

CARTAGENA DE INDIAS CONFERENCE

Name of case study : SIMCE: Lessons from the Chilean Experience in National Assessment Systems of Learning Outcomes
Date of submission : 2006
Country : Chile
Theme : Education, National Assessment
Authors¹ : Lorena Meckes, Rafael Carrasco
Contact Details : Lorena.Meckes@mineduc.cl
(056 - 2) (390 45 72)

¹ We are very grateful for the insightful reflections and helpful comments of Cristian Cox on the draft of this paper.

Table of Contents

1. EXECUTIVE SUMMARY	3
2. SIMCE: GENERAL DESCRIPTION AND IMPLEMENTATION PROCESS	4
2.1. SIMCE: Organizational and Institutional Relationships.....	5
2.2. The Implementation Process: Political Context and Changes	6
2.3. Main Changes over Time.....	7
3. IMPACT ANALYSIS	8
3.1. The Use of SIMCE in Educational Policies.....	9
3.2. The Use of SIMCE Data by Teachers and Schools to Inform their Practices	12
3.3. The Extent to which SIMCE Helps to Involve Parents and Inform their Decisions	12
3.4. The Use of SIMCE to Foster Accountability at Different Levels of the System	13
3.5. The Impact of SIMCE in Driving Policies to Increase Equity	14
3.6. Threats and Unintended Consequences	14
4. DRIVING FACTORS	15
4.1. Political Economy for Implementation and Consolidation of SIMCE	15
4.2. Institutional Factors for Implementation and Consolidation of SIMCE.....	17
4.3. Factors that Have Influenced Learning and Experimentation	19
5. LESSONS LEARNED	20
6. REFERENCES	29

1. Executive Summary

The SIMCE program is the Chilean national assessment system of learning outcomes. Its purpose is to contribute to the improvement of educational quality and equity by providing information about learning outcomes at the national and school levels in order to (a) monitor results and inform educational policies, (b) provide feedback on school and teaching practices, and (c) foment the involvement of parents and promote their responsibility in the learning process.

SIMCE carries out census-based assessments and publishes the results at both the national and school levels. The census-based assessment is complemented by international sample-based studies about learning outcomes such as PISA (OECD), TIMSS (IEA), SERCE (UNESCO), and CIVED and ICCES (IEA).

SIMCE was created in 1988 under the military regime and was developed and consolidated from 1991 onward under democratic governments. Since then, it has been used extensively to develop educational policies for diagnostic and monitoring purposes, to focus intervention programs on schools with poor results, to design and evaluate policies, and to develop incentives and set goals for improvement. However, use of the information provided by SIMCE has been more limited among teachers, schools and parents.

The main lessons learned after more than fifteen years of implementation are that a census-based assessment system impacts everyone, that published results make everyone accountable, and that both of these elements contribute to develop a culture of evaluation. It also has benefits for policymaking and the implementation of focused interventions, since a census-based assessment system makes it possible to identify specific schools in need of support. However, it is important to consider that there are trade-offs associated with the publication of results for accountability purposes. This is especially true where large social inequities exist, and in highly socially-segregated educational systems such as the Chilean one. In this sense, it is important to ensure there is coherence between the assessment system and the educational policy as a whole in order to compensate for disadvantaged backgrounds and prevent unintended consequences.

SIMCE has also learned that if the information provided by a national assessment system is to be used by teachers, it is not enough to offer school-level information about learning outcomes. Special efforts are needed to promote collective analysis of data within schools, to build a bridge between external evaluation and teacher assessment, and to involve institutions that train teachers.

Taking into account variables that contribute to the sustainability and credibility of the assessment system has also proved to be important. A legal and institutional framework that guarantees the continuity and sustainability of the program is necessary. So is involving key stakeholders in the definition of the

goals, core features and priorities of the system so that they can contribute to their legitimacy. In the case of Chile, this did not happen initially. Persistent administration of tests and commitment to transparency in the publication of results are also key to the system’s credibility.

2. SIMCE: General Description and Implementation Process

SIMCE is the Chilean Ministry of Education’s assessment system of learning outcomes. Its aim is to improve the quality and equity of education by providing data about learning outcomes at the national and school levels, making it possible to:

- Monitor and inform decision-making in educational policy.
- Provide feedback to schools in order to foster improved teaching practices.
- Foment the involvement of parents and school communities highlighting their role in the learning process.

Currently the SIMCE Unit is in charge of national testing as well as the international educational surveys in which Chile participates. The national tests that are administered provide information about students’ performance relative to the country’s National Curriculum Framework (*Marco Curricular*). International surveys complement this information by providing data about Chilean students’ performance compared to that of students in other countries and in relation to the abilities and contents that are considered relevant at the international level.

National tests are administered on a census basis. The table below shows their main features:

Table 2.1: Main Features of SIMCE: National Census Assessment

Focus of Assessment	Learning outcomes related to National Curriculum		
Grades Assessed	4 th , 8 th and 10 th		
Frequency	4 th grade: every year / 8 th and 10 th grades: every two years (on an alternating basis)		
Subjects Assessed	Spanish (reading and writing) and Mathematics in 4 th , 8 th and 10 th grades / Natural and Social Sciences in 4 th and 8 th grades		
Background Questionnaires	Parents, teachers and school directors ²		
Coverage	4 th grade: 96% of students and 90% of schools (*) / 8 th and 10 th grades: 97% of students and 99% of schools		
Information Reported and Target Groups	Audience	Information delivered	Kind of Report
	Teachers and school directors	Results at the national, municipal and school levels; school-level results compared with those of previous years, similar schools, and regional and national results (**)	School report

² Parents fill out questionnaires every year, but teachers and school directors do not do so as frequently.

