoughts follow at the end of the paper and review the questions put up for discussion in the introduction.

---

<sup>1</sup> A more detailed description of these can be found in Annex 2.



## II. Current trends: growing value of assessments

### 1. A shift from inputs to outcomes

In the past, educational policies were mainly input-oriented with emphasis being put on such quantitative indicators as student participation rates, physical facilities, curriculum materials, teacher training (Greany & Kellaghan, 2001). The input-oriented policies were based on the assumption that data supplied by these indicators could well serve as a proxy for educational outcomes, which, it was thought, were difficult to measure. Consequently, governments, development decision makers, national and international advisors embraced policies aimed at achieving quantity rather than quality focusing their efforts on educational enrollment targets (Coombs, 1968). This, in part, was also due to the fact that the past educational policies have to a large extent been dictated by political, military and religious concerns as opposed to economic goals and interests, whose main concern is efficiency (Ramirez, Luo et al, 2006).

At the turn of the century, the picture is drastically different. In recent years there has been a growing interest in measuring students' learning outcomes and their achievement levels through assessments of various kinds (Ramirez, Luo et al, 2006; Greany & Kellaghan, 1996; Greany & Kellaghan, 2001; Sahlberg, 2006; Schiefelbein & Schiefelbein, 2003; UNESCO, 2000), as substantial research demonstrated "the elusive nature between the various elements that go into schooling and student performance and also shown that the effectiveness of particular inputs varies widely from country to country" (UNESCO, 2000, p.9). As a result, policy makers have been shying away from educational inputs as a measure of educational quality and focusing more on gauging concrete learning achievements of students, such as adequate knowledge, skills, behavior and attitudes needed to succeed in adult life. It is important to note that in measuring learning outcomes a stronger emphasis has been placed not only on what students are learning in the classroom through simple memorization of facts and passive recognition of information, but also on their abilities to apply skills and competences acquired through critical thinking, clear communication and problem-solving in real-life situations. The competence-based approach to education expanded the scope of existing assessments beyond heavily content- or knowledge-based framework (Crighton, 2003).

Greany and Kellaghan (2001) name four key reasons behind the paradigm shift towards outcome-based education, which has led to the emergence of assessments:

- a) ample evidence that students during their exposure to schooling have low competency level in basic learning skills;
- b) signs of deteriorating quality in schools as enrolments increase;
- c) concern that the acquired competences will be misaligned with labor market demand
- d) growing interest on the part of ministries to procure evidence of the impact of funding invested in education.

Nowadays, in most countries, an outcome-based approach to education is realized through a combination of several types of assessments, which are used to provide support for educational objectives and standards. In accordance with their functions, assessments can be roughly divided into monitoring-oriented and selection/certification-oriented. The monitoring-oriented category includes national and international assessments, whose primary goal is to gather information on learning outcomes of a school system as a whole. More specifically, as far as national assessments are concerned, they are generally curricula-based and measure performance of the system rather than individuals and the results are used for monitoring



national standards over time. The data collection is normally conducted on the scale that either involves whole school population (census-based testing) or random samples of relevant age-cohorts. International assessments in many aspects resemble national testing, with the only difference being they are not tied to any specific curricula, are primarily sample-based and rely on standards and corresponding indicators developed by international agencies. Another obvious feature of international assessments which could be inferred from their name is that, unlike national assessments, they are mostly large-scale, multiple-country and are used for between-country comparison. Both national and international studies of student achievement are considered low-stakes and, as a result, pose fewer security and malpractice risks (Crighton, 2003).

Selection- or certification-oriented assessments typically include public examinations and school-based assessments. They are closely interrelated, as results of school-based assessments can be used for scoring public examinations and sometimes even become the sole basis for certification and selection (UNESCO, 2000). Public examinations are a common practice in most countries including developing ones, as pupils have to take some form of testing at transition points in schooling. However, there are some countries like the US, Canada and Sweden which do not have their own national public examinations. Selection- or certification-oriented assessments are typically associated with high stakes for individuals participating in them, which distorts teaching and learning by forcing teachers to practice “teaching to the test”. As stakes are high, so are security risks, which could compromise test results.

Because of existing drawbacks of examinations, it has been argued that education would be better served by less formal, continuous (formative) school-based assessment, which would be of special benefit for teachers, because it allows them to match the assessment more closely to both the curriculum and the individual student (Rehmani, 2003). At the moment, most countries use a mix of external and internal tests, scores of which are combined by the examining agency to produce the final score.

Although this paper is mostly concerned with traditional studies of student achievement, it is worth mentioning that alternative assessments have also been gaining attention. In addition to such an alternative form of school-based assessment as formative evaluation, there has been more talk of greater focus on more sophisticated and extensive models of assessments. This push for reform has given birth to such new techniques as:

**Portfolio assessment.** It is defined as "a purposeful collection of student work that exhibits the student's efforts, progress and achievements in one or more areas. The collection must include student participation in selecting contents, the criteria for selection, the criteria for judging merit and evidence of student self-reflection." (Paulson, Paulson, Meyer 1991). The key goal of the portfolio is to help a student demonstrate that he is really learning and progressing well and their greatest value is that students become active participants in the learning process and its evaluation.

**On-demand competency testing:** it allows students to be tested when they are ready rather than waiting for examination. Although this technique gives students flexibility, it may affect their discipline and lead them to take a lax approach to testing.

**Performance-based assessments.** Performance assessment is a testing technique that requires students to perform a task rather than select a correct answer from a multiple choice list. For



instance, a teacher may ask a student to explain some event, generate hypotheses, solve math problems, and do research on a given topic. This alternative form of assessment is mostly used for assessing the writer's ability based on text composed by students under test instructions.

The change in educational policies towards more focus on learning outcomes has spawned a variety of assessments worldwide. The shift to outcome-based education has by now been accomplished in many developed countries and is still spreading to developing nations, which are eagerly adopting new techniques, though some of them are reluctant to share testing results, as assessments, by and large, have become politicized and their growing popularity and eager adoption by governments can in part be attributed to the fact that politicians have realized that studies of student achievement can be used as a political weapon against opponents or, like sports, for between-country competition.

## **2. Evidence of growing interest in learning assessments: more political support**

The political support from governments of various countries for learning assessments has been strong. Unsatisfactory results on international tests helped garner political support for higher standards in US schools in the 1990s, and especially for more consistent testing and tougher accountability measures in the No Child Left Behind Act, a centerpiece of President Bush's domestic program in his first term. The president campaigned to extend that testing regime into US high schools in his second term.

Although the US assessment program NAEP has been in existence since 1969, it is only in recent years that politicians appear to have become interested in its findings (Smith, O'Day and Cohen. 1990). Political interest was heightened by the attention of the National Governors' Association to the findings of NAEP and led to the introduction of the state-by-state comparisons (Phillips, 1991). Another evidence of growing popularity of assessments with political circles emerges from the work by Bennett and Desforges (1991) who point out that "assessments in the 1980s carried considerable political weight, contributing to the major curriculum reform movement embodied in the 1988 Education Act that defined a national curriculum in England and Wales for the first time".

In Russia, more political support for outcome-based education is evidenced by the introduction in 2001 of the Unified State Examination (USE) for school graduation and for university entrance. The USE is now poised to replace the existing examination system in May, 2009 after the State Duma has passed the law on the USE<sup>2</sup>. The USE is designed to protect common standards, decrease corruption and informal payments, and increase access for students from rural areas and the disabled. Summarizing the most recent exam data, it is worth noting that in 2006 around 830,415 students participated in the annual exam which is 71.9% of the total number of school leavers; in the same year 500,000 enrolled in 1,650 universities and their branches as full-time students across the country<sup>3</sup>. Enrollments in HEIs of the rural population of the relevant age cohort that participated in the USE have grown in 2006 as compared to previous years<sup>4</sup>. In addition, Russian politicians have been very supportive of conducting PISA in the country and have paid a great deal of attention to the

<sup>2</sup> RIA News Agency ([www.rian.ru](http://www.rian.ru))

<sup>3</sup> <http://www.edu.sochi.ru/news.htm#n00010>

<http://ege.edu.ru:8080/ege/portal/news.nsf/0/2B023332A2718A4DC3256F2C0032EE44?Opendocument>

<sup>4</sup> Gvozdeva, S.I. On USE results. Presentation at the seminar in Sochi.



results. They have vested interest in PISA outcomes, as Russia's success in international assessments can be used as an indicator of the government's great performance.

In New Zealand, educational assessments have been an integral part of the national assessment strategy since 1999. The national government has sponsored a series of seminars for teachers to promote best practices, and supported the development of a CD-based computer program, Assessment Tools for Teaching and Learning, that helps teachers create classroom assessments and interpret the results, with an eye toward modifying their instruction.

In Scotland, national guidelines on assessment encourage teachers to think systematically about assessment as part of teaching and learning. The "Assessment is for Learning" program encourages teachers to make greater use of "comments only" feedback to students and more use of student self- and peer-assessment, for example, and less use of summary judgments, such as letter grades. In England, the government-sponsored "Assessment for Learning" program also provides professional development to support teachers' greater use of formative assessments in shaping the instruction of individual students.

At the system level, the further evidence of political commitment to the focus on learning outcomes can be found in the formal establishment of governing bodies responsible for controlling an assessment or an examination system. For example, such bodies include the National Matura Commission in Slovenia or national and regional assessment agencies, such as the ones found in Romania, Poland, Lithuania, etc (Bethel, 2003).

The growing political support for evidence-based policies has been demonstrated on the international level as well, as governments have repeatedly expressed their desire to pursue the course of quality education for all during the 2000 World Education Forum in Dakar, Senegal. At the Forum 164 countries signed up to the accord confirming the significance of equitable access to quality education which implies monitoring and evaluation of the education services provided. The Dakar framework was, in fact, a reiteration of political agreements signed during the Jomtien Conference in 1990 when 155 countries pledged to fight for education that can guarantee children the acquisition of knowledge and skills necessary for full participation in the life of society. The political will toward results-oriented education has been prominently displayed on the policy agendas of the OECD member countries (Cochran, 2001; Healy and Istance, 2001). Most recently, the participants of the G8 summit held in St.Petersburg in July, 2006 also pledged to "promote better understanding of foreign qualifications and educational outcomes" and to "actively cooperate to achieve high quality basic education"<sup>5</sup>, which clearly shows the commitment of major industrialized nations to the cause of promoting quality education.

