

Education Policy Analysis Archives

Volume 8 Number 32

July 11, 2000

ISSN 1068-2341

A peer-reviewed scholarly electronic journal
Editor: Gene V Glass, College of Education
Arizona State University

Copyright 2000, the **EDUCATION POLICY ANALYSIS ARCHIVES**.
Permission is hereby granted to copy any article
if **EPAA** is credited and copies are not sold.

Articles appearing in **EPAA** are abstracted in the *Current Index to Journals in Education* by the [ERIC Clearinghouse on Assessment and Evaluation](#) and are permanently archived in *Resources in Education*.

Student Assessment as a Political Construction: The Case of Uruguay

Luis Benveniste
The World Bank

Abstract

This article reveals the interplay between assessment policies in Uruguay and the nature of State-societal relations. The central State has been historically a staunch defender of public education and has championed the cause of equalizing opportunities for the most disadvantaged sectors of society. The national evaluation system of student performance has been constructed as an expression of this tradition. The Uruguayan government sought to build a wide level of consensus with respect to the assessment instruments by encouraging educators to participate and buy into the assessment initiative. Moreover, the national government shifted the focus of the national evaluation from measuring schooling outcomes to addressing the social wants that condition student learning. Hence, the national evaluation has come to symbolize an agreed-upon mechanism of social accountability by which the central government upholds its responsibility for educational provision as it intervenes on behalf of impoverished communities. (Note 1)

This article is also available in Spanish in Adobe Acrobat format at <http://www.grade.org.pe/gtee-preal/docs/Benveniste.pdf>

This study reveals the interplay between assessment policies in Uruguay and the nature of State-societal relations. The central State has been historically a staunch defender of public education and has championed the cause of equalizing opportunities for the most disadvantaged sectors of society. The national evaluation system of student performance has been constructed as an expression of this tradition.

The first section describes the educational system as a highly centralized organizational structure. Then, it provides a brief overview of the education reform initiative launched in 1995 by the National Administration of Public Education to promote and consolidate social equity.

The second section portrays the *Unidad de Medición de Resultados Educativos* (the evaluation agency of primary education) as a temporary unit created in 1996 within the framework of a project financed by the World Bank. In spite of its short history, the assessment system has garnered substantial popular support and spurred a curricular and pedagogical renovation among teachers, principals and supervisors.

The third section explores the reasons behind the public embrace of the national assessment system. This has been no slight accomplishment in light of the fact that the evaluation of student performance may potentially exert a destabilizing role by highlighting deficiencies in educational service provision. First, the central State circumscribed teacher liability over poor performance, largely assuming itself the responsibility for the character of schooling. Second, the national government built a wide level of consensus with respect to the assessment instruments by encouraging educators to participate and buy into the assessment initiative. Third, the national government shifted the focus of the national evaluation from measuring schooling outcomes to addressing the social wants that condition student learning. Hence, the national evaluation has come to symbolize an agreed-upon mechanism of social accountability by which the central government upholds its responsibility for educational provision as it intervenes on behalf of impoverished communities.

Assessment may in fact reify centralized control by imposing standards that must be uniformly enforced throughout the country. Paradoxically, in Uruguay's highly concentrated model of governance, the national evaluation proves that centralization need not be incompatible with democratic participation.

The process of education reform in Uruguay

The Uruguayan educational system

The educational system of the Republic of Uruguay is organized in three levels. (Note 2) Initial education caters to children between 3 and 5 years of age. Preschool instruction is not compulsory presently, but the government plans to make it obligatory for 4 and 5 year-old children in the proximate future. Primary education consists of six grades and services 6 to 11 year-old children. Secondary education consists of two sub-cycles. The *Ciclo Básico Único* (Unique Basic Cycle) is a three-year course common to all students between 12 and 14 years of age. Students may then opt to proceed for baccalaureate or technical-professional instruction to round off their secondary education. Training at this level may last between 2 and 7 years depending on the course. Primary schooling and the Unique Basic Cycle constitute the national compulsory educational requirements (Uruguay—Ministerio de Educación y Cultura,

1996).

The administration of the education sector is highly centralized, but falls under the jurisdiction of several independent de-concentrated councils. The Ministry of Education and Culture is responsible for devising broad national educational policies. Despite its overarching mandate, this Ministry has a subsidiary role in the operations of the education sector. The *Administración Nacional de Educación Pública* (ANEP), the National Administration of Public Education, is the agency responsible for the management of the public educational system. The ANEP is fully autonomous from the Ministry of Education and Culture and it is configured by several bodies: (a) the Central Board Council (CODICEN), (b) the Council of Primary Education (CEP), (c) the Council of Secondary Education (CES), and (d) the Council of Technical-Professional Education. The Central Board Council is the highest administrative authority in the education sector. It is comprised of 5 members elected by the President and approved by the Senate. The other three councils are subordinate to the CODICEN, but they function largely autonomously. They are responsible for imparting, administering and supervising educational services. The directors of these councils are appointed by the CODICEN (see Figure 1).

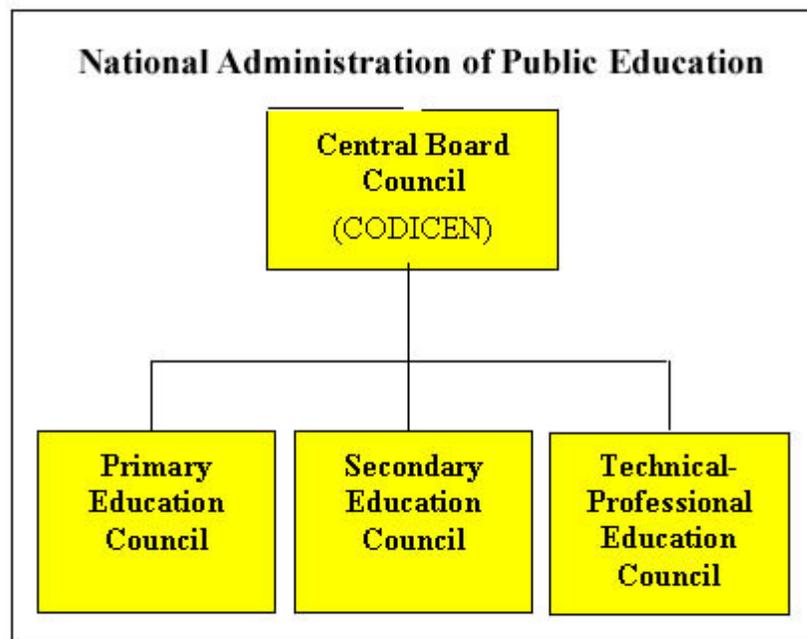


Figure 1. Organizational Structure of the ANEP

Educational policy is also shaped by several independent official advisory bodies to the ANEP. The Coordinating Commission of Education consists of the Minister of Education and Culture, the highest authorities of the autonomous councils as well as by representatives of universities and post-graduate institutions. It propounds guidelines and draft agreements for the coordination of the education sector. The *Asambleas Técnico-Docentes* (Technical-Pedagogical Assemblies or ATDs) are national and regional deliberative bodies comprised of teachers elected through secret compulsory voting. ATDs pronounce opinions regarding the conditions of education and may initiate educational policy directives (González Rissotto, 1997).

Basic education has reached universal proportions in Uruguay. In 1995, net enrollment rates at the primary school level encompassed 95% of the 6 to 11 year-old cohort. At the Unique Basic Cycle level, matriculation rates averaged 67% for the

relevant school-aged population in Montevideo and 57% for all other urban areas in the rest of the country. Participation rates drop sharply in the second cycle of high-school instruction. Net enrollments at this level were below 30%. Total expenditures in education amounted to US\$ 578 million in 1995, which represents 3.4% of the gross national product. The private sector caters to 13% of primary school students and 14% of secondary school enrollments (Uruguay—Ministerio de Educación y Cultura, 1997).

Uruguay has a shortage of teachers. The imbalance between teacher supply and demand has prompted governmental authorities to allow instructors to work double shifts. Teachers' real income has deteriorated steadily, even declining during periods of private real income recovery. Between 1960 and 1989, real salaries for teachers declined by 46.6%. Monthly wages in 1996 ranged between US\$ 270 and US\$ 407 (Uruguay—Ministerio de Educación y Cultura, 1996). Low salaries have forced teachers to search for alternative sources of income.

The Uruguayan education reform

A concern for the inequities in the Uruguayan educational system has prompted the government to embark on an ambitious reform initiative. Net enrollment rates for the population in chronic poverty reach 27% for preschoolers and 34% for high school students. The dropout rates for the poorest children in the first cycle of obligatory secondary education surpass 37%. There is also growing weariness about the deterioration of the quality of education. The national assessment of student achievement revealed that 6th graders in extreme poverty responded correctly to 37% and 17% of a language and mathematics test on average. The national means are nearly 20 percentage points above these levels. Primary school repetition rates have remained stable at around 10% during the past fifteen years. The repetition rate in the first grade, however, has reached 22%. In Montevideo, 63 out of 257 schools have a repetition rate in the first grade above 30%, and another 67 establishments between 20% and 29% (Rama, 1998).

The current administration of the ANEP has adopted four guiding principles to transform the educational system (Rama, 1998; Uruguay—Ministerio de Educación y Cultura, 1996):

1. The consolidation of social equity,
2. The appreciation of teacher professionalism and training,
3. The improvement of educational quality, and
4. The strengthening of institutional management

The *consolidation of social equity* effort directs services and compensatory actions to underprivileged children. The ANEP seeks to extend public preschool services to 95% of the 5 year-old population and conduct an outreach program to incorporate 85% of 12 to 14 year-olds to the first cycle of secondary schooling. The poorest students receive more hours of instruction, including “full-time” schooling. They also have access to a comprehensive school meal program.

The *appreciation of teacher professionalism* effort strives to double the graduation rates of primary school teachers and triple that of secondary school instructors by the end of 1999. Approximately 90% of the elementary school teacher corps and 4,300 non-certified high school instructors will receive in-service professional development training. Teacher salaries were planned to undergo an increase of 13% in 1996, 10% in 1997, 15% in 1998 and 18% in 1999. In actuality, teacher salaries did rise over the yearly inflation rates, but did not reach the goals originally contemplated. Nonetheless, education was the only social sector that received an appropriation to increase salaries

and its general operating budget in August 1998.

The *educational quality enhancement* effort focused around the widespread distribution of textbooks, instructional materials and pedagogical resources to public establishments. Curricular programs at the secondary level are also undergoing an in-depth review and renovation. In addition, the ANEP finances school-based projects to address specific needs within educational communities. Finally, the government has launched a program, “All Children Can Learn,” to reduce primary school repetition rates. This program consists of a series of integrated social activities that endeavor to facilitate the access and permanence of children in schools, to strengthen the coordination between preschool and primary education, to enhance teacher training and to use textbooks as “an instrument for open learning” (Rama, 1998).

The *strengthening of institutional management* effort encompasses specialized training for school principals as well as the creation of computerized systems to assist administrators in their functions. Rural schools with less than ten students are being consolidated in order to reduce wastage and promote a more efficient use of resources.

These four initiatives are funded by a 22% increase in the education sector appropriation. The 1996-2000 budget has grown by US\$ 75 million from the 1991-1995 budget, to US\$ 430 million. The government of Uruguay also receives substantial aid from the international donor community to implement these reforms. The Inter American Development Bank and the World Bank have lent \$140 million dollars to the modernization of the educational system. The Project for the Improvement of the Quality of Primary Education (MECAEP), (Note 3) funded by the World Bank, has contributed to the construction of preschools, the in-service training of elementary school teachers, and the provision of textbooks and pedagogical resources. It also supports the *Unidad de Medición de Resultados Educativos* (UMRE), the agency responsible for assessing educational quality at the primary level. The Project for the Improvement of the Quality of Basic Education and the Instruction and Training of Teachers (MESyFOD), funded by the Inter American Development Bank, has supported the creation of five regional teacher training centers, the in-service development of high school instructors, and the maintenance of secondary school infrastructure. In addition, MESyFOD has conducted the national assessment of student achievement at the secondary level in 1999.