	Parents	School results compared with previous years, similar schools, and regional and national results	Parents report
	General public	Results at the national level that take into account contextual information (i.e. results according to schools' student selection practices, teachers' expectations, etc.)	Press conferences, media
		Results at the national, municipal and school levels; school results compared with those of previous years, similar schools, and regional and national data	Newspapers, web page (www.simce.cl)
	Researchers	Database for research purposes	

(*) Regarding the 4th grade assessment, the 10% of schools that do not take the test are very small, rural and isolated schools that represent 3 to 4% of enrolled students. Most of these schools only offer 1st to 4th or 6th grades.

(**) From 2006 onwards, 4th grade results will be referred to *performance standards*.

SIMCE complements its census-based tests with sample-based surveys. This was the case with the national diagnostic test of students' performance in English as a Second Language that was administered in 2004. The international studies in which Chile participates are also given to samples of students.

Chile has participated in TIMSS (1999 and 2003) and CIVED (1999 and 2000) from the IEA, in OECD's PISA (2000 and 2006), and in two Latin American comparative surveys conducted by UNESCO (LLECE 1997 and SERCE 2006). The objectives of participating in these studies are (a) to evaluate Chilean students' performance in an international context, (b) to inform curriculum development processes with data about international trends, (c) to compare outcomes of different educational policies and strategies, (d) to monitor learning outcome trends at the national level and simultaneously confirm trends reported by the national tests, and (e) to improve national assessment capabilities (Ministry of Education 2004 b).

2.1. SIMCE: Organizational and Institutional Relationships

The SIMCE Unit is part of the Ministry of Education. The main institutions it works with are the Higher Council of Education³ (one of whose functions is approving the national assessment system), regional divisions of the Ministry of Education, universities and international organizations responsible for international studies about students' achievement.

Within the Ministry of Education, the SIMCE Unit is closely linked to the Curriculum Division. This organizational bond seeks to ensure the alignment of national tests with the national curriculum framework, as well as to enrich the curriculum development process with information about actual student performance. Different divisions within the Ministry of Education use SIMCE data in order to

³ This is an autonomous public institution whose main role is to oversee the quality of higher education and to evaluate the national curriculum.

select schools for targeted interventions and for the achievement-based teacher pay incentive program (SNED). SIMCE also depends on the Ministry of Education's provincial divisions for the administration of the national tests and for the supervision of the activities developed for analyzing SIMCE results within each school.

Test development, international survey administration, test marking and data analysis are frequently performed in association with national universities. Chile's participation in international surveys has created ties between SIMCE and international assessment associations and agencies such as IEA, Boston College, ACER and ETS.

2.2. The Implementation Process: Political Context and Changes

The national testing program originated during Chile's military government in the context of extensive social and economic reforms during the 1980s. Among other changes, the educational reform included decentralizing public schools from the national level to local governments (municipalities), introducing financing instruments based on subsidized demand and allowing the participation of private-sector school administrators under the same financing mechanism (Delannoy 2000; Cox 2003). This financing mechanism meant that subsidies to schools would fluctuate in direct proportion to their attendance rates and enrollment. These reforms were aimed at both increasing the participation of the private sector and changing the role of the government in education. In this context, the creation of SIMCE in 1988 had two main purposes: to inform market decisions (parents choosing schools for their children) and to provide quality control for education at the national and, especially, the school level⁴.

During the 1990s, the new democratic administrations based gradual changes on consensus building as a strategy to ensure the longevity of changes and to maintain previous transformations that could be beneficial to social policies. There was a political consensus about the strategic role of education, and public expenditures in this area nearly tripled from 1990 to 2001 (OECD 2004). Moreover, democratic authorities decided neither to reverse the "decentralization" process nor to change the financing mechanism, but instead opted to complement the subsidiary role of the government with a proactive and compensatory approach to equity and quality issues. This led to the implementation of universal targeted intervention programs to improve learning experiences and outcomes, "shifting the balance in favor of the government without abandoning the market" (Cox 2006: p. 4). In this context, the use of the data provided by SIMCE was also shaped by a dual approach to educational policies. For example, the publication of information on individual school results was complemented by the use of school data to focus interventions on the most disadvantaged schools.

⁴ However, market information and school quality control were not implemented until 1995, under a democratic government, when learning outcome information at the school level was made public (Cox 2003).

SIMCE was also used to reinforce the implementation of the curriculum reform developed during the second half of the 1990s, by means of its influence as a high-stakes⁵ assessment of school and teacher practices.

2.3. Main Changes over Time

The main changes that SIMCE has undergone over time are (a) changes in its institutional framework, (b) methodological improvements in monitoring, (c) placement of learning outcomes in an international context, (d) program intensification (in the frequency and coverage of the grades and subjects assessed), and (e) a shift from normative assessment to criterion-referenced assessment.

(a) In 1988, during the military regime, the Ministry of Education asked Chile's Catholic University (*Universidad Católica*) to develop the SIMCE. This was the same team that had previously developed the PER (Educational Achievement Evaluation Program) at the Catholic University from 1981 to 1984. As was dictated by the constitutional law of 1990 that prescribed it, SIMCE was transferred to the Ministry of Education in 1991 under the new democratic government (Schiefelbein, 1992). In 1997, the SIMCE Unit merged with the Curriculum Development Unit, giving birth to the Ministry of Education's Curriculum and Evaluation Division.

(b) In 1998, the methodology used changed from classical theory analysis to item-response theory, and a new scale was introduced in order to have comparable scores over time. The main purpose of this was to change the original goal of providing comparative information about schools' performance for each year tested to monitoring trends reliably over time. Changes in procedures were also introduced to increase the validity of the test⁶.

(c) The participation in international surveys since 1997 has been one of the most important changes in the assessment program. This has made it possible to place Chilean students' performance in an international context, with the resulting impact on the public perception about national learning outcomes.