### **3. Evidence of growing interest in learning assessments: more participants**

The general trend towards evidence-based policies illustrated by growing demand for monitoring learning results has been led by developed countries. Great Britain, France and the US pioneered national assessments in their countries thus giving birth to two general models of assessment – the British model which involves the whole population of relevant age cohorts in assessment and the American model which is primarily a sample-based one (Kellaghan and Greany, 2001). It was developed countries who not only developed their

---

<sup>5</sup> <http://en.g8russia.ru/docs/12.html>



national systems of assessment but also got engaged in designing external assessments such as TOEFL, GRE, GMAT, IELTS, FCE, CAE, CPE and others that are now globally recognized as tests which are required to be passed in order to meet the minimal requirements for entering a western university. These exams are designed to help select the best and the brightest and to weed out the underperforming students.

The accelerating pace of popularity of assessments in developed countries has spurred the cultivation of culture of assessing learning outcomes elsewhere. This is evidenced by the stunning growth in the number of countries participating in learning assessments on the national and international scale. While normally around 20 countries participated up to the 1980s in international assessments, the IEA Reading Literacy Study attracted 32 countries in 1991. In 2003, 52 nations took part in TIMSS and 41 – in PISA (30 OECD countries and 11 “partner” countries). As regards PISA, the survey was implemented in 57 countries in the 3rd assessment in 2006 and 62 countries have signed up to participate in the 4th assessment in 2009<sup>6</sup>. Moreover, all three major studies of student achievement have been conducted with a focus on monitoring performance over time and are now administered on a cyclic basis and are usually referred to as “trend” studies. Thus, not only the number of countries has increased over the years but also the nature of assessments has assumed more constancy (World Bank, 2006<sup>7</sup>).

According to Kellaghan (2001), until the 1990s, national assessments of student achievements were rare outside industrialized world, though some national surveys were carried out periodically before this time. The largest increase in national assessments happened after the Jomtien Declaration<sup>8</sup> and was generated conjointly by UNESCO-UNICEF, which in their Monitoring Project<sup>9</sup> have been responsible for initiating national assessments in about 50 countries. The most notable example of a large increase in popularity of national assessments is the system of assessments in Latin America and the Caribbean where during 1950s and early 1960s there were mainly isolated projects designed by local scholars, who often worked with North American researchers and were funded by international agencies. By 1997 practically all Latin American countries had some form of national assessment in place (Casassus, 1997). Nowadays, at least fifty nations carry out national assessments and the number is growing (Greany & Kellaghan, 1996).

#### **4. The popularity of assessments: why?**

As demonstrated above, learner assessments have become popular: they enjoy great financial and political support and an ever-increasing number of countries get engaged. But what are the reasons why assessments of all kinds are gaining in popularity and getting increased public support? The most cited factors that have led educational policy to be more evidence-based and to focus more intently on what skills and knowledge students acquire during their school studies are as follows:

<sup>6</sup> [http://www.pisa.oecd.org/pages/0,2966,en\\_32252351\\_32235907\\_1\\_1\\_1\\_1\\_1\\_1,00.html](http://www.pisa.oecd.org/pages/0,2966,en_32252351_32235907_1_1_1_1_1_1,00.html)

<sup>7</sup> National and international assessments of student achievement, 2006 Reviewer draft. Forthcoming paper.

<sup>8</sup> The declaration stressed that it was important to know to what extent students actually learned as a result of educational opportunities provided.

<sup>9</sup> The project was launched in 1992 with the goal of providing policy makers and front line implementers with the necessary conceptual and analytical tools and indicators to monitor the quality of their basic education programs in general and learning achievement in particular.



1. inadequate school learning: a concern that many children who spend a lot of time at school acquire sufficient knowledge and individuals' skills.
2. efforts to achieve greater efficiency: countries hope that a more detailed knowledge of their education systems will help them to efficiently allocate scarce resources.
3. trends to decentralize authority in the educational system, providing greater autonomy to local authorities and schools: when the tight grip of centralized system is weakened, coherent system of monitoring is viewed as necessary.
4. belief in the theory that human capital plays an important role in economic growth. Countries have been trying to focus on raising learning outcomes in the hope of increasing economic growth. There is some research evidence to support this approach, though it is not entirely consistent across countries or over time (Hanushek & Kimko, 2000; Ramirez, Luo, Schofer, & Meyer, 2006).
5. in the case of international assessments, the interest has been mainly due to comparative framework which could be utilized to assess student achievement in a country. By comparing results from different countries, assessment results can also be used to help define what is achievable, how achievement is distributed, and relationships between average achievement and its distribution.
6. international assessments have been of particular interest for their role in attracting the attention of the media and of a broad spectrum of stakeholders (politicians, policymakers, academics, teachers, the public). Comparative data provided in the studies will have more 'shock value' than the results of a national assessment. Poor results can encourage debate, which in turn may provide politicians and other policymakers with a rationale for increased budgetary support for the education sector.
7. studies of student achievement may also have a role to play in changing a culture of assessments
8. assessments may contribute to the development of local capacity in a variety of technical activities: sampling, defining achievements, developing tests, statistical analysis, and report writing.

The realization that assessments may contribute to greater efficiency within education systems, provide more accountability, help achieve strong and sustainable economic growth, especially in developing countries have brought educationalists to believe that educational policy based on concrete results of learning is the best option in today's world of scarce resources and growing pressure to stay competitive globally, as countries' economies become increasingly intertwined and interdependent. Now several decades have passed since learning assessments got a firm foothold in education policymaking for the first time. Have these years of continuous assessments and monitoring of educational quality yielded any results in the form of introduced improvements in response to interpreted assessment results? If yes, then what components of education systems were affected? The discussion of these issues continues in the next section.

### **III. Learning assessments: what is the impact on policy and existing school practice?**



There is no sense in carrying out studies of student achievement if their results are not disseminated to parties concerned, and subsequently are not acted upon to introduce improvements in education systems. Once collection and interpretation of results is completed, dissemination of information becomes the key point in predetermining a subsequent impact of assessments on education systems. Information which is distributed to the right audience and in such a way that it motivates all stakeholders involved has a better chance of effecting constructive changes throughout the education sector. However, the process of dissemination does not include just a simple distribution of results, but rather, it is a whole strategy aimed at achieving a certain effect – policy response. According to Schleicher (2006), a policy response is best effected through a two-stage process. The first stage includes procurement of the information and analysis of their policy implications. The second stage aims to raise awareness of the policy implications among the main stakeholders “in a manner that motivates the governments to respond with appropriate action” (Schleicher, 2006, p. 265). To strengthen the link between assessment results and policy impact which is reflected in actual actions of policy makers, the two-stage process needs to be complemented with understanding how governments respond to assessment findings and what leads them to develop and realize policies. The foregoing components are essential for making assessments change school practice for the better and iron out any other imperfections in education systems.

In addition, to maximize impacts of assessments, the potential benefits of assessments need to be studied by governments with a view to identify the most relevant ones and understand what they want to focus on before committing to an assessment. The advantage of this approach is that careful examination of potential benefits will ensure that resources are spent in a more efficient way and the well-designed targeting of resources will ensure that the set goal are reached. Typically, the following potential impacts are taken into account (Ross and Genevois, 2006):

- assessments can help pinpoint certain elements of a national education system that pose problems or are unusually excellent because they differ from other ‘similar’ countries;
- benchmarking embedded within assessments can help policy makers to compare their education systems against best practices or high performance in other countries;
- longitudinal studies of educational environment and outcomes of schooling may be useful in providing information on progress or declines in one or many countries;
- interpreting assessment results in terms of factors related to the conditions of schooling can be used for developing more effective approaches to school organization and resource allocation.
- assessments can encourage productive policy debates and participation in debates can help policy makers to clarify issues and highlight successful practices.
- studies of student achievement may help tweak teaching methods, instruction, teacher and professional training, curriculum, assessment practices to best suit the needs of an education system

Up till now, a great deal of assessments has been carried out worldwide and the number is growing. As will be shown further, many of them have affected educational planning, policymaking as well as school practice. In other cases, negative effects of learner assessments have also been witnessed. What follows is the review of the impacts made by monitoring-oriented (low-stakes) learning assessments on the international and national



levels. A special treatment is given to the subject of public examinations as a special kind of assessment which is drastically different from other summative assessments in many parameters such as purpose, duration, stakes involved, scoring, usefulness for further monitoring, etc. The analysis of impacts would not be complete without also considering school-based assessments and their influence on school systems. This form of assessments is of value, as it predetermines the final score given at the end of summative assessments. In addition, formative assessments accompany children throughout their schooling on a continuous basis, whereas summative ones gain importance at transition points in schooling.

### 1. Impact of international assessments

International assessments such as TIMSS, PISA, PIRLS, SACMEQ, PASEC, etc. have been instrumental in bringing about policy changes in some countries leading to the transformation of certain areas of education systems as well as individual schools.