2. Student assessment practices in Uruguay (Note 4)

A. The measurement of student achievement

Initial experiences with student assessment

Between 1990 and 1994, the United Nation's Economic Commission for Latin America and the Caribbean (CEPAL) conducted a series of studies requisitioned by the National Administration of Public Education. These studies were based on two examinations administered in 1990 to a small sample of 4th and 9th grade students in language and mathematics. CEPAL also collected socioeconomic and background information from parents, teachers and principals. The purpose of these tests was to explore the conditions of basic and secondary education in Uruguay (Comisión Económica para América Latina y el Caribe, 1994; 1993; 1992; 1991; 1990).

The primary school evaluation revealed that on average students could respond correctly to 58% of the questions (Comisión Económica para América Latina y el Caribe, 1991). The results from the secondary school evaluation were significantly inferior. Less than 22% of public school students reached an adequate level of

proficiency in mathematics or language, as opposed to over 50% in the private sector. The mathematics test showed that “students learn very little in the courses of the Unique Basic Cycle.” The language scores exposed that “the probability of success of the great majority of public establishments is so low that failure is almost certain” (Comisión Económica para América Latina y el Caribe, 1992: 90, 122).

The reports produced by CEPAL, however, abstained from making curt accounts or generic descriptions of student outcomes. Rather, test scores were the starting point for in-depth analyses of the impact of socioeconomic variables on student learning. Predictably, CEPAL found that low-income children tend to have lower levels of academic attainment. After an exhaustive review of the effect of various sociocultural indicators on school performance, the CEPAL underscored that maternal educational level is the best predictor of student achievement (Ravela, 1997b).

The research agenda of this study also included the identification of schools that, despite serving disadvantaged populations, have attained high levels of academic performance. These educational establishments were denominated “exemplary schools.” The CEPAL carried out a qualitative investigation of these schools and posited that there were four factors that explain scholastic excellence in underprivileged environments:

1. the ability of the principal to assume a leadership role in the school as well as in its community,
2. the knowledge and experience of the classroom teacher combined with the satisfaction and commitment to his/her work,
3. a dynamic pedagogical culture within the teacher cadre, and
4. the existence of significant bonds between the educational establishment and parents (Ravela, 1997a).

Finally, the CEPAL emphasized that low test scores were symptomatic of a systemic crisis in the education sector.

The reason for the results is not the fault of educational establishments or their authorities.... They are the outcome of a prolonged social process, during a prolonged historical period, during which the quality of education ceased to be a priority as an objective of State action (Comisión Económica para América Latina y el Caribe, 1992: 123).

In other words, the deterioration of educational quality was ascribed to a lack of commitment from the central State to make adequate investments in schooling services. According to this report, the reversal of this situation would follow from the initiative of the national government towards promoting policies and programs that support the labor of teachers and principals.

The construction of a national assessment system

It could be said that Uruguay does not have an institutionalized national assessment system. UMRE, the unit responsible for the measurement of academic achievement at the primary education level, is not a formal “line- agency” of the National Administration of Public Education. It is an *ad hoc* unit initially constituted to implement the evaluation sub-component of the MECAEP Project financed by the World Bank. UMRE must abide by the directives of the Central Board Council, but it is exonerated from following certain civil service regulations. Similarly, the secondary education evaluation was developed autonomously within the framework of the MESyFOD Project, funded by the Inter American Development Bank. Although there are plans to make student assessment a permanent entity within the governmental

organizational structure, the appraisal of academic performance currently operates from quasi-independent transitory agencies. This situation has provided to the evaluation of student achievement certain degree of independence and freedom—in relation to its organization, operation and personnel selection—by means of its ability to proceed outside the strict channels that regulate public offices. On the other hand, and as it will be described in a later section, this “extra-official” character has generated concern among certain sectors of the educational community, and particularly among the school inspectorate, who perceive UMRE as a parallel entity, alien to them.

The systematic and periodic measurement of schooling outcomes was not an initiative of the Uruguayan government. It was a conditional clause for the appropriation of the MECAEP World Bank loan (Interview UGN1). Although initially greeted with some resistance, the Uruguayan government eventually welcomed the creation of an evaluation unit (Interview UGN34). Germán Rama, who became Director of the ANEP in 1995, had been responsible for the design and implementation of the CEPAL study on student achievement aforementioned. Under his leadership, the Central Board Council decreed a resolution in March 1996 stipulating that “one of the prioritized lines of action of this Council is the implementation of assessment systems of [student] learning ... with the objective to appraise the performance of this Organism and the quality of service it provides to the population” (Uruguay—Administración Nacional de Educación Pública, 1996b).

UMRE has been in operation since 1994. Pilot tests for a 3rd and 6th grade evaluation were conducted late that year, with the intention to launch the first national assessment in 1995. When Dr. Rama assumed control of the ANEP in mid 1995, however, he replaced the technical leadership of UMRE and resolved to postpone the exam for one year. The national assessment underwent an important reformation. First, the ANEP would evaluate all public and private school students in 6th and 9th grades, the terminal years of the primary and secondary educational levels, every three years. Second, the test would veer from appraising curricular contents to measuring skills and competencies (such as reading comprehension or problem resolution). Third, the evaluation would incorporate a detailed sociocultural survey to be completed by parents, teachers and principals. Fourth, UMRE would seek feedback about its mission and operations from the various stakeholders involved in the provision of schooling services. Fifth, governmental authorities committed to maintaining secrecy about individual school test results. The ANEP guaranteed that only aggregate data would be made public (UMRE, 1996e).

UMRE is constituted by 3 full-time and 5 part-time professionals. It is responsible for the design, implementation, analysis, and devolution of results of the primary education assessment. From practically its inception, public and private school authorities as well as policy makers, supervisors and teachers were consulted about the development of instruments, test administration practices, and the uses of assessment results. The government also held regular informative workshops and produced several publications to raise awareness about the objectives of collecting student data (UMRE, 1996e). UMRE devoted significant effort to securing support and building consensus for the national assessment across the gamut of educational actors. In 1996, an “Advisory Group” was consolidated to review the work of UMRE and promote cooperative participation. This committee is conformed by national and regional representatives from the Council of Primary Education, the supervisory cadre, teacher training institutions, the Technical-Pedagogical Assembly, the Association of Private Education Establishments, the Uruguayan Association of Catholic Education, and the Uruguayan Federation of Teachers (the national teachers' union).

UMRE administered the first standardized evaluation in mathematics and language to all 6th grade students in 1996. Rural schools with less than six pupils in the sixth grade classroom were exempted from participation. Absenteeism rates were below 3.5% of the total enrollment in the mathematics test and 6.2% in the language test. In addition, educators and parents were required to complete socioeconomic background surveys. The rate of parental response to this survey was 98.5% (UMRE, 1998c).

The exams consisted of multiple choice items and open-ended questions. Teachers and supervisors participated in the formulation of test items, but technical staff from UMRE ultimately devised the exam. (Note 5) Independent proctors monitored the administration of the assessment. Students were allotted one hour and thirty minutes to complete the test, but those who required additional time to finish were allowed to do so (UMRE, 1996f). UMRE was responsible for correcting the exams and analyzing the results.

Forty days after the application of the test and prior to the culmination of the academic year, schools received an individualized confidential report with aggregate school results item by item. The socioeconomic background surveys served as a basis to categorize schools into five categories, from very unfavorable to very favorable contexts. Student outcomes were compared to the national average, the departmental/regional average and that of schools that service students from similar socioeconomic conditions. Educational establishments also obtained two technical manuals to interpret results. In the following academic year, educators received a second confidential report with a sociocultural profile of their school, based on background questionnaire data. UMRE also produced methodological guides with pedagogical suggestions and recommendations to redress weaknesses identified in mathematics and language (UMRE, 1997b; 1997c; 1997d; 1997f; 1997g; 1996c; 1996d; 1996g).

UMRE tailored several reports for the supervisory cadre. School inspectors participated in workshops where they received a regional profile of local schools and a “socioacademic map” that classified educational establishments under their oversight in terms of achievement levels and socioeconomic context. These instruments would allow supervisors to identify exemplary schools that exhibited high test scores in spite of being resource poor. They were also meant for targeting compensatory interventions to low performing educational establishments.

UMRE results

The national assessment of 6th grade students showed that 57.1% of students were able to respond to more than 60% of the language test correctly. The success rate in mathematics was considerably lower. Only 34.6% of students were able to answer over 60% of the questions satisfactorily. The percentage of students that did not reach the 60% “adequacy level” in both tests was 37.9%.

The first official report of results for public dissemination highlighted the role of contextual variables in the acquisition of knowledge. Students were classified into four categories according to their sociocultural context. Sociocultural context was defined in terms of maternal educational level. Schools from “very favorable” contexts were characterized as those with over 50% of students whose mothers completed at least secondary education. Schools from “very unfavorable” contexts were characterized as those where less than one out of two mothers had received only a primary education, and at most one out of ten mothers had received a secondary education.

As the CEPAL studies had demonstrated earlier, students from underprivileged backgrounds scored significantly below students from more affluent families (see Tables 1 and 2). While over 85% of children from “very favorable” contexts answered correctly

to at least 60% of the language test correctly, less than 40% of students from “very unfavorable” contexts attained the same level of achievement. In mathematics, the gap between high- and low-income children widened.

Table 1
Percentage of Students by Performance Level in
Mathematics and School Sociocultural Context

	Very Favorable	Medium High	Medium Low	Very Unfavorable	Total
Highly Satisfactory (scores above 80%)	21.0%	8.4%	3.4%	2.0%	6.8%
Satisfactory (scores 60% to 80%)	45.6%	35.3%	23.2%	15.7%	27.8%
Unsatisfactory (scores 30% to 60%)	30.6%	49.7%	60.7%	64.4%	54.5%
Very unsatisfactory (scores below 30%)	2.8%	6.7%	12.7%	17.9%	10.9%
Total	100.0%	100.0%	100.0%	100.0%	100.0%

Source: UMRE (1996g), p. 10.

Table 2
Percentage of Students by Performance Level in
Language and School Sociocultural Context

	Very Favorable	Medium High	Medium Low	Very Unfavorable	Total
Highly Satisfactory (scores above 80%)	41.9%	19.5%	9.8%	5.0%	15.8%
Satisfactory (scores 60% to 80%)	43.3%	48.1%	40.9%	32.8%	41.3%
Unsatisfactory (scores 30% to 60%)	14.0%	29.7%	43.2%	52.7%	37.7%
Very unsatisfactory (scores below 30%)	0.8%	2.8%	6.1%	9.5%	5.2%
Total	100.0%	100.0%	100.0%	100.0%	100.0%

Source: UMRE (1996g), p. 10.

UMRE produced a second report exploring the relationship between sociocultural factors and student achievement. This study categorized the Uruguayan educational system into five subsystems according to geographical and sociocultural variables. This study revealed that private schools in Montevideo generally attracted students with the highest maternal educational levels, followed by, in decreasing order of maternal educational level, private schools in the interior, public schools in Montevideo, public schools in the interior, and rural schools. School performance in these subsystems was closely correlated to sociocultural context, with the exception of rural schools that evinced academic achievement levels slightly greater than expected for their low sociocultural context (UMRE, 1997f). More importantly, this report provided proof that academic achievement levels were not directly tied to the public or private nature of schooling, but rather to the sociocultural composition of the student body. In other words, the average scores of public schools from very favorable contexts were similar to those of their private counterparts within this context. The outcomes of private schools that served underprivileged populations were also analogous to those of public schools that assisted students from very unfavorable contexts. (Note 6)

A third national report was released late in 1997 providing a meticulous institutional profile of educational establishments. This document was based on the background surveys provided by principals, teachers and parents. It depicted the attributes of building facilities, school materials, class size, years of experience of principals, teacher training, pedagogical approaches favored, staff turnover, parental involvement, and student self-esteem (UMRE, 1997g). As in previous inquiries, the analysis gravitated around the relationship between sociocultural context and schooling conditions.

Overall, the Uruguay government emphasized consistently throughout its public reports the role played by contextual factors in student learning. Average student scores, as all comparisons between geographic regions or between the public and the private sectors, were presented in direct relation to the sociocultural level in which learning took place. School-level data was kept rigorously confidential.