(d) Throughout the years, SIMCE has gradually incorporated grades and subjects to the testing program. Fourth grade was assessed from 1988 onward, and 8th grade assessment began in 1989. Starting in 1996, 10th grade was included in national tests. Spanish (reading and writing) and mathematics were assessed on a census basis from the beginning of the program. Gradually, natural and social sciences were

⁵ SIMCE can be considered a "high-stakes" assessment program because of the publication of results for individual schools and the related incentives and consequences.

⁶ Before these changes in procedures, the validity of the released items from the 1997 SIMCE test had been strongly criticized by Eyzaguirre et al. (1999).

added for 4th and 8th grades, initially on a sample basis. This was expanded to census-based assessment in the middle of the 1990s. With respect to frequency, until 1996 one grade was assessed per year (either 4th, 8th or 10th grade)⁷. But as of 2006, two grades will be tested each year – 4th grade will be assessed annually, along with either 8th or 10th grade.

(e) Finally, defining performance standards and relating them to national assessments is the main recent development at SIMCE. The 4th grade test of 2006 will report, for each school, the proportion of students performing at each of three performance categories.

3. Impact Analysis

In this section, SIMCE's effectiveness and final outcomes are analyzed in terms of the extent to which they have met the program's objectives and other purposes. Accordingly, we analyze: (a) SIMCE's contribution to the evaluation and definition of educational policies, (b) its use by teachers and schools to inform their practices, and (c) the extent to which it has helped to involve parents as well as (d) foster accountability at the different levels of the system that are responsible for educational outcomes.

In order to provide a context for this analysis, certain information about resources is necessary. Chile invests nearly 7% of its GNP in education. This cost is shared by the public sector (54%) and the private sector (46%). Public investment is approximately US\$ 3 billion⁸ (Ministry of Finance 2005 a). The budget for preschool, primary and secondary education together is approximately 85% of the total public budget for education (Ministry of Finance 2005 b). SIMCE's budget (including national and international assessments) is approximately 0.1% of this 85%, an amount similar to the average budget allocated to each of the Ministry of Education's targeted intervention programs⁹ (such as "P900," the "Rural Schools Program" and "High School for Everyone," among others).

Costs vary according to the grade assessed, because of the number of students participating, and the curriculum areas considered. Taking 8th grade test¹⁰ as a reference (around 300,000 students and near 6,500 schools participating and four curriculum areas assessed), the cost of test development is US\$ 314 thousand (including item writing, the pilot study with a sample of 3,000 students by each item piloted and psychometric analysis); the costs of the final test administration is approximately US\$ 1,409 thousands.

⁷ The exception was 1996, when 4th and 10th grade were assessed.

⁸ In this paper, costs are estimated according to USD, 2005.

⁹ This proportion is considering the 2003 budget. The last two years (2005 and 2006) the budget of SIMCE has increased considerably because of the intensification of testing frequency, the development of performance standards, the participation in international surveys and the increased costs associated to the demands of dissemination of results.

¹⁰ Based on 2004 test.

the data processing (including digital data entry, scanning of extended responses and online marking of them) is US\$ 463 thousand. The cost of analysis and reporting of results (including design, printing and delivery of personalized reports for every school and for parents, publication in the media and SIMCE web site) is approximately US\$ 121 thousand. Finally, the total cost of the test including the staff and other operational items, is around US\$ 3.4 million¹¹. This means that the cost per student assessed in 8th grade is about US\$ 11, considering four subject areas. However, if two subject areas were assessed, costs would not be proportionally reduced. Based on the different costs of 4th, 8th and 10th grades, an average of US\$8 to US\$9 per student can be estimated for the assessment of two curriculum areas.

3.1. The Use of SIMCE in Educational Policies

SIMCE has been used extensively in the development of educational policies. SIMCE data has been used for diagnostic and monitoring purposes, to focus intervention programs on schools with poor results, and to design and evaluate specific and general policies, as well as to set incentives and goals for improvement. Moreover, the findings of international surveys have also had an impact on policy making –especially in the field of curriculum development– and have been used to validate specific policies.

a. *SIMCE Data Used to Focus Targeted Intervention Programs*

Data provided by SIMCE has made it possible to channel technical and economic support from the national government to the most disadvantaged populations as defined by learning outcomes. Since the beginning of the 1990s, SIMCE has been used to define the criteria for the Ministry of Education’s targeted interventions. The main targeted programs for improving educational quality and equity either use or have used SIMCE indicators to select their target schools, as shown by the following tables:

Table 3.1: Targeted Intervention Programs for Primary Schools

Programs for improving quality and equity of Primary schooling	Period	% Primary school enrolment	Uses SIMCE results to select target schools.	Annual budget (million) ¹²
900 Schools Program ¹³	1990-present	11%	Yes	US\$ 5.3
Rural Education	1992-present	6%	No	US\$ 3.5
Priority Schools ¹⁴	2002-present	1%	Yes	US\$ 1.3

¹¹ The fixed costs (staff and others) consider the development of other tasks such as those required for international surveys and the development of performance standards.

¹² USD 2005.

¹³ The program kept its name in spite of the fact that it was expanded to 1,443 schools and to 641,316 students in 2001 (Ministry of Education 2003).

¹⁴ This program focused on 60 schools in Santiago from 2002 to 2005 and will cover 100 schools in four regions in 2006. It consists of technical assistance provided by private institutions (universities, NGOs) under the supervision of the Ministry of Education

Table 3.2: Targeted Intervention Programs for Secondary Schools

Targeted programs for secondary schools	Period	% secondary school enrolment	Uses SIMCE results to select target schools.	Annual budget (million)
Montegrande	1997-2000	5%	Yes ¹⁵	US\$ 6.0
High School for Everyone	2000-present	33%	Yes	US\$ 3.9
Priority High Schools ¹⁶	2007	5%	Yes	US\$ 2.2

The source for the costs of 900 Schools Program, Rural Education, Montegrande and High School for Everyone is OECD (2004). The costs of Priority Schools and Priority High Schools were provided by the General Education Division, Ministry of Education.