The evaluation of PIRLS (2001) and TIMSS (2003) in low and middle income countries conducted by the World Bank<sup>10</sup> has shown that PIRLS and TIMSS have made a difference for policy makers in understanding education in their countries, particularly in terms of knowledge of student achievement levels, curriculum emphasis and education resources. The recognition of benefits of assessments has translated into concrete actions aimed at carrying out improvements in the education system, such as planning of certain initiatives and introduction of changes in response to the findings of previous studies of literacy in 1991 and TIMSS in 1995 and 1999. Policy initiatives and actions to improve teaching and learning that followed as a result of assessments have largely targeted the following (Gilmore, 2005):

- a) *Education system/Curriculum*: in some countries teams were set up to develop reading comprehension skills; to balance narrative and information texts. Expert groups were established for literacy development with a view to make changes in mother tongue curriculum, especially for the reading area. In Romania TIMSS results are said to have led to substantial curricular changes: new topics of “statistics and probability”, “data analysis and representation” were added to the math curriculum. A stronger emphasis was put on problem solving. In science curriculum, the introduced changes shifted attention towards more emphasis on practical investigations, relocation of topics and scientific inquiry. Curricular changes have been reflected in the new issue of Teacher Guides for Science and Chemistry and new books for students in math and science, which were written, taking into account the TIMSS findings. Slovenia carried out reforms extending the length of compulsory primary education in order to improve the acquisition of literacy.
- b) *Teaching methods*: as far as teaching methods are concerned, a variety of texts was used to check for consistency with PIRLS texts. In addition, a host of seminars and workshops were held to disseminate the PIRLS results and to introduce new strategies for teaching reading including teaching in item and test development. In Malaysia assessments contributed to an increase in practical work load and experiments in science lab and introduced science as a subject in grade 1 in 2003.
- c) *Assessment practices*: there was substantial influence on assessment practices in that tests of reading comprehension were developed along with marking schemes. Following the pattern of international assessments, the main purpose of assessment

<sup>10</sup> Gilmore, A. (2005). The impact of PIRLS (2001) and TIMSS (2003) in low and middle-income countries. World Bank, Washington, DC



was reviewed and changed from summative assessment to the one that places emphasis on using assessment information to provide feedback to the child and to plan learning experiences which meet the child's learning needs. Following assessments, Malaysia strengthened the methods used in assessing students at the school level by using school-based assessment system. TIMSS assessment in Slovakia provided the country with the testing, sampling methodology and new practices for preparing the national examinations. In Macedonia, TIMSS conducted in 1999 provided empirical data for mathematicians, teachers and authorities helping and helped them to develop skills and understand the processes in large-scale assessment.

- d) Teacher training and professional development:* IEA Secretariat and Consortium makes sure that its assessment programs are as participatory as possible and are executed in the way that maximize the training opportunities for national coordinators of assessments and other educational staff. This being the case, the international assessments have involved a large number of research workers and educators, who had the opportunity to learn about the aims and procedures of assessment projects and to contribute to the implementation of the projects. Namely, over the course of learning assessments educators have been introduced to developing the assessment framework, sampling, item development, translation and translation verification, piloting, data entry, data management and data cleaning including quality control, reporting of results. The knowledge acquired directly from the international assessment organization is further transferred to meet national assessment needs. For example, Slovakia voiced its recognition of the fact that TIMSS has provided new methodology, and new practices for the preparation of the national exams. In Bulgaria, TIMSS results have played a major role in the national evaluation of curriculum reform and “are now considered as part of the national evaluation system” (Gilmore, 2005, p.33). In Lithuania, extensive seminars were organized country-wide for primary school teachers and educational officers on the PIRLS results. It was the first experience of such kind for teachers, which allowed them to get in close contact with the international study, gave them an opportunity to analyze the results and to “have a hands-on training in defining reading comprehension processes” (Gilmore, 2005, p.32). This had a direct impact on teaching methods and the content of reading materials for primary school teachers. Assessments carried out in Bulgaria brought about a change of attitude of some education officers toward assessment bringing them to the realization of how important it is to have a national assessment program. In Slovakia, PIRLS results contributed to the change of attitude of policy makers. Before the PIRLS results were released, government provided no financial support to cover national costs. After the results were made public, the Ministry provided support in the form of producing reports, running workshops and covering costs for three international projects. In Romania, the impact of TIMSS on teacher training was quite substantial. Seminars were conducted for Romanian teachers and released items were made available to teachers, along with the international ranking and scoring guide for each constructed response item.
- e) Funding:* TIMSS findings in some countries have induced government to reallocate funding. For example, in South Africa poor achievement results from TIMSS led to increased allocation of resources for science and mathematics (Reddy, 2005).

According to Gilmore (2005), countries that have conducted TIMSS for the first time tend to lag behind the counties that have participated in both 1995 and 1999 in the overall volume of



accumulated benefits, such as increased awareness of importance of collecting internationally comparative data about countries' education systems, greater depth and breadth of capability in assessment and greater potential to make use of the findings for informing educational practices and policies. It is worth noting that in some cases educational changes were initiated top-down (e.g., by designing new curricula) as well as bottom-up (e.g., involving teachers). In general, participation of countries in the IEA assessment studies produced multiple positive impacts and brought recognition to the international assessments at governmental level which is evidenced by the allocation of funding to support assessments, the establishment of dedicated research institutions, the increased investment and focus on educational reforms. It is evident from the foregoing summary of impacts that the general effect of participating in PIRLS and TIMSS has been substantial, multi-faceted and positive. It is also evident that many of the educational reforms would not have occurred without these international assessments.

The impact of PISA has been strong in prompting national policy initiatives. According to European Network of Policy Makers<sup>11</sup>, in countries ranked as average or below on PISA, an agreement has been taken to consider the results seriously. In those countries initiatives for reform and development as a consequence of the PISA results were formulated very quickly. Extensive initiatives can be found, especially in Luxemburg, Germany and Norway. The measures that have been taken are summarized in Table 1 below:

**Table 1. Summary of impacts of PISA in Luxemburg, Germany, Norway**

<b>Luxemburg</b>	<b>Germany</b>	<b>Norway</b>
1. improvement of reading comprehension	1. improvement of reading comprehension	1. Development of a comprehensive strategy for better teaching and learning math and science
2. Improvement of pupils' competencies with an immigration background	2. Development of a comprehensive strategy for better teaching and learning math and science	2. Development of a comprehensive strategy for the stimulation of pupils' enthusiasm for reading
3. Improvement of the recruitment of teachers in post-primary education	3. Improvement of pupils' competencies with an immigration background	3. Increase in the lessons in primary education
	4. More focus on output in relation to input	4. More focus on output in relation to input
	5. Development of standards that describe what competencies pupils should have at the end of secondary education	5. Development of a national system of evaluation
	6. Development of national tests of learning outcomes in basic subjects	6. Development of national tests of learning outcomes in basic subjects
	7. Establishment of programs of whole-day education in primary education	7. Establishment of national centers for competencies for teaching and learning of basic subjects
	8. Fostering of research activities in education and didactics	

*Source: European Network of Policy Makers*

Other countries had already started reforms in the past, so the PISA results were used as a backup for the ongoing educational programs and reforms (e.g. Spain). In the majority of the countries which ranked above average, PISA did not lead directly to initiatives for reforms on the national scale but rather, the testing results have reinforced current policies (e.g. solving the problem of educational inequities and poor achievement in key subject areas, using new

<sup>11</sup> <http://cisad.adc.education.fr/reva/pdf/ReportingEvaluation.pdf>



syllabi in the major subject areas, such as reading, math and science, increasing the number of science subjects in post-primary schools in Ireland). In addition there are accounts of indirect consequences, for example:

- the studies caused an increased awareness for the necessity of reform and an continuous monitoring of the output (Austria)
- considering reforms of the curriculum and of testing (Scotland)
- in-depth analyses on the results of immigrant children to offer better insights (Sweden)
- the insight that the evaluation system is not sufficiently used by policy makers (France)

### *Influence on schools*

According to European Network of Policy Makers, approaches to disseminate information about the PISA results to individual schools have, for the most part, been very similar in the countries for which the Network prepared case studies<sup>12</sup>. Mostly, the dissemination of information was implemented from the top on down, as Ministries or state-run institutes summarized the main findings and tried to reach schools through publishing the findings in booklets, on websites or through presentation material. Afterwards, teacher unions organized conferences to discuss PISA results or researchers actively participated in information dissemination. On the whole, there are no indications about, if, or how PISA has influenced what is going on in schools. In some answers it is assumed that the impact is negligible (France) or that immediate influence of international evaluation is very limited because there are no developed feedback strategies for schools (Austria). Few answers give report of impact. For example:

- a feeling of resignation at the very beginning (Luxemburg)
- the reaction of schools that the media, etc., painted a picture of the school which was too negative (Norway)
- the high acceptance of the outcome responsibility of schools (Austria)

In Russia, over the course of four years (2000-2004) the significance of PISA results has been recognized by the Ministry of Education and Science. Several reports were commissioned and discussions launched as to the cause of Russia's failure in teaching skills to use acquired knowledge in real life situations. The first PISA results drew a lot of attention from policy makers, educational professionals and mass media and caused an uproar in the country which prides itself in its education and raising a generation of top-notch engineers who put a man in space. This being the case, first assessment results dealt a huge blow to the country's pride and led to the reassessment of its standing in the world of knowledge. In the run-up to the G8 summit in St.Petersburg, which Russia hosted in July, 2006, Russian policy makers have come up with an initiative to develop a new assessment tool which would be better suited to the needs of developing nations. Partly in response to the PISA outcomes, Russia responded with the launch of the ambitious National Project "Education", which, among other things, is aimed at raising student achievement levels in two-year period.

Schleicher (2005) names the following impacts of assessments mostly in the OECD countries. In the case of such a top performer on PISA as Finland, assessment findings allowed for an in-depth analysis of the threats to current strengths of the education system and led to attracting, developing and retaining of best-performing teachers, maintaining flexibility of the education system and responding to increasing diversity in the student body; ensuring

<sup>12</sup> See a more detailed description of case studies of impacts of assessments in Annex 1.



strategic support and financing of schools. In response to PISA results, the National Board of Education set up a national and joint Nordic program to enhance literacy skills of weak readers, especially boys. Also the Ministry of Education took measures to improve literacy and language skills of immigrants.