Other assessment activities

In addition to the sixth grade assessment, the Uruguayan government has undertaken two other evaluation exercises. Firstly, the government conducted an experimental assessment to a stratified sample of 3rd grade classrooms late in 1998. This test was available to other educational establishments outside the controlled sample for self-administration on a voluntary basis. The Central Board Council, however, exhorted all educational establishments to take part of this initiative (UMRE, 1998a). The purpose of this evaluation was to appraise student competencies at mid-point of their primary schooling. It also pursued to signal teachers about the expected competencies pupils ought to master by the third grade and provide them with an early-warning system to reformulate programmatic contents and pedagogical strategies (UMRE, 1997a).

The exam consisted of open-ended questions that integrated concepts from a variety of disciplines (mathematics, language, social studies, natural sciences, moral education, art) without compartmentalizing them into different spheres of knowledge. In response to teachers' demands for greater participation in the formulation of the test, UMRE established working groups with educators selected by the supervisory cadre, the regional Technical-Pedagogical Assemblies, and the associations of private independent and private Catholic schools. These working groups identified curricular areas to be

evaluated and collaborated in the development of test items.

An informational document providing detailed information about the proposed testing scheme and objectives was drafted and distributed to all teachers and school inspectors. UMRE later requested teachers to respond to an opinion survey regarding the assessment instrument and competencies to be evaluated. Ninety two percent of respondents declared that the test was “adequate” and there was complete agreement about the competencies selected (UMRE, 1998b). As in the 6th grade assessment, the measurement instrument included background surveys for parents, teachers and principals in order to obtain data regarding the conditions in which student learning took place.

Every educational establishment received a report with national aggregate averages by competencies (reading comprehension, resolution of problems, processing information). Test scores were also broken down by socioeconomic context (rural, very favorable, favorable, medium, unfavorable and very unfavorable). A supplementary report detailed average background information (maternal educational level, home overcrowding, books in the house, preschool training) tabulated by sociocultural context. Schools that did not participate in the controlled sample received as well a standardized correction manual so that they could tally their own in-house results and compare them to the official national average scores.

Secondly, the 6th grade cohort evaluated in 1996 was re-tested in 1999 as students completed their 9th grade. MESyFOD, the project responsible for the administration of the test, espoused a methodology similar to that implemented by UMRE. The evaluation team sought to conduct informational sessions with supervisors, private and public school instructors, the Technical-Pedagogical Assemblies (ATD) and the teachers' union to gain their support. MESyFOD also intended to establish an advisory group conformed by representatives from every sector of the educational system that would review its operations. At the time the data collection for this study was conducted, it was unclear whether MESyFOD would be able to build consensus for the evaluation, especially from the ATDs and the *Federación Nacional de Profesores de Enseñanza Secundaria*, the national secondary school teachers' union. Secondary school teachers had adopted a more contentious stance towards the central government's reform initiatives than primary school educators. ATD representatives had refused in the past to collaborate in projects spearheaded by MESyFOD (Interviews UGN3, UGN3b). (Note 7)

The MESyFOD team, however, concedes that the national experience with UMRE had greatly eased their work nonetheless. In most instances, educational establishments offered little resistance. They had not questioned the government's rationale for conducting this initiative nor were they concerned about being penalized for poor performance.

Our undertaking has been facilitated due to the fact that MECAEP has been very careful about the confidentiality of test results, about the prompt devolution of scores, about the provision of individualized reports to each educational center. They took a series of precautions that, for instance, have encouraged private schools to open their doors. ... The realities of secondary education are not the same as those of the primary level, and there's still all the prejudices about standardized evaluations, but we're going along (Interview UGN3).

The assessment involved approximately 40,000 students. It appraised achievement in

language, mathematics, social studies and natural sciences. Tests were administered by independent proctors and corrected centrally by MESyFOD.

B. The uses of assessment data

The findings uncovered by the first national measurement of student achievement are aimed at three distinct audiences: (a) the central government, (b) the school inspectorate, and (c) teachers and principals. Parents are informed indirectly about the general conditions of schooling through the press. A few schools, mostly in the private sector, have taken the initiative to publicize their scores to the families they serve.

The central government

The national evaluation of student learning has as its official mandate:
to produce information about the extent to which primary school graduates have been able to develop the skills and fundamental understandings in Language and Mathematics that every Uruguayan child ought to have incorporated regardless of his social origin, economic condition, or local context (UMRE, 1996b: 1).

This mission statement underscores the diagnostic objectives of assessment. “To have this information available,” claims the ANEP, “is crucial to recuperate the democratizing role of the national educational system.” Equity considerations lie at the heart of the central government's involvement in the measurement of academic outcomes.

The ANEP has relied on data gathered by UMRE primarily to guide and inform compensatory policies. There are three autonomous agencies within the national government that are consumers of information generated by UMRE: (a) the Council of Primary Education (CEP), (b) the MECAEP project (which is administered independently from the CEP), and (c) the Planning Area of the ANEP, a unit that depends directly from the Central Board Council.

The MECAEP project has been the most active patron of assessment data. On one hand, MECAEP has played a key role in promoting reflection among educators regarding the results of the first national evaluation. Technical discussions about the meanings of UMRE's findings have become a standard feature of institutional planning or professional development workshops organized for school inspectors and principals (Uruguay—ANEP-MECAEP, 1997). On the other hand, test scores and UMRE's classification of schools according to sociocultural context guide many of the initiatives undertaken by MECAEP. For instance, MECAEP disburses US\$ 3,000 government grants for school-based projects. The selection process takes into account how these projects may address shortcomings identified by the UMRE evaluation. Moreover, priority is awarded to schools from “unfavorable” sociocultural environments (Uruguay—ANEP-CODICEN, 1998). Sociocultural context, as defined by UMRE, has also become a salient criterion for the allocation of resources. The official press release detailing the outcomes of the first evaluation to the general public, for example, announced that MECAEP earmarked US\$ 1 million to the purchase of pedagogical materials, targeting specifically 400 schools from unfavorable contexts (Uruguay—Administración Nacional de Educación Pública, 1996a).

“The Council [of Primary Education] permanently solicits information from UMRE,” states a senior government official. “We are interested in learning about the

strengths and weaknesses in language and math achievement, as well as about the relationship between school and family variables” (Interview UGN6). In practice, although the CEP's school inspectorate has been an important end user of test data, the central CEP office has given at best limited application to the UMRE results. School test scores have been used as educational quality indicators for the program “All Children Can Learn.” This initiative strives to reduce repetition rates below the 20% mark in 160 schools through a comprehensive set of activities that include teacher training, providing health care services, reaching out to parents, and supplying textbooks (Uruguay—ANEP-CODICEN, 1998). Achievement levels have not been a parameter for bringing schools into the program, but test outcomes are occasionally used to tailor specific remedial actions in some establishments. Outside this initiative, the Council of Primary Education does not rely on UMRE data for other purposes. This has been a source of disappointment for some UMRE officials (Interviews UGN1, UGN2).

Finally, the Planning Area of the ANEP has depended on UMRE's school socioeconomic data for several of its own activities as well. In 1998, it conducted a research project on variables associated with primary education repetition rates (Área de Planeamiento de ANEP, 1998). This study demonstrated a close relationship between sociocultural context and the likelihood that students will be held back in the first and second grades. In addition, school background information has been “a fundamental referent” in the identification of establishments that could benefit from recent government initiatives, such as in-school meals, school infrastructure maintenance, or classroom construction (Interview UGN7). It is expected that once the MECAEP project comes to its conclusion, UMRE will become part of the Planning Area of the ANEP. (Note 8)

UMRE's own policy initiatives

The Council of Primary Education maintains that UMRE's role “is bounded to describing what happens” and “providing statistical data,” so that, in turn, this knowledge can serve “the relevant organisms to make pertinent decisions” (Interview UGN6). In practice, UMRE has been more than just an information-gathering agency. It has been intimately involved in the design and promotion of educational policies for schools from “very unfavorable” contexts.

UMRE, with support from regional Institutes for Teacher Training, developed a Saturday workshop series for 541 urban primary schools serving underprivileged communities (approximately 40% of all public establishments). Participation in this four-month seminar was voluntary, but in order to qualify, at least half of a school's professional staff must have agreed to participate. Teachers were remunerated for the time they dedicated to this venture with a monthly monetary bonus equivalent to 30% of the average teacher salary.

Furthermore, UMRE established a fund to finance propositions that could enhance educational quality. Teacher training institutes received \$1,000 awards to foster “the accumulation of knowledge about [student] learning in unfavorable environments and the implementation of professional development activities in teacher training institutions around these themes” (UMRE, 1997e: 1). Low-income schools could solicit \$1,000 grants for the implementation of intervention projects destined to improve achievement levels in that educational community. The resources made available, however, would only allow for 50 school awards altogether.

Lastly, UMRE, in collaboration with the Program for the Strengthening of the Social Area (FAS) from the Office of Planning and Budget, conducted a qualitative research project in 12 schools from unfavorable sociocultural contexts. Eight of these

establishments excelled in the first national evaluation. The purpose of this study was to uncover the attributes of those establishments that inspired high attainment levels in underprivileged children. In particular, the dimensions explored were: (a) institutional characteristics, (b) pedagogical focus, and (c) linkages to the family and surrounding community. This study has become the basis for a comprehensive pedagogical proposal for “full-time” schooling to be implemented in 10% of public educational establishments serving the poorest children in the nation (Uruguay—ANEP-MECAEP, 1997).

School supervisors

The school inspectorate is organized hierarchically from the national-central level to the departmental-regional level to the local-zonal level. Although theoretically organized in a decentralized fashion (Macedo, 1995), school supervisors abide closely by the mandates established centrally at the Technical Inspection unit of the Primary Education Council (World Bank, 1994).

The supervisory cadre has a long tradition of evaluative activities at the school level. Schools are required to self-design and self-administer initial, mid-year and final exams in mathematics and language at all grade levels in order to appraise academic attainment. Inspectors must report on student test scores and specify the percentage of students that can master specific competencies, such as oral expression, orthography, reading comprehension, production of a text, resolution of algorithms, or recognition of geometrical figures. (Note 9)

In addition, inspectors are instructed to conduct their own institutional assessments in order to look beyond academic achievement as “the only objective testimonial of the level and quality” of educational services (see, for example, Uruguay—ANEP-CEP-Inspección Técnica, 1991a; 1991b; 1991c). They collect data on a wide a variety of measures related to educational quality, including student attitudinal qualities (respect, self- confidence, tolerance), absenteeism rates, repetition rates, classroom pedagogical approaches, availability of didactic materials, in-service professional development opportunities, and extent of parental involvement (see, for instance, Inspección Departamental de Montevideo, 1998). Supervisors produce a comprehensive school profile on the basis of this information and elaborate in conjunction with school authorities a strategic plan to address the shortcomings identified in this process.

The national assessment conducted by UMRE summed itself to the battery of school diagnostic information available to the inspectorate. UMRE elaborated reports tailored for the supervisory cadre categorizing schools by sociocultural context and performance level (UMRE, 1998c). Supervisors also had access to the scores of the schools under their tutelage. UMRE developed a series of workshops to familiarize inspectors with the results of this standardized evaluation and suggest potential courses of action that they may take to enhance educational quality.

Overall, the inspectorate gives high marks to UMRE. They underscore that it “has been extremely useful” (Interview UGN16) and has spurred a transformation throughout the educational system at various levels.

We discovered that, it is important to have these data at the national level. In second place, this information is not only useful for the [educational] system, but for schools themselves. There were certain guarantees respected of all operations conducted. [The assessment] is not assigning blame in the face of potential deficits or anything like it. It is simply an objective measure that goes beyond [curricular] contents, and looks at much broader

processes. ... In general terms, everybody is conscious that this is something valuable (Interview UGN9).

The assessment was a starting point to begin to understand the weaknesses in schooling, particularly for low-income children. Furthermore, it paved the way for the adoption of specific remedial actions to address these shortcomings.