Furthermore, SIMCE data has been used in gradually - implemented universal programs in order to start with schools with greater educational needs. This was the case of the MECE programs during the 1990s.

Moreover, JUNAEB, the institution in charge of providing free school meals and other assistance to poor students, also uses SIMCE data to estimate educational risk indicators and focus resources on particular students. JUNAEB aims its efforts at 40% of primary school students (JUNAEB, 2005), and its budget is 5% of the overall public-sector education budget¹⁷.

b. SIMCE as a Tool for Designing Educational Policies

Recently, the government submitted a bill to modify the structure of the school financing system. The purpose of the bill, known as “Preferential Subsidy” (*Subvención Preferencial*), is to increase the financial resources provided to primary schools with the most disadvantaged students. Currently in Chile, approximately 70% of the budget for primary and secondary education is supplied directly to schools’ administrators, providing a predetermined amount (a subsidy or voucher) calculated on the basis of the number of students and their attendance.

In the first stage of the project’s implementation, it is estimated that the subsidy will increase by 50% for 30% of the poorest students from preschool through 4th grade (Ministry of Education 2005 a).

This reform of the financing system was based on evidence provided by SIMCE that revealed that schools with a high proportion of socially disadvantaged students obtained lower learning outcomes. The design of the project also took into account SIMCE data when calculating the amount of resources required to compensate for the socioeconomic disadvantages of these students and thereby allow them to achieve learning results comparable to students of higher socio-economic status (Gonzalez et al. 2002).

¹⁵ Montegrande used SIMCE data not as a criterion to focus the program, but to select schools that showed diversity in their learning outcomes.

¹⁶ In 2006 this program will commence with 50 high schools and will expand to cover 120 high schools in 2007. The total estimated budget takes into account the intended coverage of the program in 2007.

¹⁷ Ministry of Finance (2005).

c. *SIMCE Data Used for the Evaluation of Policies*

Since the early 1990s, SIMCE data has been used as an important indicator for evaluating specific policies and programs, such as P900, Priority Schools and the extended school day, among others (Ministry of Education 2000, Ministry of Finance 2001, Raczynski 2004, Kenneth et al. 2003). These evaluations have been used to make decisions about the coverage and continuance of these programs, in accordance with their contribution to the improvement of test scores (OECD 2004). In the case of P900, for example, the program's purpose was explicitly "to improve the SIMCE scores of schools with low performance and high social vulnerability and contribute to the goals of greater quality and equity in learning outcomes" (Raczynski 2004, p. 122).

The annual publication by the media of national and international assessment results also generates a public evaluation of educational policies as a whole. It has often been stated that SIMCE is considered by the public to be the "thermometer" of educational quality, indicating either "failure" or "success" of the national reform.

d. *SIMCE Data Used for Incentives and for Setting National Goals*

The main incentive program using SIMCE data is SNED, a system of merit-based rewards targeted at teachers. This system increases the salaries of teachers in schools with better educational outcomes and processes. With an annual budget of US\$21.2 million, it reaches approximately 34,000 teachers in 1,800 schools¹⁸. The calculation used to select the schools receiving additional resources is based 65% on SIMCE results (Ministry of Education 2005 b).

In 2000, after the 1999 results were published, a national target based on SIMCE was set for fourth graders in 2005. The proposed goal was to reduce the number of students in the lower performance category by 50%.

e. *SIMCE Data Used for Research and Policy Analysis*

An important part of the national and international reflection and research on Chile's education is based on information provided by SIMCE. One of the reasons that Chile's educational policies are studied internationally is the existence of information that makes it possible to observe the effects of different initiatives on learning results. Research comparing different administration models or the effects of school choice and vouchers use SIMCE data as its (nearly) sole output indicator. Bellei (2005) summarizes the main studies comparing public and private schools in Chile - all of them use SIMCE data to measure effectiveness.

¹⁸ USD 2005.

f. The Contribution of International Assessments to Policy Design

The findings of international assessments have had an impact on policy development. For example, the current proposal for curriculum development was influenced by information about the gaps identified between the Chilean curriculum and the international consensus, as represented by the items selected for TIMSS. Additionally, results and comparative information have been used to confirm and legitimize specific policies, some of which have already been implemented. For example, comparative information about teachers' training and certification provided by TIMSS 2003 has supported policies that reinforce initial and in-service training for teachers in math and science.

Evidence about Chilean students' performance has also been systematically used in the development of the national performance standards that will be tested in 2006.

It is too early to evaluate the use of results from international assessments for monitoring purposes because Chile has only participated in two assessments in the case of TIMSS (1999 and 2003) and thus can only analyze trends in learning outcomes in this one case¹⁹.

3.2. The use of SIMCE Data by Teachers and Schools to Inform their Practices

According to CIDE (2003) only 15% of teachers and school directors considered SIMCE data "not very important or not important at all." The same question asked a year earlier indicated that 40% of teachers agreed with this statement (García Huidobro 2002). Recent information gathered through teacher surveys suggests that approximately 80% of teachers said that their schools had developed activities to analyze the results of the 2003 test (SIMCE teachers' questionnaire, 2004), and between 90% and 95% responded that they had used the materials that SIMCE provides to better understand the test from the year in question and that they had used sample tests to diagnose their students' learning level (SIMCE teachers' questionnaire, 2005).

Despite the evidence shown above, it is necessary to stress that teachers initially perceived SIMCE as a tool for external control. As the OECD review (2004) suggested, reports were not as user-friendly as they should have been, and performance standards are needed to clarify the meaning of scores.

Nevertheless, it is not clear to what extent knowledge and analysis of SIMCE data have an effect on school practices.

3.3. The Extent to which SIMCE Helps to Involve Parents and Inform their Decisions

¹⁹ In 2006, two additional international tests will be administered: PISA 2006 and SERCE. PISA 2006 will provide the opportunity to compare results with PISA 2000.