In Japan, PISA caused a national debate and improved student motivation and engagement with learning. Denmark has launched an OECD-led comprehensive review that resulted in specific recommendations; introduced assessments of educational progress at key stages; turned PISA into an asset for professionalism. In Germany PISA has led to a 30% increase in federal funding for education. The Lander ministers of education have moved to establish “standards” in the form of educational goals for specified grade levels. At the same time, ministers have established an agency responsible for working out tests in order to supervise the standards. This agency is an institute at the Humboldt University. As a result of PISA, high schools have changed from a 13-grade to a 12-grade system. With each new round of PISA results, the discussion resumes about the segregation of 10-year-old children into different types of schools.

In response to PISA results, Italy came up with the first coherent action plan to address a large regional and between-school variation in the quality of learning outcomes; established national and regional task force aimed at individualizing learning, teacher professional development, school leadership and accountability/autonomy. Almost in all participating countries PISA contributed to changes in national assessments, as nations tried to adjust their own assessments to international standards and criteria shifting focus from rote learning to application of acquired knowledge in real life situations.

The international study of student achievement SACMEQ held in 1995 and 2000 produced a series of recommendations for policy makers. In response to SACMEQ results, the Tanzanian national report recommended that the government investigate gender disparities in school enrollment and identify options to help eliminate the gender gap. SACMEQ results have featured prominently in presidential and national commissions in Zimbabwe and Namibia and national education sector studies in Zambia, etc. In some countries, the results were interpreted to indicate a need to provide standards for resources in education. For example, benchmarks for the provision of classroom facilities were introduced in Kenya. In Zimbabwe, special funds were provided for classroom supplies. In Kenya, SACMEQ findings on gender, regional disparities and internal inefficiencies were used to guide the development of action plans to implement Education for All at national, provincial, and district levels (Murimba, 2005a). Two countries, Namibia and Kenya, are worth dwelling on, as there is substantial evidence of SACMEQ impacts in these countries (Nzomo and Makuwa, 2006).

*General reaction to the results:*

In Kenya, the SACMEQ results have alarmed educationists, when it was found out that female teachers are underrepresented in education in the North Eastern Province. Also it was observed that female reading teachers in North Eastern Province constituted only 28% and 33% in 1998 and 2000, respectively. The discrepancies found in the performance of Kenya’s pupils in Grade 6 on the SACMEQ reading and mathematics tests and the results of the national examination at the upper end of primary school led to a widespread debate among policy makers regarding the performance standards that were supposed to be established by Kenya’s examination system. Although the results of SACMEQ in Namibia showed major improvements in resource allocation to schools, disparities still remained among regions in



terms of inputs to schools. This situation raised concern that despite some success in providing formal access to education, the issue of access to quality education was unresolved.

*Education system/Curriculum:*

In Namibia, the SACMEQ assessment led to some interventions in the form of new programs (English Language Teacher Development Program, the expansion of the Basic Education Teachers' Diploma, the Efficiency Program, the Management Policy Coordinating Committee including the establishment of the Presidential Commission of Inquiry on Education and Training). Further debates resulted in a five-year strategic plan. Deteriorating quality of education in Namibia and poor conditions of schooling formed the foundation for the establishment of a national inspectorate to monitor the quality of education in the country based on national standards. SACMEQ research results prompted the division of schools into inspection circuits and into clusters for administrative and support services. These measures helped provide multi-level assistance allowing schools to share resources, good practice and expertise within a cluster.

In Kenya, SACMEQ I results led the ministry to design a comprehensive set of acceptable standards to be followed by schools to ensure efficiency and effectiveness in the education system. In addition, the assessment results and education reviews introduced the tradition of continuous curriculum review. The Kenya Institute of Education was directly involved in the review of the curriculum to make it more affordable to parents along with raising quality and relevance. The number of subjects was reduced along with examinable subjects.

*Publications and research:*

The findings of SACMEQ I project in Kenya were extensively used for the World Bank Education Sector Analysis for the country and led to the development of proposals that targeted deficiencies in the country's education system. The Education Sector Preview led to the development of the Education Sector Strategic Plan. The SACMEQ were further used in financial planning.

*Teacher development:*

In Kenya, lower than expected achievement levels and the lack of textbooks prompted the government to develop a school-based teacher development program.

The Program on the Analysis of the CONFEMEN Education Systems (PASEC) was first launched in 1991 and since then has helped to analyze key factors that can explain variations in educational quality. PASEC's impact has been rather multi-faceted: it tried to inform policy debate and provide sound arguments to convince the stakeholders. In Senegal, the assessment results drew a lot of attention from the print and electronic media and prompted the organization of seminars to discuss the results. In Togo and Guinea PASEC evaluations mainly focused on teacher training and non-civil-servant teacher contracts (Bernard and Michaelowa, 2006). PASEC results have also encouraged the Togolese authorities to retain the BEPC (lower secondary attainment) as the minimum academic entrance requirement for primary teachers. In Cameroon, a special evaluation unit was created to ensure the institutionalization of educational evaluation.

Over the years major international assessment tools such as PIRLS, PISA, TIMSS, SACMEQ and others have been changing educational systems worldwide. Their impact was felt in curricula, teaching methods, assessment practices, teacher training and professional development, funding in the countries where they were carried out. International studies of



student achievement revolutionized the notion of quality of education and brought about institutional reforms in some countries and led to targeted support in others.

## 2. Impact of national assessments and public examinations

On the national level, assessments typically take two forms: monitoring-oriented assessment which is used primarily for collecting data on student achievement and selection/certification-oriented assessment conducted typically as a national examination. National systems of assessments were created to achieve greater transparency in the outcomes of educational management and practices in schools and to provide professionals with useful tools to improve their practices. National assessments in both forms provide important feedback on what students learn and what they fail to learn. Thus, mere data become substantive information on students' educational achievement, which may be used to introduce improvements or positively impact education systems. The impact of national assessments on education systems depends on those responsible for dissemination of information and its interpretation, reorganization and presentation to all parties involved. Among other factors that affect the prospects of the data having an impact are the time periods within which results are provided and the regularity with which they are delivered. Participation of local actors in the assessment process, the quality of the reports and political will to make informed decisions also play an important role in the strength of the impact assessments are likely to make. The analysis of data on national assessments shows that over the years they have brought about changes in the following educational components:

- a) *Targeting of scarce resources*: some countries used results of national studies of student achievement to detect learning problems and to direct funding along with other resources to help deal with deficiencies. For example, in Chile, the national assessment helped identify 900 schools in need of improvement. These schools were given a supply of teaching materials, textbooks; teacher development activities were also established through workshops. Following these measures, there is evidence of achievement gap between the schools and other schools diminishing over time (Benveniste, 2000b). Ogawa and Collom (1999) point out that Chile test results led directly to the design of a compensatory program to provide the schools that scored lowest with textbooks, assistants and teacher training. Schiefelbein (2003) notes that in 1997 testing data helped convince the Chilean Congress to approve the resources for expanding the annual school schedule by 40% in order to catch up with the quality deficit.
- b) *Impact on classroom teaching*: assessments have contributed to the expansion in the content taught and helped to at least partially eliminate the practices of some education systems "to teach to the test". A good example is Thailand, where a national assessment program was established which helped to widen the scope of study content beyond what student needed to pass a public examination. In other countries assessment projects has also affected teaching: China used its assessment to provide for learning needs of students who were not planning on entering a university. The Uruguayan domestic assessment UMRE made a big impact on teachers' acceptance of testing results, as teachers appear to place a high value on the national assessment, which is demonstrated through their requests that UMRE provide
- c) *Impact on textbooks and curriculum*: Learning assessments have played a key role in providing valuable information to writers of textbooks and guides for teachers. For example, in Chile SIMCE findings were used to detect problem areas and to deal with these through new teaching material. Moreover, some national assessments had a major



impact on curricular changes, as happened in Thailand when the results of the national assessment indicated that students were performing poorly in science and math. The Department of Curriculum then used the quantitative data collected through assessment procedure “to revise the curriculum in order to emphasize process skills” (UNESCO, 2000, p.58). In Brazil, the results of national assessments have led to some curriculum revisions and pedagogical innovations as well as changes in human resource training and policies regarding school funding. In response to assessment findings, Kenya included a much broader spectrum of cognitive skills as compared to the past and made efforts to promote teaching and learning of competencies. The assessment led to the shift from just measuring students’ ability to memorize factual information to testing of comprehension and application of knowledge (Greany & Kellaghan, 1992).

- d) *Public support for education:* national assessments have contributed to raising public awareness of problems in education. In Chile SIMCE results were publicized by media which led to publishing manuals to explain results and provide information for parents to make informed decisions about which school to send their children to. The popularization of the program in the mass media have secured a large degree of public support for the education sector. Kenya and Sudan are also among the countries that enjoy public support for education as a result of assessments that allow parents to compare schools. In Colombia, national test (SABER) results were distributed to the political and technical leaders in the education sector and widely publicized, which resulted in organization of interpretative workshops and other public events.
- e) *Teacher training:* Argentina has held its assessment and came up with a number of reports which were then disseminated among teachers, principals and school administrators. The communication of assessment results to professional audience has helped to identify school deficiencies, and to provide recommendations for dealing with them. These recommendations were even incorporated into pre-service and in-service teacher training courses. Dissemination of the national assessment results in Uganda was designed to help education authorities to come up with their own strategies to apply to worsening education system. The results distributed through a series of workshops and seminars prompted a consideration of their usefulness and value for teaching and their possible usage to adapt the process of instruction to improve learning outcomes. Uganda National Examination Board expected that the distributed results would induce teachers to shift their emphasis in how they presented their curriculum material and to pay more attention to the areas of knowledge and skill that were identified as weak during the national assessment. In Ecuador, APRENDO results led to the development of a proposal on teacher training curricula on the basis of APRENDO results (however, this proposal failed to bear fruit later on). APRENDO results were useful to private schools, which used the data for institutional marketing purposes. Overall, APRENDO assessment significantly contributed to the development of local assessment capacities and provided data on academic performance that had not been studied before in Ecuador. The assessment program in Honduras (UMCE) has contributed a great deal to developing and training technical and professional staff and to developing a psychometric model of high pedagogical value, particularly in the light of the lack of robust and updated curricular referents at the national level.
- f) *Policymaking:* Although SIMECAL (Bolivia’s national assessment) results have had only a limited impact on the policy agenda, it brought numerous experts in the assessment to the agreement that the SIMECAL data should be used as one of the main criteria in deciding on the best sectoral (including curricular and pedagogical) and intersectoral policies. The Guatemalan assessment PRONERE made it possible to develop local technical capacities in educational assessment. In addition, some private



sector stakeholders have mobilized to demand that the assessments continue and their results be used effectively for policymaking.