This mass evaluation of [student] achievement has put on the table quite clearly what all teachers have been perceiving for many years: how little children in situations of social exclusion learn. The evaluation took into consideration the educational level of the mother, home crowding, or the number of children in the family. [This systematized information] gave us, at the educational system level, some tools to correct in part this situation of low student achievement by updating teachers ... and proposing useful strategies in the areas of psychology, language and mathematics. From this point of view, it served an important professional upgrading role throughout the nation. It allowed many teachers to connect [with their students], because many knew that things were going poorly but it wasn't clear the reason why. It was useful to find new pathways (Interview UGN10).

Supervisors praise the technical reports and pedagogical recommendations put forward by UMRE. They are described as “filled with proposals for action” and “based on solid theoretical foundations” (Interview UGN15). “For me, [UMRE] has been very advantageous because of the exchange of materials. Their contributions are very helpful ... Really, they have been a great technical support” (Interview UGN13).

The national assessment has also served as a model towards a new educational paradigm. Traditionally, educators have emphasized memorization drills of curricular contents. The UMRE test, instead, moved away from appraising curricular contents to assessing competencies. A supervisor suggests that the UMRE test “took place precisely at a time when other pedagogical changes were taking place, and UMRE was able to appropriate itself of all this ... and motivate a re-elaboration of [educational] processes” (Interview UGN23).

Inspector 1. [UMRE] moved us. It put us into contact with [new] literature, with another modality of evaluation that in turn implied another modality of [curricular] planning (Interview UGN19).

Inspector 2. The results obliged us to think about the way curricular proposals were being implemented in educational establishments and how children were learning. The failure of students ... suggested that perhaps it was necessary to reformulate the educational project (Interview UGN14).

The inspectorate has played a crucial role in bringing the lessons from the first national evaluation into the classroom. Across regions, the supervisory cadre was required to organize in commissions to reflect upon student outcomes and devise plans of action that responded directly to the needs identified. These sessions focused on “the role and mission of the inspector” as a catalyst for change (Interview UGN12).

The departmental inspector asked [us] to conduct a study, an analysis of the results, and see what we, as a departmental inspectorate, could do. I was recently reviewing this, and we had accorded to work with institutional

projects ... Every supervisor, following these general guidelines, could request for funds to implement an intervention project in reference to the [UMRE] test results (Interview UGN30).

Inspectors were encouraged to adapt the guidelines outlined in departmental commissions to the social realities of the establishments they oversaw. In certain localities, supervisors organized 2- to 3-month seminars “to support educators with the findings of new research, and a theoretical framework” that delved not only on how students learn but what is relevant learning (Interview UGN13). In most districts, the favored approach has been to intercede directly with school administrators. “We work on specific proposals with our principals, who in turn pour this effort into institutional projects developed together with their teachers” (Interview UGN21). Lower scoring schools have received preferential attention over higher achieving establishments.

There is a growing sense that UMRE has imbued the educational system with a reflexive and renovating spirit. Regardless of the actual transformations that may have occurred as product of the first national evaluation, supervisors concur that UMRE has been responsible for bringing to the fore a national dialogue on the effectiveness of educational services and practices.

Personally, I perceive that there have been changes. Changes in the good sense. There has been an evolution, in theory and in practice. There is a theoretical discussion about [educational] issues, which gets translated into daily activities. ... I have never seen such quick change. I believe this is positive (Interview UGN20).

Despite this strong endorsement to the work and outcomes of UMRE, inspectors do express reserve towards the national evaluation system. First, they underscore that the measurement of student achievement is not a new activity in the Uruguayan educational landscape “We have always evaluated,” attests one supervisor unequivocally (Interviews UGN22).

Inspector 1. In terms of evaluation, I believe that teachers have been working a lot previously on this subject. And so have inspectors. Yes, I share with others that the [UMRE] materials we received have triggered reflection among educators, but I believe that we have been working continuously on evaluation (Interview UGN17).

Inspector 2. I suggest that it is not new to evaluate. [The UMRE assessment] is not new nor is it the only kind of evaluation. Of course, this was an evaluation at the macro level and by an external agent to the school. But we have never stopped evaluating within schools because this is inherent to teachers' practices: evaluating, planning, and researching (Interview UGN15).

Second, the supervisory cadre is concerned about the lack of coordination between the central Technical Inspection and the national assessment. Although all levels of the inspectorate (national, departmental and zonal) are represented in UMRE's advisory council, some supervisors protest that there has not been sufficient participation or communication between the two agencies.

There is a need to polish certain instances [of participation] so that they are

truly effective. Sometimes it is not enough to say that we are participating, that we want to participate. It is necessary that these spaces be created. The possibility is not always present. ... The will has been there, but the spaces are not instrumented so that we can actually share our opinions (Interview UGN12).

Ultimately, the inspectorate is wary of the overlap between UMRE's and their own functions. Supervisors stress that the national evaluation does not supersede their role in the education sector. "I believe [the UMRE evaluation] was a new thing for the educational system, but under no circumstance it precludes the other kind of evaluation that we have been conducting. They are complementary" (Interview UGN15). Some suggest that UMRE is an external agency that has unfairly arrogated their jurisdiction.

The fact is that UMRE belongs to an organism that is called MECAEP and that is parallel to the normative system. It is alien to the Primary Education Council and to the [educational] system. Even though one may value some of the actions that they perform, we can't stop feeling this way. It is not an evaluation generated within the Primary Education Council. It comes out of an external organism. I believe this is one of the issues that produces great aggravation (Interview UGN10).

Others remark that UMRE has been unabashedly displacing them with an agenda of which they claimed to have no knowledge. "[UMRE] was coming above us. Sometimes we didn't even know what they were doing" (Interview UGN16). And yet others claim that UMRE oversteps the separation of responsibilities between the autonomous councils of the National Administration of Public Education (Interview UGN12).

Inspector 1. Over the entire evaluative history in our country the ones that always performed a pedagogical review, a study, were the supervisory cadre and the Primary Education Council. Presently, that review is being done externally. We now wonder repeatedly, as inspectors, to what extent it is valid that somebody else comes along, with other possibilities, with other mechanisms, with more people, to do what we are doing. The measurement performed by UMRE is parallel to the functions of this deconcentrated authority (Interview UGN13).

Inspector 2. The issue is that [UMRE and the inspectorate] each have their own lines of action. The inspectorate has a very clear agenda. But these lines of action get intercepted. Supposedly, UMRE ought to be an advisory or collaborative board in support of our activities. But if their actions are intercepting ours, or we are being displaced by UMRE, then that is where things are starting to become unwound (Interview UGN16).

In summary, supervisors object to the fact that UMRE is an external agency to the inspectorate with comparable functions. They resent that UMRE has had the ability to act independently, the authority to command the attention of educational establishments and the resources to implement directly remedial activities. To some extent, UMRE has come to embody a potential threat to the supervisory cadre. In a few schools, teachers even give credence to the rumor that the supervisory cadre will disappear or that it will be restructured. These criticisms notwithstanding, the general consensus is that UMRE

has been a positive asset and ought to continue the work that it has begun. “A system that does not evaluate itself cannot improve,” remarks an inspector (Interview UGN17). According to the supervisory cadre, it is its organizational structure and relationship to the Primary Education Council that, in their eyes, begs to be redefined.

Principals and teachers

In hindsight, teachers and principals believe that the first national evaluation was an important experience. Private and public schools, as well as low- and high-income establishments concur that the UMRE assessment “was very useful, because it helped us to see where were our flaws, what we can do about them, and how we can change” (Interview UES2).

In its inception, teachers were suspicious of the UMRE test (Interview UGN9). Some expressed concern about whether student performance would be a means to appraise their own professional performance. Others feared that if their students did not attain high marks, they might be transferred to another grade (UMRE, 1996a). The Association of Teachers of Montevideo (ADEMU) expressed its rejection and opposition to UMRE. ADEMU protested that this was a test devised by an entity external to the Primary Education Council and supported by international donor agencies. “The economic expenditure that [the evaluation] supposes,” the teachers' union announced in a newspaper communiqué, “does not conform to the austerity criteria that govern the education budget” (*El País*, 1995). The Uruguayan Federation of Teachers (FUM) also declared deep reservations towards the national assessment.

In the second semester of 1996 and just prior to the measurement, the teachers' union picked up the debate [on the UMRE evaluation]. We reiterated certain existing reparations, about its expense and the degree of dependency to the World Bank's orientations. New elements of concern were also incorporated, like ... the possibility of using the results to categorize schools, to provide differentiated salaries to teachers according to test scores, to stigmatize a certain group of teachers or schools. Also, that it may favor the private sector in some way or other to the extent that it was predictable that public schools would have worse results than private ones. Another series of criticisms were directed to the pertinence of the instruments and the appropriateness of administering one instrument to measure processes in different social realities. Finally, there were concerns about the operational organization in itself, who was going to apply the tests, the access teachers would have, which guarantees existed about the formulation of the tests, the trustworthiness of correction criteria. The criticisms varied from highly ideological considerations, to reserve and distrust, to concerns about the everyday operations of the classroom. There was a wide scope of opinions (Interview UGN37).

Over time, these misgivings were assuaged. Although ADEMU remained defiant to the first national evaluation and encouraged educational establishments to forestall the entrance to exam proctors, teachers and principals collaborated with this governmental initiative. The Uruguayan Federation of Teachers recognizes that UMRE's “open attitude and desire to consult with the teachers' unions and technical-pedagogical assemblies” led to their participation in the Advisory Group and cooperation with the national test (Interview UGN37). An instructor from a rural area recounts that “at the beginning, teachers were not invested [in the evaluation], but during the past year, people started to

talk positively about it” (Interview UES13). A representative from the Association of Private Education Establishments describes a similar experience:

When UMRE appeared, we had a brick on each hand. I was ready to kill them. I had all my reasons against them ready. Little by little they convinced us. Now, after all that has happened and as we get more results, they convince us even more. It is OK that the test is obligatory. It has been a valuable experience (Interview UEM32).

The sense of trust and confidence garnered by UMRE among the teacher cadre can be attributed to four factors:

1. strict confidentiality of test results,
2. prompt devolution of student outcomes to school authorities,
3. contextualization of test scores by sociocultural background, and
4. abstention from holding teachers directly accountable for academic attainment.

Private school principal. Teachers [initially] felt on the spot. There was talk ...that instructors who did not reach certain scores would be removed from office, that there was going to be a public ranking of schools, that this was an attempt to regulate teachers. The people from UMRE were quite clear in explaining what the objectives of the test were. But nobody believed them. Everybody feared that behind this there was something that somehow would harm teachers. ... It is now clear that they kept their word, that it was useful, that it helped us to review things, that two years later we are still working with the results (Interview UEM33).

Public school instructor. Teachers feared that their school would be identified in some manner. And if the school was identified, so would their classroom. And from the classroom, the teacher [would be recognized] ... But the data were confidential. Only we got to know the scores. And the schools were later categorized according to their environment (Interview UEM16).

The national assessment has taken place within an education reform context that has espoused “teacher-friendly policies.” “The appreciation of teacher professionalism and training” has been one of the four pillars of the reform (Rama, 1998). Real average teacher salaries have also risen progressively and consistently starting in 1993, after a period of decline between 1988 and 1992 (Domingo, 1998). (Note 10) This general setting might have contributed to generate a positive disposition among teachers towards the objectives of UMRE and a sense of trust that the evaluation had not been established to monitor their performance or increase their productivity.

Moreover, the attention of educators has not been focused on student achievement measures exclusively. The sociocultural data collected by UMRE were featured as an salient explanatory factor behind student performance. School background information has become a key justification to account for the level of academic achievement attained and an important consideration in the design of relevant remedial actions.

The system of evaluation also conducted a family survey that took into consideration the role of the home in the educational process. We need to take into account that children only spend four hours at school, and twenty at home. The role of the family is fundamental in terms of the contributions

that it can dispense to reaffirm educational processes (Interview UET21).

Teachers report that test scores were subject of repeated discussion and reflection sessions among school inspectors, principals, and the teacher cadre. The organization and participation in these initiatives was mandated by the central government.

The following year, in 1997, when we came back to school, we were required to study the results of the UMRE evaluation, point by point, during our 'administrative days.' Then, we had to draw joint conclusions. It was an obligation to read them. [The order] came to the school in the form of an [official] act (Interview UES15).

The outcomes of the 6th grade assessment were the starting point of a process of pedagogical reflection for a wide range of public and private schools.