According to data gathered through SIMCE surveys, parents have increased their knowledge about results from their children's schools. 21% of 10th graders' parents responded that they knew what the 2001 SIMCE results were, and 26% said the same regarding the 2003 test. Meanwhile, 41% of 8th graders' parents said they knew what the SIMCE results were (SIMCE parents' questionnaire, 2004).

Regarding the use of SIMCE data to select their children's schools, the evidence shows that it is not the main criterion used by parents. According to Elacqua (2004) less than 1% considered SIMCE scores as the main indicator for choosing a primary school for their children. This is consistent with SIMCE data (parents questionnaire, SIMCE 2005). As SIMCE questionnaires indicate, parents of secondary students seem to give similar relevance to SIMCE data when selecting a school. Nearly 5% considered SIMCE scores as one of the three main indicators for choosing a primary school for their children in 2003. The last CEP survey (CEP 2006) indicates that 11% of parents considered SIMCE results as the main criteria for choosing a school.

Finally, it is necessary to stress that parents' knowledge of SIMCE data varies according to their socioeconomic status, being higher for higher socioeconomic families (Ministry of Education, 2003).

3.4. The Use of SIMCE to Foster Accountability at Different Levels of the System

Every year, results at the national level produce a public debate about educational policies as a whole, and the government is literally called to account for the outcomes. In fact, after SIMCE results from 2005 were released in March 2006, the Chilean congress requested that the new minister of education justify and account for the results and what was considered to be too modest an improvement in test scores. Over the years, the demand for improvement in scores has increased. According to a recent survey, 74% of those participating consider that "it is not natural to have no improvement in SIMCE scores when comparing a test to the previous one, and all possible measures should be taken to improve results" (CEP 2006).

Local authorities are also evaluated in terms of SIMCE improvement in their areas. For example, the comparison between regions has been used to call to account and evaluate the success of regional authorities.

At the school level, when scores are high, teachers and school directors use them to promote their schools to parents. However, in the case of schools with low scores, accountability to the community is limited by the degree of knowledge that parents have about the SIMCE scores obtained by their children's schools.

3.5. The Impact of SIMCE in Driving Policies to Increase Equity

Despite efforts and improvement in the reduction of poverty levels, equity issues are still a challenge for Chile. In this context, there is widespread consensus in Chilean society about the relevance of good quality education as a necessary condition for social and economic development. As was stated before, the annual publication of test results has contributed to making education an unavoidable issue on the public agenda.

In addition, SIMCE data (and previously, PER results) has provided and publicized evidence about the close relationship between socio-economic inequities and learning gaps. This information has provided the basis for policies focused on equity during the 1990s, and its publication has increased the social demand for equity in access to quality education. It is revealing that the head of the national teachers union, a member of the Communist Party, stressed SIMCE data as an indicator of equity: “SIMCE results call us to reflect deeply upon Chilean education. We have serious problems of quality and equity”²⁰.

3.6. Threats and Unintended Consequences

The most frequently mentioned unintended consequences of a high-stakes testing program such as SIMCE are teaching for the test, narrowing the curriculum that is implemented, schools rejecting or expelling students in order to raise scores, and labeling of poor performing schools, among others. In the case of SIMCE, there is no systematic evidence regarding these effects.

However, it is important to mention that a decade (1990 to 2000) of publication of SIMCE results directly comparing the performance of public versus private subsidized schools with no consideration of the socioeconomic backgrounds of their students has had an impact on the general perception about municipally-administered schools, reinforcing the general belief that “private services are better than public ones.” In fact, 50% of parents in 1988 agreed that they would prefer a private subsidized school rather than a public (municipal) school. In 2006, 69% of parents agree with the same statement (CEP 2006).

Finally, public expectations about quick improvement could entail the risk of discouraging the development of long-term policies. Similarly, establishing national goals without sufficient evidence of how much improvement can reasonably be expected, given the resources invested and strategies developed, could discredit public efforts in the educational field.

²⁰ “Facing SIMCE results” (2006). http://www.colegiodeprofesores.cl/noticiero/declaraciones_publicas.htm

4. Driving Factors

This section analyses the political, institutional and external factors that have influenced the implementation, continuity and development of SIMCE, as well as the experimentation and innovation factors that have induced change.

4.1. Political Economy for Implementation and Consolidation of SIMCE

a. *SIMCE Initiated by the Military Regime: Easy Implementation and Limited Legitimacy*

As previously mentioned, PER (1981-1984), and subsequently, SIMCE were initiated during the military regime. In fact, it was the government that commissioned the Catholic University to design and develop the assessment system. The authorities did not consult or negotiate with teachers. This approach facilitated the implementation of the system, but limited its legitimacy.

The dominant approach of policymakers and planners during that period considered an assessment system to be essential for providing comparative information on schools to inform parents' choice so as to ensure the functioning of an educational system based on vouchers.

The continuity of SIMCE was ensured by the Military Government through a constitutional law passed in 1990, one day before the democratic government came to power. This law also guaranteed the inclusion of the program in the national budget:

The Ministry of Education must design the instruments for establishing a national system to regularly assess the achievement of the objectives of the national curriculum at the primary and secondary levels. The assessment system must be approved by the Higher Council for Education. The Ministry of Education must implement this assessment regularly and apply the tests, at a minimum, at the end of primary education and at the end of secondary education. The ministry must analyze the data at the regional and school levels. The results must be published in a nationally- or regionally-distributed newspaper and also be placed in a visible location in each school assessed. Under no circumstances shall the publication of results include information on individual students (LOCE 1990).

b. *Legitimizing SIMCE: Continuity, Initial Phase of Low-Stakes Assessment, and Use of Data for Equity and Improvement Purposes*

The previous assessment programs (PER and the initial phase of SIMCE) had already built a tradition of national testing before the assessment system was administered by the Ministry of Education (from 1991 onwards).