As far as selection/certification-oriented assessments are concerned, the subject of public examination merits special treatment, as in many aspects it is different from other types of assessments mentioned at the beginning of the paper. The key difference lies in the purpose of public examinations - to select the most able on the basis of merit and certify achievements of individual students, which usually involves high stakes for participants as opposed to national/international and regional summative assessments where the primary purpose is to monitor and assess the performance of the whole educational system (or part of it) and where the stakes involved are low, as measurement results are mainly collected for providing feedback to policy makers and do not expose individuals to risks of being barred from getting a diploma, moving on to the next level of study or getting the desired job. These fundamental differences in the level of aggregation of data and use of information predetermine characteristics of public examinations and their likely impact on education systems.

Public examinations have been shown to have an impact on what goes on in the classroom in terms of “what” and “how” teachers teach and students learn (Gipps, 1994 & 1996 in Little and Wolf; Black, 1998; Greaney and Hasan 1998). They have been considered to be most useful in highly competitive situations, as they allow “greater standardization of tasks and conditions and, hence, greater comparability of results” (Crighton, 2003, p8). They are well-suited to serve their purpose of selection and certification and their focus on individual students makes them indispensable when trying to monitor individual student achievement rather than that of the education system as a whole. There has been a growing interest in the use of examination results to monitor the effectiveness of schools and to hold them responsible for their performance, as is the case in England where there is a reporting system in place which uses examination and national test data toward this end (Crighton, 2003). In addition, an examination can be a great means of ensuring that the new courses are introduced or the new subject matter is taught if it is well-matched to the curriculum development. For example, in France the diversification of the *bac* helped to introduce technological courses in the lycees and the diversification of the curriculum. Also, exams can well serve the function of clarifying and maintaining standards. In some cases, the examination creates a climate for pedagogical innovation (for example, in the US the introduction of “authentic assessment” has encouraged new teaching and learning strategies). Some countries, such as Chile and Mexico, have used results of high stakes assessments as complementary indicators for teacher incentives. In Chile, this has been introduced through the National Performance Assessment System for subsidized schools, which provides bonuses annually for the teaching staff in schools that raise the academic performance of their students relative to previous assessments. In Mexico, teachers can file an application to the Teaching Career Program, a system of wage incentives that uses national test results as an indicator of professional performance.

However, examinations have a downside as well as benefits: their role in monitoring learning outcomes on the system level and improving quality of education has been rather limited. High stakes attached to most public examinations put a lot of pressure on everyone involved thus resulting in distorting teacher instruction through “teaching to the test” and causing the so-called “backwash effects” (Griffin and Nix, 1991). High-stakes testing has been blamed for problems in the recruitment and retention of teachers. Teachers tend to abandon schools in which student performance is poor, and try to get into schools in which they will receive rewards for good performance. Being based on the belief that knowledge is absolute, adhering to strict deadlines and marking procedures, public examination determines what is taught and



leads teachers to drill students for examination and encourages rote learning. These effects on school practice are far from positive. Declarative knowledge and out-of-context learning becomes more important than life skills. According to Satterly (1989), negative impacts of public examination include giving false impression of the status of learning in school, being biased against some minority or cultural groups of students; spending enormous time in preparing and administering tests – all of these cause much damage to the curriculum limiting it to only what is measured during the test rather than what is really needed for the future.

Claims of negative effects of public examinations on curriculum development are backed up by abundant evidence. In Colombia, for example, there have been attempts, as part of the movement for Renewal of the Curriculum, to extend curricular objectives beyond simply cognitive skills to affective domain. However, existing ICFES examinations cultivate the culture of “teaching to the test” and do not allow for curriculum expansion. There even exist published texts with examples of questions from the ICFES examinations and advice on how to succeed in the examination (Becerra Ed, 1991, 1992), which openly encourage the practice of focusing solely on passing the examination.

According to Christine De Luca (1994, p.43), in Egypt, the national examination Thanawiya ‘Amma has “emphasized memorization of facts and this has affected both the curriculum content and methods of teaching and learning. Most teaching could be accurately described as lecturing and most learning as rote learning.

The situation is similar in many countries (France, Japan, Scotland, etc): curriculum along with curriculum development are closely aligned to the examination system and are limited by it as to the content taught.

Continuing the list of drawbacks of public examinations, Amrein and Berliner (2002) point out that as there is no clear evidence for positive impact of high-stakes assessment on student learning (based on their analysis of 18 US states), there is a need for transformation of existing high-stake policies (Sahlberg, 2006).

However, despite the shortcomings, it cannot be said that public examinations are utterly useless and should be substituted for. They, nevertheless, play their narrow but important role in providing certification and selecting students for future studies or work rationalizing the distribution of scarce places.

Given the distortions and limitations described above, it is all but surprising that most countries tend to employ two (or more) parallel systems of assessment, as is the case in France or Russia – one for selection and certification which is best served by a public examination; the other one is either a sample-based international/national or regional summative assessment, which provide better conditions for monitoring aggregated student outcomes and are considered low-stakes which effectively does away with limitations of a public examination and proves to be less costly. The necessity of having a parallel system of assessment that would complement public examinations is reinforced by the fact that examinations are likely, by their nature, to remain subject-bound, whereas national/international assessments (in Thailand and Chile) have shown tendency to develop beyond cognitive to affective and other domains (Himmel, 1996; Prawalpruk, 1996) and are more attuned to measuring higher order skills.



Formative assessments, which accompany pupils throughout their school studies and are tightly coupled with public examinations, have also been found playing their role in impacting school systems in various countries.

#### **IV. Concluding thoughts**

The explosion of interest in assessments of learning outcomes in recent years has been prompted by economic concerns as opposed to the past, when other interests were actively pursued. The greater preoccupation with learning outcomes has come to dominate the educational policy agenda bringing policy makers, education authorities and a broad spectrum of other stakeholders to act upon findings supplied by national, regional and international systems of assessment. These action plans have effected significant changes in various aspects of education systems, such as curriculum and curriculum development, teacher professional development, assessment practices, teaching methods, targeting of funding, resource reallocation, and public support for education. Some assessments, mostly high-stake ones, have been acting as a restraining force on education systems preventing their further development and imposing considerable limitations. Although assessments have contributed, to a large extent, to changes in education systems and school practice around the world, it seems premature to assess the impact precisely. The fact that learner assessments have prompted countries to react to findings does not signify that these actions have necessarily led to positive results, such as higher learning outcomes or better quality of education, though some scarce evidence pointing to that is available (Benveniste, 2001). The results of country-led actions in response to assessment findings are in need of further monitoring and research.

In addition, it is important to add that assessments discussed in this paper include mostly standard forms of learner evaluation including formal school examinations, university entrance exams or sample-based national/international assessments. However, recently there has been growing pressure to pay more attention to alternative forms of assessments in order to reflect new approaches to teaching and learning. These forms of assessments include portfolio assessments, formative, on-demand competency and computer-based forms of assessment. Will these drastically change learning practices around the world as 'traditional' ones did? This is a topic for further analysis.

#### **References**



- Amrein, A.L. & Berliner, D.C. (2002, March 28). High-stakes testing, uncertainty, and student learning *Education Policy Analysis Archives*, 10(18). Retrieved February 18, 2007 from <http://epaa.asu.edu/epaa/v10n18/>
- Becerra, R. (Ed) (1991, 2<sup>nd</sup> Ed 1992) *PasalCFES: La Clave para lograr altos puntajes* Bogota Educar Editores S A
- Bennett, N., Desforges, C. (1991). Primary education in England: a system in transition. *Elementary School Journal* 92:6178
- Benveniste, L. (2000). The politics of student testing: a comparative analysis of national assessment systems in Southern Cone countries. Doctoral dissertation, Stanford University
- Bernard, J., Michaelowa, K. (2006). How can countries use cross-national research results to address ‘the big policy issues’? In *Cross-national studies of the quality of education: planning their design and managing their impact*. Ed. By K. Ross and I. Genevois Paris: International Institute for Educational Planning
- Bethel, G. (2003). Examination and assessment reform: an overview of experiences in Central and Eastern Europe. Retrieved March 15, 2007 from [http://www.soros.org/initiatives/esp/articles\\_publications/publications/examinations\\_20030701/examinations\\_20030701.pdf](http://www.soros.org/initiatives/esp/articles_publications/publications/examinations_20030701/examinations_20030701.pdf)
- Black, P. (1998) *Testing: friend or foe? Theory and practice of Assessment and Testing*, London: Falmer Press.
- Casassus, J. (1997). Eleven issues challenging international studies. Rio de Janeiro: MEC-INEP-OREALC
- Cochrane, D. (2001). Economics and demographics in metropolitan communities: An argument for international equity indicators. In: W. Hutmatcher, D. Cochrane, N. Bottani. (Eds.), *In pursuit of equity in education: Using international indicators to compare equity policies* (pp. 343-372). The Netherlands: Kluwer Academic Publishers.
- Cohen, M. (1990). Key issues confronting state policymakers. Pp. 251-288 in *Restructuring schools: The next generation of educational reform*, edited by R. Elmore. San Francisco: Jossey-Bass
- Coombs, P. (1968). *The World educational crisis*. Oxford: Oxford University Press
- Crighton, J. (2003). Assessment: general overview of international practice. Retrieved March 23, 2007 from [http://www.soros.org/initiatives/esp/articles\\_publications/publications/general\\_20030801/overview\\_20030801.pdf](http://www.soros.org/initiatives/esp/articles_publications/publications/general_20030801/overview_20030801.pdf)
- De Luca, C. (1994). *The impact of examination systems on curriculum development: an international study*. Paris: UNESCO
- European Network of Policy Makers for the Evaluation of Education Systems. (2004). Reporting evaluation findings to policy makers. Retrieved April 5, 2005 from <http://cisad.adc.education.fr/reva/pdf/ReportingEvaluation.pdf>
- Ferrer, G. (2006). Educational assessment systems in Latin America: current practice and future challenges. Washington, DC: PREAL (Retrieved April 5, 2005 from <http://www.preal.org/Archivos/Bajar.asp?Carpeta=Preal%20Publicaciones%5CLibros&Archivo=Ferrer%20assessment%20systems%20-%20final%20with%20covers.pdf>)
- Gilmore, A. (2005). The impact of PIRLS (2001) and TIMSS (2003) in low and middle-income countries. World Bank, Washington, DC
- Gipps, C. (1994) *Beyond Testing: towards a theory of educational assessment*, Lewes: Falmer Press (reprinted 1998).
- Gipps, C. (1996) ‘Assessment for Learning’ in Little, A. & Wolf, A. (eds.) *Assessment in Transition*, Oxford: Pergamon. Chapter 15, pp. 251-261.