Medium-income school. On the basis of the [exam results], we developed a plan for the following year. For instance, the discussion over problem resolution was very important for us in order to go deeper into this issue, to work more on reasoning. I don't know if this took us further away from the [official curricular] program, but ... Also, we've been working on the language [curriculum] in teacher meetings. ... In these sessions we analyzed some of the test items (Interview UET4).

Medium-low income school. [UMRE] identified those competencies that experience the greatest problems. ... We studied the results and worked together with other teachers. We presented the findings in teacher meetings, and discussed the pros and the cons. We devised our [classroom] diagnostic tests at the beginning of the year on the basis of the test outcomes in order to give teachers the opportunity to continue working on these competencies (Interview UES7).

Low-income school. The contents and approach [of the UMRE test] challenged a great deal of ideas that we had. When we saw the exam and what they were after, we came to realize that we were working wrong, that we were working differently, that we were behind, that we were traditional. ... The results and the design of the test (which was a very good proposition) led teachers to realize of everything that we lacked. From here on, we started to review everything, not because we did well, but because we could have done even much better (Interview UEM18).

Rural school. [The UMRE test] does not evaluate for the sake of evaluation, to just get some numbers back. It is meant to improve [our future practices], and to provide feedback. These are problem areas that require hard work and a different approach (Interview UET10).

Private school. I believe this was a very positive experience. It allows teachers to question if they are working well, along the lines they should be working, or if their approach is satisfactory (Interview UES21).

In some establishments, instructors aligned the course curricula according to the

competencies measured by the UMRE test. Others describe that the evaluation triggered greater coordination between grade curricula. And in yet some others, they allude to the design of specific institutional projects to strengthen curricular objectives where students scored poorly.

Low-income school. I liked the approach of the [UMRE] test. It was an interesting proposition. The following year, we planned [our curriculum] on the basis of the approach forwarded by the test. We worked together with another sixth grade teacher on reasoning, geometry, numbers. We did all this basing ourselves on the UMRE test. Last year there wasn't an evaluation, but we administered the '96 exams at the end of the school year. We got a completely different result. In mathematics, it was very good. In language, it was low, but not as low as in the previous year. We even used the same methodology. You could not ask questions to a classmate or the teacher (Interview UEM3).

Rural school. I think that [the UMRE exam] was highly positive to shake teachers up a bit. It led us to question ourselves about many competencies and [curricular] areas that, perhaps, we were not developing well. Throughout the school cycle, students do not receive the same type of education. We might have missed a few steps. These concepts may not have been grasped at the right time and kids drag this handicap into the sixth grade. So the teacher covers the sixth grade curriculum, but oftentimes students do not have clear the concepts or processes necessary to sustain these new concepts. This is all very positive, so that we can all reflect. We are all responsible for specific areas. We have to make sure that students learn certain topics so that the teacher that follows can continue to build upon them (Interview UET11).

Low-income school: Principal. We observed that we needed to start all over again in language, particularly in reading comprehension. One of the factors that exerted incidence on this question was the lack of books at home. ... So we developed a project. *Instructor.* Yes, we developed a project that sought to overcome the current deficits. We called it "A Vegetable Garden to Learn." Through this project we are trying to address the problems detected in the evaluations. ... We find unsatisfactory or insufficient levels in competencies such as production of texts (... which comes to 52%), also algorithms (52%) and problem solving (48%). Those are the competencies with the lowest scores. Hence, we are trying to find solutions to those problems. At the same time, we see the need to continue working on discipline and the formation of good habits. The data from UMRE were particularly useful here. The study showed that we had a 47.6% of aggressiveness and misconduct, and lack of motivation or interest in a 29.9%. Through our little great project "A Vegetable Garden to Learn" we are trying to bring parents into the school and integrate them. Our school is from an unfavorable sociocultural context, and one of the problems that affects much of our functioning is that parents are not involved in student learning (Interviews UES9, UES10).

The impact of the UMRE test on academic practices can be appraised most overtly by how quickly it has become a standard for in-school evaluative practices. Public and

private schools that cater to children from diverse communities manifest that they have modeled their own student assessments after the UMRE test. Some establishments have photocopied the UMRE test and re-administered it. Others have prepared a different test with a comparable methodological approach.

Private school teacher. Some teachers used the [UMRE] test again [the following year]. It was like a re-application. We've also used it as a model for other tests (Interview UES21).

Public school teacher. I came to this school in 1997. I am the sixth grade teacher. Last year we administered something similar [to the UMRE test]. It was prepared here, together with other sixth grade teachers. ... We started to evaluate like UMRE. If the proposal is good, let's do it! I liked the narrative and argumentative text parts of the test in particular (Interview UES8).

Without entering into the discussion regarding the appropriateness or desirability of standardized evaluative practices in the classroom, it is apparent that UMRE's assessment experience reveals the influence nationwide examinations may exert in schooling practices, even in cases where these assessments do not involve *high stakes* testing. Uruguayan teachers adopted the evaluative approach proposed by UMRE despite of the fact that there were no incentive mechanisms or penalties openly associated with this test. It is also opportune to highlight that teachers did not experience this alignment an imposition of the central government or as a restriction to their pedagogical autonomy. They welcomed this methodology for finding it interesting or innovative.

Educators underscore that the type of evaluation proposed by UMRE epitomizes a novel pedagogical approach. Teachers find the emphasis on skill areas and problem solving particularly attractive. On the other hand, they recognize that they lack the know-how to implement it properly. That is, the methodological guidelines forwarded by UMRE are at best an initial referent; in order to be truly effective, they ought to be complemented with specific training.

Instructor 1. The methodological guides say "this should not be like this," but they don't explain how we should do it. *Instructor 2.* There have been radical changes. We studied all our lives one way, under certain methodology. Suddenly, and especially in reading and writing, everything changes. *Instructor 1.* The explanations are very theoretical. Experts prepare these materials, but they remain up there, in theoretical issues. They are not very practical, or clear about how to apply them. *Instructor 2.* [They] first have to come to terms that we are not mathematicians or linguists (Interviews UES14, UES15).

The difficulties experienced in implementing change in classroom practices, according to teachers and principals, have centered around two broad predicaments: (a) lack of capacity, and (b) institutional organizational impediments. These obstacles afflict more acutely the public rather than the private sector, and low-income rather than high-income contexts.

Moreover, teachers lack the institutional space and time to master new techniques or ponder about educational practices. There are few opportunities for in-service training, team curricular planning and professional development. A notable exception

was the seminar organized by UMRE for urban establishments from unfavorable sociocultural contexts.

There are establishments that take into consideration the UMRE data, but there are also establishments that do not take advantage of [this information] not because they do not want to, but because they lack the institutional space for teachers to meet. There is no time for instructors to come together and reflect. It is all left in the hands of the good will of teachers to benefit from the results. ... This is an obstacle. There is an enormous quantity of information, but oftentimes it is wasted. It does not reach the teacher as it should (Interview UGN12).

Public establishments also undergo a frequent and dramatic staff turnover every few years. This has been a standard feature of the Uruguayan educational system. Educators are assigned permanent posts that they periodically vacate to fill in for temporary more desirable positions. This shift causes a ripple effect, encouraging another educator to leave her current post to fill in for that position now open. This permanent flux of school staff interrupts medium-term institutional processes as well as hinders educators from becoming intimately acquainted with local educational and social conditions.

[In 1996,] I was the sixth grade teacher. It was my first year in the school. That year every teacher in the school was new. We had no knowledge of those kids. And neither did the principal, who had been assigned to the school the year before. It took us a year and a half to get to know the school integrally. The only original thing that remained in the school were the students (Interview UES18).

Educators have also professed some objections to UMRE's instruments and methodology. Classroom teachers, and especially those who work with low-income children, criticize the first national evaluation on three counts primarily. First, they object that UMRE depends upon the same instrument to evaluate disparate social realities. It is conceived as intrinsically "unfair" that children from underprivileged backgrounds must face the same exigencies as children that have access to plentiful resources. (Note 11) Second, they protest that exams were administered by outside proctors. The presence of an unknown person in the classroom allegedly distressed and distracted students.

What I objected to was that the classroom teachers could not be the exam proctors. They did not trust us. The job of the proctors was only to distribute the tests, and we could have done that perfectly well. ... There was too much formality, and children are not used to it. ... And that had a negative impact. ... Children were neither at ease nor comfortable in that environment, and that was truly detrimental for them (Interview UET5).

Third, educators claim that unfamiliarity with a multiple choice methodology encouraged students to guess answers or select responses randomly.

In spite of these reservations, UMRE has managed to establish itself quite quickly within the Uruguayan educational landscape. This is a remarkable achievement provided that the evaluation system is barely a few years old. The words of a trade union leader capture this sentiment persuasively:

I believe that at the [educational] system level, [UMRE] furnishes very valuable and interesting information. Although with some difficulties, it has been effectively incorporated into the school culture. The results are valued. The lack of discussion about the application of the third grade assessment immediately demonstrates that it has been incorporated into the school dynamic (Interview UGN37).

Teachers concur that participation in this experience has been beneficial. The UMRE test has, at worst, successfully fostered a dialogue about classroom practices and, in the best case scenarios, stimulated a renewal in pedagogic approaches.

I speak sincerely. Sometimes, when teachers have many years of experience, we find that we must take on other activities outside school. The poor economic conditions oblige us to search for other activities so that we can live with dignity. Hence, suddenly we fossilize in certain aspects, certain methodologies. This test allowed us to see that we can evaluate in a different way. It has become a model. And it gave us bibliography so that we can continue along the path paved by UMRE (Interview UET21)

The first national evaluation has become a model on how to emphasize competencies rather than straight curricular contents. Many educators, in fact, argue that UMRE has taken the lead in educational matters, leaving the old official curricular designs to recede into the background and prompting teachers to challenge long-held assumptions.

Our [curricular] program says Venn diagrams, it says operations, it says reasoning, it says application of knowledge, it says grammar, it says written expression, it says oral expression, it says reading. That is how our programs are currently structured. In the [UMRE] test, it said something else: mother tongue, reflection on language, text production. In the program, it says composition, it does not say written expression. Argumentative text is nowhere. In other words, the program is not what was evaluated. ... The program talks about sentence grammar, ... it talks about subject and predicate, but [UMRE] measured it as contextual grammar. ... We were convinced that we were teaching, but we had not realized that what we had in front was [expected of us too]. With UMRE, we came to realize that not everything that we did was right, that students were not quite responsible [for their shortcomings], that *we* needed to change behaviors (Interview UET7).

In summary, the assessment of educational quality in Uruguay went beyond a mere description of the conditions of schooling throughout the country. It was decidedly a call to action.

3. National assessment and the character of the Uruguayan nation-state

A. Assessment, rationality and State legitimacy

Assessment for rational decision-making.

The UMRE assessments have been designed as a recurrent diagnostic instrument of the characteristics of the Uruguayan education sector. “The evaluation of student learning ... is conceived as a systemic evaluation for feedback purposes” (UMRE, 1997a: 6). The main objective of UMRE is to supply educational actors—policy makers, school inspectors, principals and teachers—with relevant and updated information about student academic performance and the sociocultural variables that may condition it. This information will promote educational quality and equity through two channels. First, it identifies the strengths and shortcomings in education provision. Second, it sets the stage for school actors and government officials to take the necessary steps to correct deficiencies in the efficiency and distribution of educational services on the basis of systematically collected and objective data.

What we endeavor is to produce information regarding ... which skills [students] have mastered and which ones they have not, what pedagogical and institutional strategies have succeeded to instill fundamental learnings in students from the neediest sectors and, finally, where it is still necessary to invest and provide technical assistance to attain a more democratic educational system that benefits all Uruguayan children without socioeconomic distinctions” (UMRE, 1996b: 1-2)

The national government has employed assessment outcomes to shape remediation policies and direct technical and economic resources to those segments of the population in greatest need. Student achievement measures and sociocultural context considerations have played a modest role in the allocation of didactic materials, technical assistance, and funds for school-based projects. The central State, however, has prioritized socioeconomic variables over strict performance standards for redistributive purposes.