During the first five years of democratic government, the above-mentioned law was not fully complied with since results at the school level were not published until 1995. Consequently, SIMCE, in its initial stage, could have been described as a "low-stakes" assessment for teachers and schools. The confidentiality of school outcomes during this period probably reflected the emphasis of the democratic governments of the early 1990s on the use of census-based data for improvement purposes (especially for

the identification of schools in need of support) rather than accountability. These uses of the information strengthened the value of the assessment system for the democratic authorities and mitigated resistance from schools.

The uninterrupted operation of the assessment system, along with the initial confidentiality of test results at the school level and the use of data to focus support strategies for schools, were all factors that helped to create a tradition of national assessment and to legitimize the program.

c. *Publication of Results at the School Level*

Ever since results at the school level began to be made public, the use of this data in nationally - designed interventions began to be complemented by the goal of public accountability at the school level. The publication of results caused mistrust among teachers, especially among those working in disadvantaged social contexts, since their low scores were simplistically interpreted by the public as being the result of poor performance on the part of their schools²¹. SIMCE was considered to be essentially a tool for external control (Ministry of Education 2003). This perception, together with the lack of legitimacy of its origins, probably influenced the poor use of test results among teachers.

From 2000 onwards, SIMCE results at the national level received increased attention in the media, and the lack of improvement in results was interpreted as being due to a “lack of impact” of policies on learning outcomes. Policymakers and some academics started to question the usefulness of an assessment system in which publicized results contributed to the deterioration of the self-esteem of teachers, as well as of the educational system as a whole.

d. *The Reassessment and Consolidation of SIMCE: Involvement of Key Stakeholders through a National Committee*

In 2003, the Minister of Education created a national committee to review and provide recommendations for SIMCE. This committee was diverse, including representatives from different stakeholders (research centers, think tanks controlled by governing and opposition parties, school administrators, policymakers, and the teachers union), and furthermore, it gathered testimony from additional parties. After five months of analysis, the committee concluded that SIMCE was already considered by the public to be a consolidated institution and that its main features should be preserved. Although not every member of the committee was in favor of the publication of results at the school level because of its impact on low-performing schools, it was finally suggested not to reverse this policy for

²¹ This was clearly detrimental for municipal schools since they have a higher percentage of students from lower socio-economic strata.

legal reasons and political considerations. Not publishing results would be perceived as lack of transparency.

The recommendations reflected a compromise between two different views, combining suggestions about the development of SIMCE as an instrument for accountability with its development as an instrument to support teaching practices. On one hand, it was recommended that SIMCE increase the frequency of testing and develop value-added indicators in order to improve the use of data for accountability purposes. On the other hand, it was recommended that published data incorporate performance standards in order to provide teachers with more understandable, and consequently, more useful information.

The conclusions of this diverse committee contributed to the legitimacy of SIMCE. The government presented a development strategy for SIMCE taking its recommendations into account and placing emphasis on the use of SIMCE to improve teaching practices.

e. Looking Forward: Political Demand for More Assessment

In the Chilean case, once the assessment system was installed, the trend has been to increase assessment, not to reduce it. This is probably reinforced by the steady rise of testing observed in the USA, UK and other countries and by the fact that any step backward is likely to be presented by the media and judged by the public as a reduction of transparency or a strategy by policy makers to reduce the impact of poor outcomes on both schools and the government itself. As a case in point, the Chilean Ministry of Education decided to alternate its participation in PISA and TIMSS, omitting the 2003 PISA and the 2007 TIMSS²². The media presented this decision as an attempt by the government to avoid the political consequences of predictably low scores, and congress requested the ministry reconsider the decision.

Considering that “what is assessed becomes valued,” there is political demand to include new subjects as a strategy to increase their relevance. This has been the case with English as a second language, physical education and other subjects.

The political costs of taking a step backwards are likely to be increasingly high, and conversely, intensifying the role of assessment is likely to become a popular strategy to drive educational change.

4.2. Institutional Factors for Implementation and Consolidation of SIMCE

In this section, we explore the main institutional variables, besides political and social factors, that have influenced the implementation and development of SIMCE.

²² This decision was made in advance in light of the need to concentrate limited technical expertise and resources on developing performance standards and intensifying testing while preserving its quality.

a. *Institutional Framework to Achieve Stability, Credibility and Technical Quality*

The question of the best institutional framework for an assessment system and the optimal degree of dependence on the central government has been an issue not only in Chile²³ but also in the rest of Latin America. While stability, political relevance and the use of results seem more feasible when the assessment system has closer institutional ties to the ministry, technical quality and public legitimacy are easier to achieve by autonomous institutions (Ferrer 2006).

The first iteration of SIMCE, developed by the Catholic University, created an environment for research that facilitated the creation of the program and guaranteed its technical quality and independence. However, this institutional framework had already proven to be too fragile to survive political change. That was the case with PER (the “predecessor” of SIMCE), which was terminated in 1984 by a minister’s decision to end its financing.

In 1990, SIMCE was transferred to the Ministry of Education as mandated by constitutional law. This assured the stability and continuity of the program. However, in this context, technical quality and credibility are much more dependant on policymaking priorities. For example, during the first five or six years in which SIMCE was part of the ministry, the national assessment was implemented and became a tradition, but policymaking priorities were not focused on its continuous development or on building the national capacities required to keep it updated. Afterwards, the reform of the national curriculum in Chile immediately underlined the need to align national assessment to its emphasis. This led to a policy decision to further develop the assessment division: expansion of and change in the team in charge of it and participation in international surveys.

The public credibility of SIMCE has not been undermined in spite of the fact that it is run by the Ministry of Education, which is responsible for educational policies that are evaluated with the data generated by SIMCE. Transparency and the persistent publication and dissemination of the results of national and international assessments by the Ministry of Education –regardless of how good or bad these results are– have been key factors in maintaining SIMCE’s credibility with the public.