- Greaney V. & Hasan, P. (1998) 'Public Examinations in Pakistan: A system in need of reform' in Hoodbhoy, P. *Education and the State: Fifty years of Pakistan*, Karachi: Oxford University Press.
- Griffin, P. & Nix, P. (1991). *Educational assessment and reporting, A new approach*. New South Wales: Harcourt Brace Jovanovich.
- Gvozdeva, S.I. On USE results. Presentation at the seminar in Sochi.
- Hanushek, E., Kimko, D. (2000). Schooling, labor force quality and the growth of nations. *American Economic Review* 90 (December): 1184-1208
- Healy, T. and D. Istance (2001). International equity indicators in education and learning in industrialized democracies: Some recent results and avenues for the future. In: W. Hutmatcher, D. Cochrane, N. Bottani. (Eds.), *In pursuit of equity in education: Using international indicators to compare equity policies* (pp. 195-215). The Netherlands: Kluwer Academic Publishers.
- Himmel, E. (1996) National assessment in Chile, in: P. MURPHY et al National Assessments: testing the system (Washington DC, World Bank).
- Kellaghan, T. (2001). The Globalisation of Assessment in the 20th Century. *Assessment in Education: Principles, Policy & Practice*. 8(1):87-102.
- Kellaghan, T. & Greaney, V. (1992). *Using Examinations to Improve Education: A case study in Fourteen African Countries*. Washington, D.C.: World Bank
- Kellaghan, T. & Greaney, V. (1996). *Monitoring the Learning Outcomes of Education Systems*. Washington, D.C.: World Bank
- Kellaghan, T., Greaney, V. (2001). *Using assessment to improve the quality of education*. Paris: UNESCO: International Institute for Educational Planning
- Marzano, R. J. (2001). *Classroom instruction that works: Research-based strategies for Increasing student achievement*. Alexandria, VA: Association for Supervision and Curriculum Development
- Murimba, S. (2005a). The impact of the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ). *Prospects*, 35, 91-108
- Nzomo, J., Makuwa, D. (2006). How can countries move from cross-national research results to dissemination, and then to policy reform? In *Cross-national studies of the quality of education: planning their design and managing their impact*. Ed. By K. Ross and I. Genevois Paris: International Institute for Educational Planning
- Ogawa, R., Collom, E. (1999). "Educational Indicators: Instruments for Bounding Rationality or Grounding Legitimacy". University of California, Riverside.
- Paulson, L. F., Paulson P. R., & Meyer C. (1991). What makes a portfolio a portfolio? "Educational Leadership," 48(5), 60-63
- Phillips, G. (1991). Benefits of state-by-state comparisons. *Educational Researcher* 20 (3): 17-19
- Pravalpruk, K. (1996) National assessment in Thailand, in: P. MURPHY et al (Eds). National Assessments: testing the system, pp. 137-145 (Washington DC, World Bank).
- Ramirez, F., Luo, X., Schofer, E., Meyer, J. Student achievement and national economic growth. *American Journal of Education* 113 (November, 2006). Retrieved March 19, 2007 from <http://www.journals.uchicago.edu/AJE/journal/issues/v113n1/113101/113101.web.pdf>
- Reddy, V. (2005). Cross-national achievement studies: learning from South Africa's participation in the Trends in International Mathematics and Science Study. *Compare*, 35, 63-77.



- Rehmani, A. (2003). Impact of public examination system on teaching and learning in Pakistan. Retrieved March 5, 2007 from <http://www.aku.edu/akueb/pubexam.pdf>
- Ross, K., Genevois, I. (2006). Cross-national studies of the quality of education: planning their design and managing their impact. Retrieved March 19, 2007 from [http://www.unesco.org/iiep/PDF/pubs/G117.pdf?class=IIEP\\_PDF\\_enhp&page=Cross-national%20studies%20of%20the%20quality%20of%20education&estat\\_url=http://www.unesco.org/iiep/PDF/pubs/G117.pdf](http://www.unesco.org/iiep/PDF/pubs/G117.pdf?class=IIEP_PDF_enhp&page=Cross-national%20studies%20of%20the%20quality%20of%20education&estat_url=http://www.unesco.org/iiep/PDF/pubs/G117.pdf)
- Sahlberg, P. (2006). Education policies for raising student learning: the Finnish approach. Washington, DC: World Bank
- Satterly, D., 1989. Assessment in Schools. Oxford, UK: Basil Blackwell Ltd.
- Schiefelbein, E., Schiefelbein, P. (2003). From screening to improving quality: the case of Latin America. *Assessment in Education*, 10 (2), July 2003.
- Schleicher, A. (2005). Conference presentation on PISA results
- Schleicher, A. (2006). How can international organizations work with the media to manage the results of cross-national studies? In *Cross-national studies of the quality of education: planning their design and managing their impact*. Ed. By K. Ross and I. Genevois Paris: International Institute for Educational Planning
- Smith, M., & O'Day, J. (1991). Systemic school reform. Pp. 233-267 in *The politics of curriculum and testing: The 1990 yearbook of the politics of education association*, edited by S. Fuhrman & B. Malen. New York: Falmer Press
- UNESCO. (1990). Basic Education and Literacy: World Statistical Indicators. Paris: UNESCO
- UNESCO. (2000). Educational for All: Status and Trends 2000: Assessing learning achievement. Retrieved March 19, 2007 from <http://unesdoc.unesco.org/images/0011/001198/119823e.pdf>
- UNESCO. (2000). The Dakar Framework for Action. Education for All: Meeting our collective commitments. Paris: Author.
- World Bank. (2006). National and international assessment of student achievement. Reviewer draft. Forthcoming paper.

## **ANNEX 1. Impact of PISA: National Case Studies (adapted from the report by European Network of Policy Makers<sup>13</sup>)**

### **Austria**

<sup>13</sup> <http://cisad.adc.education.fr/reva/pdf/ReportingEvaluation.pdf>



*Reception by policy makers*

The National PISA centre organised had conferences and releases for the press, for the ministers, for the staff of the ministers and for educational interest groups. Ministers also have been informed by papers of their staff. These papers included the most important result and background information.

The overall reaction of the educational authorities to the PISA results was positive (see above). The ministry tried to influence the public opinion by taking over opinion leadership in conferences and the official website ([www.bmbwk.at](http://www.bmbwk.at)) which was, however, only partly successful.

Despite the tendency to emphasize the positive aspects of the PISA results, the international achievement studies also caused an increased awareness on the side of politics for the necessity of reform and a continuous monitoring of the output of the education system

- The poor results in TIMSS III gave way to a costly national initiative for improvement and innovation in mathematics and science teaching, lead by the University of Klagenfurt.
- Following PISA 2000, several initiatives to promote reading competence, especially in primary school children, were started.
- The new awareness of the need for reform resulted in the appointment of a commission of educational scientists to work out a proposal for a comprehensive reform initiative, which was published a few weeks ago and submitted to the public for a broad discussion.
- In the centre of this expert paper, the need for a continuous monitoring and evaluation of the results of education at the class-, school- and system level is stressed.

*Influence on schools*

The information of schools and teachers was mainly through the media. Besides this, schools became acquainted of the notion of “PISA”, in that until today almost every political argument in the educational field starts with the phrase “as PISA has shown ...”

The immediate influence of the international evaluation and achievement studies on schools is nevertheless very limited, as there are no developed feedback strategies to schools. However, the current stress on output-related evaluations strategies is gradually leading to a higher acceptance of the notion of “outcome-responsibility” of schools and teachers. More and more, schools and teachers also become accustomed to the possibility to have outcome evaluations not only at the system level but also in schools and classes, which has been a somewhat strange thought until recently.

## **Flemish Community of Belgium**

**Reception by policy makers**

We informed our Minister just before the international data release in December 2001 in order to ensure confidentiality of the results. We are thinking of changing this procedure a little for the PISA 2003 release. We would like to inform our Minister a bit earlier, in order for



him/her<sup>1</sup> to be able to understand the results better and in order to prepare better for possible policy actions.

The main reaction of the political forces was happiness with the results. In Flanders we do not have many output data for our schools, especially not in secondary education, and surveys such as PISA and TIMSS give us an idea about the effectiveness of the Flemish schools, also in an international context. The main topic for discussion was the difference with the French Community results and later on the differences within our own education system.

No concrete initiative for reform has been taken. At the Education Department we think that good results risk to put asleep instead of keeping awake.