UMRE expects to bring about a renovation in pedagogical practices and classroom activities on the basis of the data it collects. Specifically, the assessment system propounds the following objectives:

To make information available about [student] competency levels in areas considered to be fundamental; [and]

To provide that information to teachers so that they can search for pedagogical alternatives that may revert situations prior to the exit of students from the primary educational system (UMRE, 1997a: 5).

Teachers and principals have been formally instructed to review the findings of the first national evaluation and devise compensatory strategies in response to them. The supervisory cadre has been closely involved in this process too, particularly in schools from unfavorable sociocultural contexts.

Assessment and State legitimacy

UMRE has consistently reported and analyzed student achievement outcomes in relation to socioeconomic measures. The first national report underscores the link between test results and background variables (UMRE, 1996g). The second national report is exclusively devoted to the impact of socioeconomic factors on academic performance (UMRE, 1997f). In other words, in Uruguay, the concepts of educational quality and equity are inextricably intertwined. The national evaluation system embodies

another conduit for the central government to fulfill its obligation to reduce the gap between the privileged and underprivileged sectors of society.

[I]t is considered that having information about fundamental skill levels is crucial to recuperate the democratizing role of education. The results obtained in the first national evaluation corroborate that strong inequalities in the quality of learning opportunities exist among students from social environments with great deficits. Although it is known that this is due to a multiplicity of factors, oftentimes external to the educational system, we assume our responsibility for the permanent improvement of the quality of learning. In socially disadvantaged sectors, the mediating function of the school becomes all the more necessary in order to contribute to the personal and social development of children (UMRE, 1997a: 6).

The contextualization of average test scores has become standard practice not just in official documentation, but in the collective mind of educators throughout the country as well. Educational establishments are keenly aware of their own location within the “socioacademic map” and have learned to interpret test results in relation to the social conditions in which the school is inserted.

What is ultimately fundamental? To evaluate [student] linguistic and mathematical competencies, and to precise their family contexts. We have to see what the incidence of the [family] background is [on student achievement] (Interview UET14).

The identification of UMRE with the plight for educational equity has been instrumental for the legitimation of the State's evaluative activities. The collection of student achievement data has validated the reform initiatives of the national government by providing scientific proof of the erosion in the quality of educational services while furnishing rational-technical justifications for the pursuit of these compensatory measures. But perhaps more importantly, UMRE has bolstered the image of the central State as an interventionist agency supporting and tending for the neediest sectors of the population. As a teachers' union leader attests,

[UMRE] ended up inspiring satisfaction. That is, it supplied schools with a depiction of their [academic] situation cross-referenced to sociocultural variables, repositioning results in terms of their contexts. This allows for a type of public stance that is congruous with the trade union's habitual position. Isn't it true? [It refers to] the degree of predetermination and conditioning faced by children as they enter the school. ... In short, there was a national test and there were results of that test that did not merit objections (Interview UGN37).

There are two additional factors that have ratified the validity of the assessment instrument and, *ergo*, the evaluation of educational quality as a legitimate State activity. First, the national evaluation apparatus has been construed as the fruit of a consensual process that has incorporated all of the actors in the educational system, including central government officials, regional and local school inspectors, teacher representatives from the Technical-Pedagogical Assemblies, trade union leaders, and private sector delegates. Second, the State has secured the support of educators by largely circumscribing their liability over test outcomes. “There is not going to be an

index finger accusing anybody,” declared Germán Rama, the National Director of the ANEP, upon the dissemination of test results (*El Observador*, 1996).

Obviously, the deterioration [of the educational system] was not the unique or principal responsibility of teachers. **A multiplicity of factors external to the educational system has been in operation for this to occur:** the mass expansion of education, the deterioration of the quality of life of families, the retraction in educational investments during the military regime, etc. **However, it is necessary to recognize that there are variables internal to the system that affect the quality of student learning:** the pertinence of pedagogical strategies, the relevance of the curriculum, the modalities and expectations inherent in academic evaluations, the fact that schools from the poorest areas are the gateway to the teaching profession, among others (UMRE, 1996b: 1, bold in the original).

The circumscription of teacher liability was accomplished in two ways. First, by showcasing background variables as explanatory factors of academic attainment. “Student learning,” UMRE (1997e: 2) attests, “is strongly stratified as a function of the sociocultural context within which each school operates.” And secondly, by the central State acknowledging accountability over the conditions in schooling services. As established earlier, the national government accepts “*its* responsibility for the permanent improvement of the quality of learning” (UMRE, 1997a: 6, my italics).

The premise that the assessment of academic achievement legitimizes the central State potentially encompasses within itself a paradox. On one hand, evaluation endorses State action by making public its commitment and responsibility over educational processes and outcomes. On the other hand, the measurement of student learning implies a high risk: that poor test performance may provide irrevocable evidence of governmental inefficiency in educational service provision. Thus, if the central State is directly accountable for schooling processes and outcomes, doesn't evaluation jeopardize State legitimacy by calling attention to the deficiencies in schooling?

Sociological institutional theorists posit that assessment is primarily a *symbolic* activity (Meyer and Rowan, 1978). Its main objective, according to this paradigm, is not to produce results or provide relevant data for a diagnosis of the conditions of the education sector, but rather to *appear* that it does. That is, assessment strives to imbue the policy-making process with a guise of scientific rationality. The measurement of academic performance is foremost a legitimizing mechanism of State action by associating the policy-making process with scientific analysis.

Institutional sociologists underscore that attention to test scores may have a deleterious effect by uncovering inefficiencies within the educational system. Consequently, the relationship between assessment and legitimacy depends upon a *loose coupling* between evaluative processes and outcomes. In other words, assessment plays predominantly a figurative role, where the act of evaluating has greater salience than the findings it may uncover. This disjunction blurs the inconsistencies between educational goals and the existing conditions of schooling. In summary, institutional sociologists profess that assessment systems prescribe officially acceptable standards of behavior and operation that uphold State action. On the other hand, these principles that educational establishments professedly embrace are in fact decoupled from the actual organization of schooling.

What do we observe in the Uruguayan case? The central government has reported aggregate test results from the UMRE evaluation at the national level. Student

achievement data were not broken down by department or educational establishment. This practice differs significantly from evaluative experiences in other countries in the region that report testing outcomes by school or by region. Although withholding individual school data may indeed hide inconsistencies in educational service provision, it does not absolve the central government from liability over test outcomes. On the contrary, protecting individual school variability makes the central State the sole publicly accountable agent for educational quality. This strategy would appear to contradict the predictions of sociological institutional theorists. The Uruguayan government's approach to give ample dissemination to test results and advocate reflection over student outcomes, within a context where the central State has accepted responsibility for the quality of educational services, could give way to a crisis of legitimacy for the central government.

National test scores in the first national evaluation were, at best, substandard. Over 65% of students scored unsatisfactorily in mathematics and 43% performed poorly in language. (Note 12) Despite this inferior record, and contrary to common wisdom, UMRE did not delegitimize central State action. The central State, as predicted by sociological institutionalism, shifted the focus of attention from student outcomes to the role of sociocultural variables in academic achievement.

Assessment data fostered a national debate about the impact of socioeconomic forces in educational services. Evidence of the decay of the education sector was primarily a backdrop to champion governmental compensatory initiatives and vindicate the participation of the central State in social policies. The central government could afford to expose the deterioration in schooling because the root causes of the present educational landscape preceded the current administration. These had been already documented in detail in the student achievement studies conducted by the CEPAL in 1990 (Comisión Económica para América Latina y el Caribe, 1993; 1991; 1990).

Moreover, assessment data demonstrated that, controlling for sociocultural context, the performance of public sector schools is equivalent to their private sector counterparts.

[W]hen we take into consideration the sociocultural context within which schools carry out their activities, results vary: public schools that operate in the most favorable contexts obtain results as good as private schools in the same contexts. At the other extreme, rural schools obtain results similar and sometimes even better than urban establishments from contexts equally unfavorable (UMRE, 1997f: 5).

Hence, UMRE asserted the value of public education and, consequently, of State-run educational service provision.

Then, if assessment does not jeopardize State legitimation, are evaluation practices in Uruguay an instance of a loosely coupled system as predicted by sociological institutionalism? That is, is the measurement of student achievement primarily a symbolic activity where evaluative processes are of greater consequence than their outcomes?

Interview and observational data collected for this research study suggest otherwise. In fact, school actors manifest that there is significant coincidence between State mandates around the UMRE evaluation and actual school behavior. In other words, there is evidence that central State action has successfully elicited organizational alignments. Teachers, principals and supervisors alike express a high level of familiarity with assessment policies. In most cases, they have largely complied with regulations to

review and analyze test results. Furthermore, educators concur that this assessment has triggered reflection and some renovation in educational practices. As it has been documented in the section above, some schools have devised institutional projects in response to the findings of the evaluation. Other educational establishments have modeled their classroom and evaluative activities after UMRE's appraisal, focusing for example on competencies rather than on curricular contents.

The degree of influence of the UMRE evaluation on educational practices stands out provided that this is a low-stakes test. There are no incentives tied to performance standards. Neither are educational establishments liable to the public or to the government for student scores. Similarly, a comparison between UMRE's appraisal and other school-based diagnostic evaluative exercises confirms that the former has had quite a distinct impact on classroom activities.

Low-income school instructor. When we conducted our own evaluations, we tested concepts. The type of evaluation of UMRE, it makes you think and balance things out. It leads you to wonder what lies behind [a question]. It is evaluating the process itself. And it is providing feedback to our work. That is what we need to do ... We need to change (Interview UET15).

Two factors can account for this budding transformation in the classroom brought about by the national assessment. First, the evaluation was built and designed with the support and participation of the education community at-large. This process has fostered among educational actors a sense of appropriation and commitment to the work of UMRE. Second, UMRE accompanied its evaluative activities with in-service training workshops for teachers, principals and inspectors. Professional development has catalyzed the patronage and implementation of novel curricular and pedagogical propositions.

In summary, the Uruguayan central State is responsible and accountable for the conditions of the educational system. Assessment may potentially delegitimize State action by underscoring the weaknesses in the education sector. In spite of the shortcomings in schooling services exposed by UMRE, the national government did not suffer a crisis of legitimacy (Weiler, 1990). In fact, the central State was able to rally a wide basis of support behind this initiative. As sociological institutionalism predicts, the central State shifted the focus of public attention from testing outcomes to a comprehensive policy initiative addressing the socioeconomic wants that condition student learning. This displacement, however, did not necessarily decouple assessment from schooling practices. This is particularly striking given that the national evaluation was not designed a high stakes test for students, teachers or principals. The UMRE evaluation acted a conduit to channel the might of the State apparatus behind a pedagogical and curricular transformation.

B. Assessment and State ideology

Uruguay has a long-standing tradition of public support of social sector activities. It has the highest per capita spending on social sectors among Latin American countries. Social expenditures comprise approximately 50% of total government expenditures (World Bank, 1994). The State has been an ardent defender of public education and a champion of the conception of the *Estado docente*—the State as teacher (Fernández, 1997).

Uruguayan education reform program has leaned on two principles: (a) the pursuit

and defense of basic social entitlements, and (b) the resolute participation of the State in the attainment of these entitlements through social promotion and redistributive policies. “The history of Uruguay shows that if you want to change qualitatively a social sector, it must originate from a strong State presence,” remarks a high-ranking government official. “It is unimaginable to think of education reform without the State being an important protagonist” (Interview UGN7).

At an historical junction when the Keynesian Welfare State has been pronounced to be “in terminal decline” (Jessop, 1993: 34), the ANEP frames its vision for central State action in the education sector within this very paradigm. Renato Opertti, the National Coordinator for the Planning Area of the ANEP, portrays the current efforts to transform the educational system along this vein.

The [education] Reform is rooted on a vindication of the Welfare State, in its objectives as well as in its contents; indirectly, it defies reform programs steered by the idea of an auditing and regulating State that “delegates” onto the market the direct provision of services (Opertti, 1997: 146).

Santiago González Cravino, another high-ranking government official, tempers this model of State action, while reaffirming the irrevocable duty of the national government to support the neediest sectors of the population.

In order to attend to disadvantaged people, we need an Interventionist State. In order to favor and sustain the middle class, it is essential, sometimes, the intervention of the State. But the emphasis ought to lie in giving it a more positive and active role, using the private sector as a motivating instrument (González Cravino, 1995: 10).