In summary, a constitutional law creating SIMCE together with its placement within the Ministry of Education have produced continuity, while credibility and technical quality have been a result of transparency of information and policymaking priorities that focused on the deliberate allocation of resources and capacity building.

b. *Institutional Mechanism for External Regulation*

²³ This was one of the issues discussed by the National Committee for the development of SIMCE in 2003.

The Ministry of Education makes the final decisions about international surveys, the timing of tests and the assessment of new curriculum areas. However, these definitions have to be reported to the Higher Council of Education, which has, among other tasks, the mission of approving both the national curriculum and the assessment system. For example, the modifications of SIMCE derived from the curriculum reform and the recommendations of the 2003 committee were reported to and approved by this Council.

c. Strategic Partners and Composition of the Team

At the micro-organizational and management levels, two factors have been important in the implementation and development of SIMCE: the interdisciplinary composition of its team of professionals and the balance between internally- developed tasks and outsourcing of some stages of the process. Currently, the national team is composed of economists (10%), engineers (25% some of whom have experience in the private sector), software developers, psychologists (30%) and teachers (30%). One third of them have postgraduate qualifications. The early introduction of a computer-based marking system (2001), and more recently, an item bank, as well as the contracting of private companies specialized in the printing and distribution of secure products, are all in one way or another a consequence of the composition of the team and the diversity of its members' backgrounds.

Working with external partners (such as prestigious universities and private companies) has extended the institution's capabilities, and also has been a source of innovation and development.

4.3. Factors that Have Influenced Learning and Experimentation

In this section, the main factors that have influenced learning and experimentation practices are addressed.

The national assessment unit's participation in international surveys has been an important source of improvement in learning, as well as a benchmark of the quality of assessment in Chile. The implementation of methodologies for the rigorous study of longitudinal trends (equating), the use of models for marking open ended questions, progress in the quality control of instruments and data and improvement of background questionnaires are all consequences of participation in international surveys. The way international studies report results categorizing performance in levels that describe what students know and how they perform on tests has provided experts, as well as the public at large, with a model with which to compare SIMCE reports of outcomes, engendering expectations of more and better information.

Increased comparative information about assessment systems and their role in the policy agenda of different countries has also been a guiding factor. Debates about experiences in other countries, the frequency of testing, the subjects and grades assessed and the target groups that receive information are common in policy analysis debate and are also a criterion used to assess SIMCE and a potential source of pressure toward change and innovation. The demand for an assessment tool that reports results in terms of performance standards is clearly inspired by an international trend.

Political authorities willing to endorse experimentation and carefully - crafted change have also facilitated learning. In fact, although implementing an assessment system based on performance standards is one of the priorities for the development of SIMCE, the time and resources needed to consult international experts have been provided in order to ensure quality. This has provided the opportunity to discuss the pros and cons of changes and to restrain the quick implementation of solutions that are not sufficiently mature.

A driving factor in this coincidence between political and technical agendas is perhaps experience with the political risks arising from mistakes in assessment.

5. Lessons Learned

LESSON 1: A census-based assessment system impacts everyone and provides information to focus on specific schools, both of which help to develop a culture of evaluation.

Assessment can be considered a communication device connecting the activities carried out in the field of education with society as a whole. National assessment as a communication device is greatly enhanced when every school is affected and issues at the national level have a counterpart at the local level. A census-based assessment system that publishes results on a regular basis has the power to drive public attention toward learning outcomes in a way that sample-based assessments do not. In Chile, the public impact of SIMCE's results is clearly greater than the attention paid to the results of international surveys²⁴.

Making results available at different levels –the national, regional, administrative body and school levels– also highlights the responsibility that each level of the system has for the results

²⁴ For international surveys, additional efforts must be developed in order to impact schools at the local level. An example of this in Chile was the publication of a special guide for teachers to analyze the results of CIVED 1999 (IEA) at the national level and its relationship to curriculum and teaching practices.

achieved²⁵. Public scrutiny of national results and government policies is accompanied by an examination of specific schools which have achieved high results or improved their outcomes.

LESSON 2: Coherence between the assessment system and educational policies must be achieved.

The impact and characteristics of SIMCE demand alignment and coherence between the assessment system and: (a) the national curriculum, (b) intervention strategies for improvement, (c) incentives tied to performance, and (d) social and political demands for accountability

A census-based assessment system with public results provides the opportunity to inform and coordinate diverse policy strategies.

a) *Alignment with the national curriculum.* In Chile, SIMCE has been used to support the implementation of the new curriculum by aligning tests to it. However, aligning test items to the new curriculum is not in itself sufficient to communicate to teachers that there is coherence between SIMCE and the aims of the curriculum reform. The reporting of results, which until 2005 were focused on scores, still does not stress what high or low results mean in terms of the competencies emphasized by the curriculum reform.

b) *SIMCE and intervention strategies for improvement.* Using school assessment data to guide policymaking has the potential to increase programs' effectiveness. An example of this in Chile is the focusing of improvement programs on poorly performing schools and using data to evaluate these programs. However, this demands that the assessment system provide explicit criteria for adequately interpreting the data while taking into account its limitations, by indicating –for example– when score improvement can be considered significant.

c) *Incentives and consequences tied to performance.* When linking consequences or incentives to assessment results in order to foment improvement, it is necessary to make sure that the inferences made about causal relationships and changes in scores are valid. For example, SNED²⁶ takes into account the socio-economic background of students when evaluating school effectiveness²⁷. It is also necessary to make sure that incentives actually produce the desired actions, and not undesirable ones, to obtain rewards²⁸.

²⁵ Although the relationship between central and local responsibilities is a contentious one and reflects the way the role of the state is conceived in education (Benveniste, 2004).

²⁶ An incentive program for teachers consisting of a salary bonus for teachers who work in schools demonstrating effectiveness and improvement in the learning outcomes of their students.

²⁷ Mizala and Romaguera (2002).