### **Influence on schools**

The schools that participated in PISA 2000 each received their own school report, comparing their pupils' results with the results of pupils from similar schools (i.e. similar education programmes offered). The University of Ghent, that also implemented PISA 2000 in Flanders, prepared these school reports. The schools were very happy with the feedback they received. In fact they received two school reports in two phases: the first one only included the raw results of their own pupils; the second one consisted of the controlled data and the comparison with similar schools.

After that, all Flemish secondary schools received a publication containing the brochure with the PISA results of Flanders compared to the other participating countries; an explanation of the PISA 2000 framework; and the released items on reading literacy, mathematical literacy, scientific literacy and problem-solving. These items were put in a ring binder, so that teachers could use them in class.

The Education Department, the researchers of Ghent University, a panel of experts and an OECD representative also presented the PISA 2000 results in a seminar in Brussels, in which all secondary schools could participate.

## **England**

### **Reception by polity makers**

The absence of any concrete initiatives for reform as a result of international studies has been a cause for concern in certain quarters in England. It could be argued that PISA, for example, provided further evidence on areas of performance which we already guessed at intuitively or were aware of through other research and that the policies to tackle the issues highlighted by PISA were already in place. Certainly, England's mediocre performance in the last two TIMSS studies in 1995 and 1999 had some part to play in the establishment of the National Numeracy Strategy for primary age children in England.

### **Influence on schools**

*How have schools and teachers been informed about the results of the international studies?*

#### *Response*

In the case of the PISA and PIRLS studies, the Department for Education and Skills produced booklets summarising the findings for teachers and headteachers. Electronic versions of the



booklets were put on Teachernet, the Department's website for teachers and they were advertised in Spectrum, a monthly communication from the Department to schools. Copies of the booklets were also distributed amongst the teacher and headteacher unions who alerted their members to the findings. Two of the main teacher unions, NASUWT and NUT, held a joint conference on the PISA findings in January 2003 which was very well attended by teachers. Speakers from the Department (including the Minister of State, David Miliband) and OECD also attended.

*How did the results of international studies influence schools and instruction? What was the reaction of teachers?*

*Response*

We have no evidence to suggest that they have.

## **France**

### **Reception by policy makers and politicians**

*What is the reaction of the main political forces?*

Because for both PISA and PIRLS the French results are consistent with the results of national tests and because France's rank is either average or just above average, there was no specific party political reaction.

*Are there concrete initiatives for reform and development as a consequence of the international studies? Of what kind (e.g. standards, testing, curriculum)?*

One of the weaknesses of the French evaluation system is that it is not sufficiently used by policy makers to initiate changes. This is also the case for international studies. The policy makers responsible for the curriculum or teaching/learning policies do not normally act on them.

In the case of PIRLS, and following the presentation to the press, the national inspectorate decided to organise a meeting with the assessment directorate about reading at the primary level (such a meeting is unusual) in order to see what could be changed in the pedagogy or curriculum. Whether anything will happen in practice beyond that is another matter.

*How do the results of international studies influence schools and instruction? What is the reaction of teachers?*

No information is available on these points but it can safely assumed that the impact is negligible.

## **Germany**

### **Reception by policy makers**



The National PISA consortium organised several press-conferences for the press, for ministers, for the staff of the ministers and for educational interest groups. Ministers have also been informed by papers from their staff. These papers included the most important result and background information.

The educational authorities have been shocked by the PISA results. In 2002 the conference of ministers agreed on several measures from which the following activities have a very high priority:

- measures for the improvement of reading comprehension, not only in primary and secondary education but also in the pre-primary section
- development of standards for mathematics, German language, English as a foreign language, that describe what competencies pupils should have at the end of secondary education; these standards go along with new core curricula for the subjects mentioned above
- nearly all states in Germany started to develop large-scale assessment tests for the end of primary and secondary education
- measures for the improvement of students', with a migration background, competencies
- some states started programs for whole day education in primary education

Additionally pilot projects for the improvement of classroom instruction in mathematics and science have been established nation-wide. Also, a nation-wide initiative to foster the research activities in education and didactics started in 2002, in which more than 50 research projects are included.

### **Influence on schools**

Schools involved in the PISA study (219 schools in the international and 1466 schools in the national PISA study) got the results of their school in comparison to the national average. Results of single schools have not been published. Ministries, state -institutes and universities prepared presentations (e.g. power-point presentations) of the main results on internet. We don't have empirical evidence about the influence of PISA on schools, on individual teachers or on teaching methodology. We estimate that a lot of teachers feel a certain pressure because of the tremendous public reaction and the new concepts which the ministries recently started. Indeed, there is an interesting influence on people working in the scientific community: nearly all articles started with the phrase: 'As PISA showed...'

### **Ireland**

*Are there concrete initiatives for reform and development as a consequence of the international studies? Of what kind (e.g. standards, testing, curriculum)?*

Rather than suggesting the need to launch any new initiatives immediately, the outcomes of PISA 2000 have reinforced current policies. These include policies in the areas of

- addressing educational disadvantage and underachievement in reading through schemes such as those described earlier, and the Junior Certificate School Programme (a programme for at-risk students in the Junior Cycle of post-primary schooling)
- addressing early school leaving



- implementing new syllabi/courses in the areas of mathematics (a new Junior Certificate syllabus was implemented in 2000), and in science (a new Junior Certificate syllabus was launched in 2003)
- increasing uptake of science subjects in post-primary schools (a government Task Force on the Physical Sciences made heavy use of the PISA results in its report published in March 2002 – see <http://www.education.ie/> and go to Reports and Publications).
- Curriculum review – the National Council for Curriculum and Assessment draws on the frameworks underpinning the PISA 2000 assessment in its work on advising the Minister for Education and Science on matters related to curriculum and assessment.

### **Influence on schools**

The impact of PISA on schools has been relatively low to date. This may be due, in part, to the fact that PISA doesn't involve teachers directly in the assessment (by, for example, asking them to complete a Teacher Questionnaire). In order to increase teachers' awareness of the outcomes of PISA 2003, the Educational Research Centre published a report titled, *A Teacher's Guide to the Reading Literacy of Irish 15-Year Olds*, in November 2003. This is currently being disseminated to schools and can also be downloaded at <http://www.erc.ie/pisa>

## **Luxembourg**

### **Reception by policy makers**

The educational authorities, the public, pupils and teachers were shocked by the PISA results. The political opposition mainly used PISA results in order to criticise the governmental policy. Main lines: school failure, complexity of the structure of the school system, content of the school programs

Starting in 2002, several measures were taken in order to improve the pupil's educational situation: luckily, some reform projects had been decided before publication of the PISA results; their necessity has been confirmed by the analyses of the results.

- National project on the organisation of preschool and primary school education
- National project on the reform of apprenticeship and vocational education
- National project on the organisation of classical and technical secondary schools
- National project on the review of the curriculum of the lower technical secondary school
- Measures to improve the recruitment of teachers in post-primary education
- Measures to improve reading comprehension, not only in primary and secondary education but also in the pre-school section
- Measures to improve the competencies of pupils with an immigration background

### **Influence on schools**

There is no empirical evidence about the influence of PISA on schools, on individual teachers or on the teaching methodology. At the very beginning a feeling of resignation could be noticed in some cases. However, the PISA shock wave passed through Luxembourg quite quickly and forces could be concentrated on improving the organisation and perception of the study for the oncoming second cycle.



## Netherlands

### Reception by policy makers

To get mass media attention for important research reports a press-conference is organised (including a press release) where the report is offered to the minister of Education by the research institute and the press can pose questions to the researchers and to the minister. The minister has received a summary of the background of the report and of the results and is prepared for possible question that can be posed to here. The political forces, the members of parliament did what they are used to do: questioning the minister on what they have read in the papers : the absence of the Netherlands in the OECD report. Since the Netherlands scored well in PISA there were no consequences in terms of reform or development.

### Influence on schools

Schools and teachers were informed by the media and by the journal and the website of the ministry. There was no feedback to schools on their own scores, since a sample of 30 pupils of the same age per school is not representative for the separate big Dutch secondary schools. There is no influence on instruction. PISA 2003 got enough responses (over 80% instead of the 50% in 2000) This means that a notion of the importance of PISA is established.

## Norway

*Are there concrete initiatives for reform and development as a consequence of the international studies? Of what kind (e.g. standards, testing, curriculum)?*

The "diagnosis" of international studies, in particular PISA and PIRLS, has been used in support for a number of political initiatives. Among them are:

- The number of lessons in Norwegian has been increased from a total of 3040 to a total of 3154 lessons in grades 1-4 in primary school.
- A comprehensive strategy for better teaching and learning of mathematics and science
- A comprehensive strategy for the stimulation of pupils' enthusiasm for reading.
- Scholarships for teachers for further education in Norwegian (focusing on reading and writing abilities) and mathematics.
- The establishment of national centres of competence for the teaching and learning of basic subjects (science, mathematics, reading).
- National tests of learning outcome in basic subjects (reading, writing, mathematics and English) from 2004/2005. The testing will be part of the new national system of evaluation (operative from 2004) and will be based on the same type of methodology as the PISA and PIRLS studies.
- The establishment of a national system of evaluation has a long history, dating back to 1988. Norway's ranking in international studies over the last years has undoubtedly contributed to an accelerated speed in the development.

### Influence on schools

From the national media debate it seems that some schools, teachers and researchers meant that some others (media, politicians, researchers etc.) painted a too negative picture of the school. So far we do not have information on the effect of international studies on the practise



in schools. With the new national evaluation system there is a possibility to trace influences over time.

## Sweden

### Reception by policy makers

The minister and some of his staff were informed a few days before the official publication of the international and the national report. From the ministry's point of view the results were important and useful. Afterwards they have been used in different analyses and reports from the ministry, not least in the preparation of a national analysis and follow-up of the common education goals within European Union. The public discussion in media on the lack of discipline resulted in a political debate on education in for instance the parliament, before the election to parliament in September 2001.