The education reform program, an initiative born in the context of “budgetary limitations” and “commitments and conditions generated by international organizations,” has been target of harsh criticisms from those that believe that the central State has relinquished its historic role.

This “State” has had no incidence in overcoming the sociocultural deficiencies of increasing student cohorts. Neither have “compensatory” or “focalization” policies demonstrated any ability to surmount ... the true causes of pauperization, marginality and social exclusion (Pallares, 1998: 64).

President Sanguinetti, however, has staunchly defended his agenda for the transformation of the educational system as a “new form of humanist liberalism based precisely on the promotion of equity” (*El País*, 1997).

UMRE has evolved and operated within this framework of State-societal relations. Hence, the assessment system has sought to align its activities with a model of governmental action that promotes the production and distribution of social well-being.

The Uruguayan education reform is statist in its defense of public education. The [UMRE] evaluation is very much linked to this. It is an attempt to promote social policies, to provide services. It is not symptomatic of a retracting [State] (Interview UGN3).

The national evaluation, as already documented earlier in this article, has stressed

the utilization of student achievement information in support of remediation programs intended for disadvantaged communities. As test results have come to light, the central government has assumed responsibility for the conditions of schooling and voiced an institutional commitment to enact a policy agenda to address the shortcomings identified. In this sense, the national evaluation has proceeded in the spirit of *social accountability* and under the currency of social equalization.

The construction of UMRE in these terms has been a deliberate choice. The World Bank, who currently finances UMRE activities, had originally proposed an evaluation system based on a consumer accountability paradigm. The assessment would have operated under a different logic where parents, as consumers of educational services, rely on test results to select an educational establishment for their children (*Últimas Noticias*, 1996). When the administration of Germán Rama took charge of the ANEP in 1995, however, there was a change in strategy and UMRE was shaped after the assessment model that Rama had developed earlier at CEPAL (Comisión Económica para América Latina y el Caribe, 1991)

The association between the assessment system and the World Bank has inspired some mistrust regarding the credibility of the model of State-societal relations espoused by UMRE. In fact, this partnership has threatened to interfere with the legitimacy of the national appraisal. The Uruguayan Federation of Teachers and the Technical-Pedagogical assemblies have expressed opposition to the evaluation of student achievement “because of its international perspective,” associated with neoliberal policies that seek to reduce governmental intervention in the provision of social services (Interviews UGN6, UGN37).

The [Central Board Council] has deteriorated the autonomy of the [Primary Education Council] by assigning functions to a parallel organization (MECAEP). [MECAEP] operates with resources conditioned by international loans and imposes EDUCATIONAL policies that do not respond to the needs forwarded by the NATIONAL TEACHER CADRE (Asamblea Nacional Técnico Docente, 1998: 30, caps in the original).

The national government, in turn, underscores its independence from the multilateral organization and reaffirms to the public opinion its defense of public intervention in the social sectors. “We are not dominated by [the World Bank],” asserts Claudio Williman, the vice president of the Central Board Council, “We are an underdeveloped country where State involvement is vital. ... Education is a competency of the State” (*El Diario*, 1996).

C. Assessment and State control

The Uruguayan educational system is structured in a greatly centralized and hierarchical fashion. All decisions—from administrative matters to curricular frameworks—are determined in Montevideo and uniformly enforced throughout the country. “Teachers in Uruguay behave like an army,” remarks a government official. “If you give them an order, they will follow it” (Interview UGN3). There are extremely limited instances of organizational decentralization or institutional autonomy (Fernández, 1997).

World Bank report ascribes to this “extreme” concentration of power a profoundly deleterious effect.

The highly centralized public primary education system hinders undertaking the required changes to achieve greater sectoral efficiency, equity, and

quality. Centralization has restricted teachers and local managerial authority and initiative, reduced teacher-pupil interaction, discouraged personal growth and professional advancement, and limited the extent to which managerial staff and teacher opinions in pedagogical and administrative matters are solicited and recognized by those in charge of their workplace. On the other hand, centralization has overburdened policymakers and higher level staff with routine tasks and decisions, depriving them from having a more long-term strategic and prospective approach to the sector. The 19 [Departmental Inspectorates] are more concerned with transmitting centrally adopted policies and guidelines and collecting data on behalf of ANEP's central offices than with enforcing activities to enhance the quality of education for which they are ill equipped and trained (World Bank, 1994: 11).

Uruguayan scholars concur that the educational system may benefit from greater flexibility and autonomy in its governance (Pallares, 1998; Fernández, 1997; Macedo, 1995). An initial step in this direction has been the disbursement of small grants to educational establishments for the implementation of school-based initiatives that can enhance educational quality (Uruguay—ANEP-CODICEN, 1998).

Undoubtedly, the national assessment supports the concentration of authority at the central level. As Hans Weiler (1993: 76) proposes, “evaluation is not merely the gathering and dissemination of information; it also has something to do with the authoritative interpretation of standards of knowledge and is endowed with a considerable amount of force, both real and symbolic.” UMRE reinforces curricular mandates pronounced by the ANEP. It also fosters the alignment of school practices with centralized prescriptions. A government informant even claims that UMRE was an attempt from the central State to exert greater control over the flow of information on academic achievement after the release of the highly-critical CEPAL studies (Interview UGN3).

On the other hand, the organization and implementation of the national evaluation defy this depiction of closed centralized control. UMRE has dedicated great effort to the incorporation of an ample array of voices and opinions into this process. It has steadily encouraged the systematic and continuous participation of all levels of civil society. The UMRE Advisory Group consists of representatives from the public and the private sectors, as well as central, departmental and local jurisdictions. Teachers, principals and supervisors have been repeatedly consulted on a wide variety of topics, from the design of the curricular matrix to be appraised to the development of test items.

UMRE's experience serves as a model of centralized governance sustained and enriched by democratic cooperation. The involvement of the Technical-Pedagogical assemblies and the teachers' union in the national assessment is living proof that even unpopular policies may garner the consent of reticent social actors in an environment that nurtures open and effective dialogue.

4. Concluding remarks

UMRE incarnates a model of social—as opposed to consumer—accountability where the central State must respond for the conditions of schooling. In this paradigm, the national government not only functions as a guarantor of educational quality and equity, but it also upholds its obligation as provider of educational services. The

evaluation of student performance is an avenue to defend the role of public education as an equalizing social force and reaffirm the central government's support to the neediest sectors of the population.

However, student performance measures, as already expressed repeatedly, may potentially exert a destabilizing role by highlighting deficiencies in educational service provision. The central State averts the potential crisis of legitimation (Weiler, 1990) by shifting the character of assessment from the measurement of student outcomes to the remediation of the ills in student learning.

The conceptualization of education as a governmental responsibility has largely insulated the assessment process from finger-pointing or assigning blame. It is not teachers or schools that are being tested, but the educational system as a whole. This approach has generated the potential for educators to identify with and participate in evaluative activities. Democratic participation, in turn, buttresses the legitimacy of the assessment scheme.

UMRE has spurred the beginnings of a curricular and pedagogical transformation throughout the Uruguayan educational landscape. This is a promising first step for an evaluation system in its formative years. As new data are collected and UMRE consolidates its role within the education sector, the central State will face a new challenge: It will have to account for the effectiveness of its own policies in reducing existing inequalities.

Notes

1. I would like to thank the Unidad de Medición de Resultados Educativos for their support in this research project. I am indebted to Patricia Arregui, Martin Carnoy, John Meyer, Kathleen Morrison, Karen Mundy, Pedro Ravela and Michel Welmond for helpful comments and suggestions. All remaining errors are my own. This research project was supported by a summer fellowship from the Center for Latin American Studies at Stanford University and a Spencer Foundation Research Training grant. A Spanish version of this article has been published by the Working Group of Standards and Evaluation of GRADE-PREAL and can be accessed in Adobe Acrobat format at <http://www.grade.org.pe/gtee-preal/docs/Benveniste.pdf>
2. This analysis excludes tertiary education.
3. In 1998, the ANEP signed a loan agreement with the World Bank for an additional US\$ 28 million for the second phase of the MECAEP project.
4. This section draws largely from personal interviews conducted with government officials involved in the design and implementation of the primary and secondary national assessment systems.
5. In the 1998 evaluation, teachers were invited to participate in the formulation of test items.
6. Correlation coefficients and standard errors were not provided by the source document.
7. A government informant explains the reasons behind secondary teachers' more contentious attitude in this manner:

The secondary education teacher cadre is very different to primary school educators. The latter is a professionalized group. One hundred percent of [primary school] teachers obtain their degrees. They all went to normal institutes. They all have the title hanging somewhere

at home. Hence, they have a positional culture that is more homogeneous. In secondary schooling, only 30% of the people teaching have specific preparation for being a 'professor.' There are university professionals, university students Thus, the heterogeneity is much greater. ... Secondly, secondary teachers have adopted a "let's see" attitude towards the education reform. Primary teachers were "calmer," more easy going, less opposition. That is why MECAEP, and more specifically UMRE, has been able to secure an active collaboration, inclusive of the teachers' unions and the ATD, the Technical-Pedagogical Assembly. In the case of secondary schools, the unions were more in opposition from the get go, more combative because the education reform was deeper. The ATD is also more politicized. The ATD leaders have emphasized their own ATD position over the stance of the [teachers'] union (Interview UGN3).

8. There is currently some uncertainty regarding the transfer of UMRE from the MECAEP project to the ANEP due to potential changes in the organizational and institutional structure of the evaluation system.
9. The characteristics of these evaluations vary from school to school. Thus, average student test scores are not comparable across schools.
10. Real average teacher salaries, however, are still slightly below their 1988 level nonetheless.
11. This stance contradicts another argument that points at the inherent inequity of holding different expectations for students from dissimilar sociocultural contexts, and particularly of holding lower expectations for children from lower-income backgrounds. The challenge would be not to "veil" the differences among social groups, but rather to introduce the necessary compensatory measures so that all students, regardless of their sociocultural context, can equally reach high achievement levels or national standards.
12. Unsatisfactory test performance was defined by UMRE as inferior to 60% of correct answers.

References

Áde Planeamiento de ANEP (1998). *Incidencia de los Factores Socioculturales en la Repetición Escolar*. Montevideo: Consejo Directivo Central, ANEP.

Asamblea Nacional Técnico Docente (1998). *Resoluciones*. Montevideo: mimeo.

Comisión Económica para América Latina y el Caribe (1990). *Enseñanza Primaria y Ciclo Básico de Educación Media en el Uruguay*. Montevideo: CEPAL.

Comisión Económica para América Latina y el Caribe (1991). *Que Aprenden y Quienes Aprenden en las Escuelas de Uruguay: Los Contextos Sociales e Institucionales de Éxitos y Fracasos*. Montevideo: CEPAL.

Comisión Económica para América Latina y el Caribe (1992). *¿Aprenden los Estudiantes? El Ciclo Básico de Educación Media*. Montevideo: CEPAL.

Comisión Económica para América Latina y el Caribe (1993). *Escuelas Productoras de Conocimientos en los Contextos Socioculturales Más Desfavorables*. Montevideo:

CEPAL.

Comisión Económica para América Latina y el Caribe (1994). *Los Bachilleres Uruguayos. Quienes Son, Que Aprendieron y Que Opinan*. Montevideo: CEPAL.

Domingo, R. (1998). *Evolución Salario Docente Promedio*. Montevideo: Universidad de la República, mimeo.

El Diario (1996). *Estado No Puede Entregar la Educación a Manos Privadas*. Montevideo, November 6-7.

El Observador (1996). *Los Resultados No son Satisfactorios*. Montevideo, November 29.

El País (1995). *ADEMU Informa*. Montevideo, November 14.

El País (1997). *Sanguinetti Ratificó Apoyo para la Reforma Educativa*. Montevideo, July 30.

El País (1998). *Dinero para Universalizar la Enseñanza Preescolar*. Montevideo, October 6.

Fernández, T. (1997). "Educación post-vareliana: Evidencia para nuevas prioridades." *Revista Educación* 6: 63-78.