PISA 2001 has, however, not resulted in any initiatives for national reform etc. However, National Agency for Education has taken an initiative to make an in-depth analysis on the results of immigrant children, to offer better insights than the initial reports could give. The report has been published in Swedish, but soon there is an English version as well, since we have compared Sweden with nine other countries. A Nordic report on the PISA results has also been published – *Northern Lights on PISA. Unity and diversity in the Nordic countries in PISA 2000*.

In fact, we now have a strategy to publish PISA-results not only once but use parts of the results in new publications. We also try to make the analyses initiated by PISA BPC known and discussed, more than earlier international comparative studies, by translating chapters into Swedish or making Swedish summaries.

### Influence on schools

Schools and teachers have been informed about the results primarily via media and via the website of National Agency for Education. We have no indications on if or how PISA has influenced what's going on in schools. We know that the teachers' unions have been interested in the study and that the results from the first as well as later reports have been published in their magazines.

## Spain

### Reception by policy makers

A paper was prepared by the National Institute for Evaluation and Quality of the Education System (*Instituto Nacional de Evaluación y Calidad del Sistema Educativo-INECSE*), official body responsible for the implementation of the PISA program in Spain, for the Office of the Minister of Education, Culture and Sport, and for the Press Office of the Ministry. Also, the Minister of Education, Culture and Sport, together with higher officials of the Ministry, attended two meetings where the PISA 2000 results were presented.

On the other hand, several meetings, attended both by politicians of the governing party and the parties of the parliamentary opposition, jointly with several stakeholders of the education community, were organized by different official and private institutions. Such meetings reflect the amount of interest and political debate aroused by PISA in Spain. The main meetings organized in Spain around PISA results were:



- a) Meeting organized by the governing political party (*Partido Popular*) (with the attendance of the Prime Minister of the Spanish Government and the Minister of Education, Culture and Sports)
- b) Meeting organized by the Government of the Autonomous Region of Catalonia
- c) Meeting organized by the agency of educational evaluation of the Basque Government (ISEI-IVEI)
- d) Meeting organized by the National Association of Private Schools (*ACADE*) (with the attendance of the Minister of Education, Culture and Sports) (In all these meetings the international and Spanish national results were presented by the person responsible for PISA at the OECD)
- e) Meeting organized by the *Santillana* Foundation (with the attendance of the General Secretary of Education of the Ministry of Education)
- f) Meeting organized by the National Institute for Evaluation and Quality of the Education System (*INECSE*) of the Ministry of Education, with the attendance of representatives of the 17 Autonomous Communities (Regions) of Spain.

As expected, different political parties highlighted different sets of results and, subsequently, interpretations, explanations and valuations of PISA results varied in a wide range among the different political views. The interpretation of PISA 2000 results by the governing authorities backup the ongoing educational programs and reforms already put in force by the government previously to the release of PISA 2000 results. It is to note the ongoing governmental program for the improvement of reading engagement among Spanish students.

### **Influence on schools**

The edition in Castilian language of the Executive Summary of the results of PISA 2000 survey and the two books where the PISA 2000 assessment framework is presented, translated to Castilian, were distributed to all the Spanish participating schools in PISA (both in pilot and main studies). Also, schools, teachers, families and students were informed by the media of the overall results and conclusions on the Spanish students results.

## **ANNEX 2. Description of major international assessments (adapted from ‘*Assessment of Student Learning in International Development Association Countries*’ by Robert Prouty and Ndeye Fama Diagne)**



## **International Assessments**

International assessments compare the learning achievements of students in a country relative to the learning achievements of students in other countries. The international assessments we look at in this report are listed below. It is important to note that MLA, PASEC AND SACMEC are national assessments however; they are used as international assessments as well

### **Monitoring Learning Achievement (MLA)**

This assessment was carried out in 38 out of 81 IDA countries, far the greatest coverage for low income countries. However, the program was closed in 2004 following a general review that reveals a range of technical shortcoming. The objective is to enable countries to identify factors that promote or hinder learning in primary school, analyze problem areas and develop policy changes and new practices to improve the quality of education.

The first MLA assessment of literacy and numeracy achievement of 4<sup>th</sup> and 5<sup>th</sup> grade students and mathematics and science achievement of 8<sup>th</sup> grade students began in 1992. In addition to testing traditional academic subjects, MLA looks at life skills such as knowledge of health, hygiene and nutrition. MLA studies administer questionnaires to pupils, class teachers, head teachers and parents. Gender parity in learning outcomes at the lower grades of basic education is an achievable goal. Most surveys indicate also that in some cases it is the same at secondary level mainly in literacy but not always in mathematics and sciences.

### **Program for the Analysis of Education Systems of CONFEMEM Countries (Francophone West Africa) (PASEC)**

PASEC was established in 1991 by the ministers of education of French-speaking countries in sub-Saharan Africa. It has attempted to use assessment as a way of identifying efficient models in schools, and its major goal is to build capacity in each nation to carry out on-going assessments. Studies have taken place in twelve IDA countries. Each country has implemented at least 2 major assessment projects focusing on achievement in mathematics and French in Grades 2, 3 and 5. Tests were administered at the beginning and the end of the year beginning in 1991 (Grades 2 & 5) and 1993 (Grade 3) and ending in 2000.

PASEC explored the impact on achievement of 16 in-school factors (such as teacher training, class size and availability of text books) and 8 environmental influences (such as parental education, distance to school and language spoken at home).

### **Latin American Program for the Assessment of Quality in Education (Laboratorio)**

The Laboratorio project is sponsored by UNESCO. It is a network of assessment systems in 18 Latin American countries. The general aim is to improve policy development at both macro and micro levels. In 1997, two IDA countries, Bolivia and Honduras participated. Tests of language and mathematics were given to representative samples of students in Grades 3 and 4 in each country.



## **Progress in International Reading Literacy Study (PIRLS)**

PIRLS was directed by the International Study Center at Boston College (ISC) and to date, four IDA countries, Georgia, Indonesia, Moldova and Nicaragua are planning to participate in PIRLS 2006. Each of them participating in 2006 PIRLS will be appointed a National Research Coordinator (NRC) who will carry out the study. The NRC is responsible for selecting the school sample, collecting data, conducting scoring and data entry, and preparing a national report of study results.

PIRLS consists of a carefully-constructed test assessing a range of reading comprehension strategies for two major reading purposes - literary and informational. PIRLS collect extensive information about home, school, and national influences on how well students learn to read. Of special interest, parents and caregivers complete questionnaires about their children's early literacy activities.

## **Programme for International Student Assessment (PISA)**

This survey of 15 year olds was launched in 1998 by the 29 member countries of OECD and Brazil, China, Latvia and the Russian Federation. The goal of PISA is to assess whether education systems are providing the tools for continued learning that today's young people will need over the course of a lifetime for full participation in society.

The first PISA assessment took place in 2000 with 27 countries and Brazil, Latvia, Liechtenstein and the Russian Federation participating. Thereafter, assessments were to occur every 3 years. Reading literacy, mathematical literacy and scientific literacy were assessed, with two-thirds of testing time in each cycle devoted to a 'major' domain and assessed in depth. The testing schedule for the major domains is reading literacy in 2000, mathematical literacy in 2003 and scientific literacy in 2006. PISA 2000 shows a pattern of gender differences that is fairly consistent across countries

Albania and Indonesia participated in the PISA in 2000, Moldova in 2001, Serbia Montenegro in 2003, Indonesia and Kyrgyz Rep will participate in the 2006 PISA.

## **Southern Africa Consortium for Monitoring Educational Quality (SACMEQ)**

SACMEQ was undertaken by the Ministries of Education in co-operation with IEP of UNESCO to undertake educational policy research with the main aim of generating reliable information that can be used by decision-makers to plan the quality of education. The main goal was to enhance the research and evaluation capacity of each of the national education systems. The first round of studies occurred in 1994 and involved four IDA countries (Kenya, Malawi, Zambia and Zimbabwe), and the second round of studies in 2000 involved six IDA countries (Kenya, Lesotho, Malawi, Mozambique, Tanzania and Uganda). The reading levels of 6<sup>th</sup> Grade students were tested. The 59 reading test items were split into three subsets (narrative, expository, and documents). Research studies in many countries have indicated that there is a relationship between availability of books in the home and pupil reading achievement. The pupils in the survey were asked to indicate the approximate number of books in their homes



## **Third International Mathematics and Science Survey (TIMSS)**

TIMSS, the most extensive international comparative study to date, began in 1994 - 1995 and was coordinated by the IEA and the international cooperative of research centers and departments of education in more than 50 countries with headquarters in the Netherlands. The goal of TIMSS is to measure, compare and explain learning in mathematics and science at five grade levels (the third, fourth, seventh, and eighth grades, and the final year of secondary school). Tests were designed by task forces from participating countries, but each nation cooperated by being responsible for its own data collection and analysis. In addition to asking questions in the two subjects at the three grade levels, researchers asked students, teachers and school administrators to fill out questionnaires aimed at producing better understanding of how contextual factors affect achievement (i.e. students were asked about their attitudes towards their schooling, and teachers about how they structured their classrooms).

Albania, Ghana, Moldova, Serbia Montenegro and Yemen participated in the 1999 – 2003 TIMSS. Nine IDA countries are planning to participate in the 2007 TIMSS: Armenia, Bosnia-Herzegovina, Djibouti, Georgia, Ghana, Mongolia, Uzbekistan and Yemen.

## **Pacific Islands literacy level (PILL)**

As part of the UNDP-funded Basic Education and Literacy Support Programme (BELS), a Pacific Islands literacy level test was developed and administered to sample groups of primary children in most of the South Pacific Island countries in 1993. The areas tested were English and Vernacular languages. The result of the survey showed that a large number of children were underachieving in general literacy. Four IDA Countries, Kiribati, Solomon Islands, Tonga and Vanuatu took part of the test. The PILL 1 test was administered to 1327 grade 4 (9 -11 years old) students in 1994 and PILL 2 to 846 grade 6 (13 -15 years old) students in 1995.