González Cravino, S. (1995). "La reforma de los servicios sociales del Estado: Una visión global." In Ministerio de Educación y Cultura, (Ed.) *La Gestión en la Educación, Política Educativa, Administración y Financiamiento*. Montevideo: Oficina de Planeamiento y Presupuesto, FAS, 6-10.

González Rissotto, R. (1997). *Sistema Educativo Nacional de Uruguay: 1993*. Montevideo: Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura.

Inspección Departamental de Montevideo (1998). *Evaluación*. Montevideo: ANEP-CEP- Inspección Departamental de Montevideo, mimeo.

Jessop, B. (1993). "Towards a Schumpeterian Workfare State? Preliminary remarks on Post-Fordist Political Economy." *Studies in Political Economy* 40(Spring): 7-39.

Macedo, B. (1995). "Descentralización de la Educación Básica: Las Experiencias de Chile y Uruguay." In U. C. d. U. D. A. Larrañaga, (Ed.) *Problemas y Desafíos en la Educación Básica*. Montevideo: Oficina de Planeamiento y Presupuesto, FAS, 27-29.

Meyer, J. and B. Rowan (1978). "The structure of educational organizations." In M. Meyer, (Ed.) *Environments and Organizations*. San Francisco: Jossey-Bass, 78-109.

Operti, R. (1997). "La reforma educativa: Reinvidicación del Estado Benefactor." *Cuadernos del CLAEH* 22(78-79): 139-159.

Pallares, R. (1998). "Fundamentos y consideraciones para una política alternativa en la educación media." In F. V. Trías, (Ed.) *Reforma Educativa: Análisis Crítico y*

Propuestas. Ediciones de la Banda Oriental.

Rama, G. (1998). *La Reforma Educativa en Uruguay*. Montevideo: Administración Nacional de Educación Pública, mimeo.

Ravela, P. (1997a). "Escuelas productoras de conocimientos en los contextos socioculturales más desfavorables." In Organización de Estados Iberoamericanos, (Ed.) *Cumbre Iberoamericana: Memoria, Parte II. Aspectos Cualitativos y Cuantitativos en la Evaluación Educativa: Una Aproximación al Rendimiento Escolar*. Buenos Aires: OEI, 99-125.

Ravela, P. (1997b). "La búsqueda de escuelas productoras de conocimientos en el marco de la evaluación nacional de aprendizajes en Uruguay." In Organización de Estados Iberoamericanos, (Ed.) *Cumbre Iberoamericana: Memoria, Parte I. Aspectos Cualitativos y Cuantitativos en la Evaluación Educativa: Una Aproximación al Rendimiento Escolar*. Buenos Aires: OEI, 113-135.

Últimas Noticias (1996). *El Banco Mundial Sugiere que ANEP Reduzca sus Controles*. Montevideo, November 5.

UMRE (1996a). *Acta No. 8. Reunión Grupo de Consulta*. Montevideo: UMRE, mimeo.

UMRE (1996b). *Manual del Aplicador*. Montevideo: Administración Nacional de Educación Pública.

UMRE (1996c). *Material Informativo para Docentes: III. Manual de Interpretación Prueba de Lengua Materna*. Montevideo: Administración Nacional de Educación Pública.

UMRE (1996d). *Material Informativo para Docentes: III. Manual de Interpretación Prueba de Matemática*. Montevideo: Administración Nacional de Educación Pública.

UMRE (1996e). *Material Informativo para Maestros y Directores: I. Fundamentos*. Montevideo: Administración Nacional de Educación Pública.

UMRE (1996f). *Material Informativo para Maestros y Directores: II. Aspectos Organizativos*. Montevideo: Administración Nacional de Educación Pública.

UMRE (1996g). *Primer Informe de Difusión Pública de Resultados*. Montevideo: Administración Nacional de Educación Pública.

UMRE (1997a). *Fundamentos y Objetivos de la Evaluación Muestral de Aprendizajes en 3er Año Educación Primaria*. Montevideo: Administración Nacional de Educación Pública.

UMRE (1997b). *Lenguaje: Especificaciones y Sugerencias Didácticas*. Montevideo: Administración Nacional de Educación Pública.

UMRE (1997c). *Matemática: Especificaciones y Sugerencias Didácticas*. Montevideo: Administración Nacional de Educación Pública.

UMRE (1997d). *Primer Informe sobre Producción Escrita*. Montevideo:

Administración Nacional de Educación Pública.

UMRE (1997e). *Programa de Mejoramiento de los Aprendizajes en las Escuelas Públicas Urbanas de Contextos Desfavorables*. Montevideo: Administración Nacional de Educación Pública.

UMRE (1997f). *Segundo Informe de Difusión Pública de Resultados*. Montevideo: Administración Nacional de Educación Pública.

UMRE (1997g). *Tercer Informe de Difusión Pública de Resultados*. Montevideo: Administración Nacional de Educación Pública.

UMRE (1998a). *Acta No. 2. Reunión Grupo de Consulta*. Montevideo: UMRE, mimeo.

UMRE (1998b). *Evaluación de Aprendizajes en Tercer Año de Educación Primaria. Boletín Informativo para Maestros*. Montevideo: Administración Nacional de Educación Pública.

UMRE (1998c). *Uruguay--Unidad de Medición de Resultados Educativos*. Montevideo: mimeo.

Uruguay—Administración Nacional de Educación Pública (1996a). *Nota de Prensa. Primeros Resultados de la Evaluación Nacional de Aprendizajes en 6to Año de Enseñanza Primaria*. Montevideo: Consejo Directivo Central, mimeo.

Uruguay—Administración Nacional de Educación Pública (1996b). *Resolución, Marzo de 1996*. Montevideo: Consejo Directivo Central, mimeo.

Uruguay—ANEP-CEP-Inspección Técnica (1991a). *Circular No. 3*. Montevideo: ANEP-CEP- Inspección Técnica, mimeo.

Uruguay—ANEP-CEP-Inspección Técnica (1991b). *Circular No. 26*. Montevideo: ANEP-CEP- Inspección Técnica, mimeo.

Uruguay—ANEP-CEP-Inspección Técnica (1991c). *Circular No. 30*. Montevideo: ANEP-CEP- Inspección Técnica, mimeo.

Uruguay—ANEP-CODICEN (1998). *1998: La educación Uruguaya. Situación y Perspectivas*. Montevideo: Consejo Directivo Central, ANEP.

Uruguay—ANEP-MECAEP (1997). *Informe de Actividades No. 3 1997 1er Semestre*. Montevideo: ANEP- MECAEP.

Uruguay—Ministerio de Educación y Cultura (1996). *Desarrollo de la Educación: Informe Nacional de Uruguay*. Montevideo: Ministerio de Educación y Cultura, Dirección de Educación.

Uruguay—Ministerio de Educación y Cultura (1997). *Anuario Estadístico de Educación 1996*. Montevideo: Ministerio de Educación y Cultura, Dirección de Educación.

Weiler, H. (1990). "Decentralisation in educational governance: An exercise in contradiction?" In M. Granheim, M. Kogan and U. Lundgren, (Eds.) *Evaluation as*

Policymaking: Introducing Evaluation into a National Decentralised Educational System. London: Jessica Kingsley Publishers, 42-65.

Weiler, H. (1993). "Control versus legitimation: The politics of ambivalence." In J. Hannaway and M. Carnoy, (Eds.) *Decentralization and School Improvement: Can We Fulfill the Promise?* San Francisco: Jossey-Bass Publishers, 55-83.

World Bank (1994). *Uruguay, Basic Education Quality Improvement Project.* Washington, DC: World Bank.

About the Author

Luis Benveniste

The World Bank, 1818 H Street, NW, Room MC8-414B
Washington, DC 20433

Tel.: 202.473.8495

Fax: 202.614.0935

Lbenveniste@Worldbank.Org

Email: luis_benveniste@post.harvard.edu

Luis Benveniste is an education specialist at the World Bank. His recent research work has focused on the politics of student achievement testing in Argentina, Chile and Uruguay. He has also co-authored articles on multigrade schooling and school accountability mechanisms in public and private schools. He has worked in a wide variety of international education projects in Latin America, Western Africa and East Asia.

Copyright 2000 by the *Education Policy Analysis Archives*

The World Wide Web address for the *Education Policy Analysis Archives* is epaa.asu.edu

General questions about appropriateness of topics or particular articles may be addressed to the Editor, Gene V Glass, glass@asu.edu or reach him at College of Education, Arizona State University, Tempe, AZ 85287-0211.

(602-965-9644). The Commentary Editor is Casey D. Cobb:

casey.cobb@unh.edu .

EPAA Editorial Board

[Michael W. Apple](#)

University of Wisconsin

[John Covalleskie](#)

Northern Michigan University

[Sherman Dorn](#)

University of South Florida

[Greg Camilli](#)

Rutgers University

[Alan Davis](#)

University of Colorado, Denver

[Mark E. Fetler](#)

California Commission on Teacher Credentialing

Richard Garlikov
hmwkhel@scott.net

Alison I. Griffith
York University

Ernest R. House
University of Colorado

Craig B. Howley
Appalachia Educational Laboratory

Daniel Kallós
Umeå University

Thomas Mauhs-Pugh
Green Mountain College

William McInerney
Purdue University

Les McLean
University of Toronto

Anne L. Pemberton
apembert@pen.k12.va.us

Richard C. Richardson
New York University

Dennis Sayers
Ann Leavenworth Center
for Accelerated Learning

Michael Scriven
scriven@aol.com

Robert Stonehill
U.S. Department of Education

Thomas F. Green
Syracuse University

Arlen Gullickson
Western Michigan University

Aimee Howley
Ohio University

William Hunter
University of Calgary

Benjamin Levin
University of Manitoba

Dewayne Matthews
Western Interstate Commission for Higher
Education

Mary McKeown-Moak
MGT of America (Austin, TX)

Susan Bobbitt Nolen
University of Washington

Hugh G. Petrie
SUNY Buffalo

Anthony G. Rud Jr.
Purdue University

Jay D. Scribner
University of Texas at Austin

Robert E. Stake
University of Illinois—UC

David D. Williams
Brigham Young University

EPAA Spanish Language Editorial Board

Associate Editor for Spanish Language
Roberto Rodríguez Gómez
Universidad Nacional Autónoma de México

roberto@servidor.unam.mx

Adrián Acosta (México)
Universidad de Guadalajara
adrianacosta@compuserve.com

Teresa Bracho (México)
Centro de Investigación y Docencia
Económica-CIDE
bracho dis1.cide.mx

Ursula Casanova (U.S.A.)
Arizona State University
casanova@asu.edu

J. Félix Angulo Rasco (Spain)
Universidad de Cádiz
felix.angulo@uca.es

Alejandro Canales (México)
Universidad Nacional Autónoma de
México
canalesa@servidor.unam.mx

José Contreras Domingo
Universitat de Barcelona
Jose.Contreras@doe.d5.ub.es

Erwin Epstein (U.S.A.)
Loyola University of Chicago
Eepstein@luc.edu

Rollin Kent (México)
Departamento de Investigación
Educativa-DIE/CINVESTAV
rkent@gemtel.com.mx
kentr@data.net.mx

Javier Mendoza Rojas (México)
Universidad Nacional Autónoma de
México
javiermr@servidor.unam.mx

Humberto Muñoz García (México)
Universidad Nacional Autónoma de
México
humberto@servidor.unam.mx

Daniel Schugurensky
(Argentina-Canadá)
OISE/UT, Canada
dschugurensky@oise.utoronto.ca

Jurjo Torres Santomé (Spain)
Universidad de A Coruña
jurjo@udc.es

Josué González (U.S.A.)
Arizona State University
josue@asu.edu

María Beatriz Luce (Brazil)
Universidade Federal de Rio Grande do
Sul-UFRGS
luceb@orion.ufrgs.br

Marcela Mollis (Argentina)
Universidad de Buenos Aires
mmollis@filo.uba.ar

Angel Ignacio Pérez Gómez (Spain)
Universidad de Málaga
aiperez@uma.es

Simon Schwartzman (Brazil)
Fundação Instituto Brasileiro e Geografia
e Estatística
simon@openlink.com.br

Carlos Alberto Torres (U.S.A.)
University of California, Los Angeles
torres@gseis UCLA.edu