²⁸ Carnoy et al. (2005) analyse teachers and school incentives related to SNED index.

Publication of results can also be regarded as a direct consequence for schools. From 2000 onwards, the publication of test results in the press has made it possible to compare similar schools, while taking into account socioeconomic factors.

Besides these considerations about the valid interpretation of results, a policy based on linking incentives or consequences to school results needs to insure, and not only to assume, that the resources and capabilities required to affect or control such outcomes are in place.

d) *SIMCE and social and political demands for accountability.* The public reporting of results draws public attention to equity issues, the rate of improvement and the effectiveness of policy interventions. Therefore, it has the potential to generate new and more informed social demands. This has various implications. The existence and intensification of the assessment of learning outcomes must be well balanced, with comparable strategies to build capacities and plans for improvement. Although there are differing views on this matter, the mere existence of an assessment system that provides public results at the school level is not in itself sufficient to trigger improvement.

LESSON 3: Avoid defining national or school learning targets which are not evidence-based.

Having and communicating well-founded expectations of improvement at the system and local levels, based on empirical evidence and taking into account the resources and capabilities that are available, is important for policymaking. In the case of Chile, the 2000 SIMCE results were used to establish national goals for improvement by 2005 without sufficient evidence²⁹. The government promised that in five years the number of students in the lowest achievement category would be reduced by 50%. This goal turned out to be too ambitious, and thus, had political costs and will presumably lead to reluctance to establish new national targets.

LESSON 4: Do not allow the unintended consequences of national assessment systems to go unaddressed.

Preventing unintended negative consequences from causing members of society to see the assessment system and policymaking as acting explicitly against them is very important. Otherwise, the system loses legitimacy and its potential use is reduced. Although there is no systematic research about the magnitude of the likely unintended negative consequences of SIMCE, mere anecdotal evidence is enough to affect its public image. Thus, even without impact data, it is important to prevent unintended negative consequences when they can be anticipated from other countries' experiences in areas such as selecting students, narrowing the implemented curriculum or inducing low-performing students to be

²⁹ Chile does not stand alone in this challenge. The targets set by the NCLB for 2013 could also be considered to lack evidential support.

absent the day of the test in order to obtain better results. In the case of SIMCE, we can cite two examples of efforts to prevent negative effects: students with special educational needs are identified and their scores are not considered in the average published for the school. In this way, schools are not disadvantaged because of their integration policies. Additionally, parents are encouraged to report cases in which students have been induced to be absent (and they have played a role in this), and the process of scoring considers a comparative analysis of the grades of students absent the day of the test.

However, assessment system efforts are not enough; they require a coherent policy strategy. A good example of this is a bill that would introduce differential vouchers for socially disadvantaged students and prohibit schools from selecting the students they choose to admit based on academic criteria if they make use of this system

LESSON 5: Involve key stakeholders in the definition of the system's goals, central features and priorities.

The literature on assessment systems often stresses the need to clarify goals and the difficulty of conciliating contradictory goals, emphasizing the problems involved in combining accountability and monitoring for improvement objectives (Forster 2001).

Different stakeholders put forward different and frequently contradictory requirements. In the Chilean experience, appointing a national committee with diverse views about education and the role of the assessment system has been crucial to systematically discussing purposes and agreeing on priorities for the future development of SIMCE. For example, it was agreed that the focus of SIMCE should be to assess learning outcomes and not to provide a complete account of education and quality indicators or evaluate intervention programs³⁰. It was also agreed that the first priority in the development of SIMCE should be to develop and assess *performance standards* so as to give meaning to results and thereby make them more relevant for teachers, policymakers and the general public.

This agreement not only made it possible to design a focused development plan; it has also facilitated its public endorsement by different stakeholders³¹.

LESSON 6: There is a potential trade-off between the publication of results for accountability purposes and their use by teachers.

In a highly socially-segmented educational system such as Chile's (OECD 2004) there is an intrinsic trade-off –if not a contradiction– between the publication of learning results by school and the

³⁰ Although it was agreed that the assessment system should provide data for the subsequent evaluation of intervention programs.

³¹ The topics discussed are a good guide for the main issues that any assessment system must solve. See *Evaluación de Aprendizajes para una Educación de Calidad*, Chilean Ministry of Education, 2003.

use of these results by teachers, as those teaching in lower-income areas perceive and interpret SIMCE as being fundamentally unfair (in spite of SIMCE's efforts to the contrary). The contradiction here is between the use of SIMCE by the public and in policymaking and its use by teachers: the more the first demand 'publicity for accountability purposes', the less the latter feel inclined to learn about and use the information and lessons provided by the yearly SIMCE assessments in their own planning and actual teaching.

The lesson here may be that in contexts of strong structural inequalities, publication for accountability purposes should come second to the use of information about students' learning by teachers. This is especially the case if accountability of poorly performing schools is not adequately balanced with sufficient resources and effective strategies to improve those results and to compensate for adverse contexts.

LESSON 7: Build a (strong) link between external assessment and assessment developed within schools.

If the data provided by the assessment system is to inform teachers' practices and contribute to their improvement, a strong link between this information and school-based assessment is needed (National Research Council 2003).

Publishing results and delivering them to each school is not sufficient to ensure their use by schools and teachers. Only recently has the Ministry of Education developed a national strategy for the analysis of SIMCE results. Since 2003, it has requested that schools have their staffs focus on this analysis. However, schools and teachers could benefit still more from the information provided by SIMCE if its linkage to the assessments they develop and the curriculum they teach was made more explicit to them. This linkage could be strengthened by fostering the analysis of test questions and the abilities valued by the national curriculum that are needed to answer them, by developing a common framework for performance standards measured by SIMCE and school-based assessments, and publicizing successful school practices, among other things. Furthermore, "a comprehensive training program to improve teachers' understanding of what SIMCE measures (...)," as an OECD review (2004) recommends, should link this comprehension to the development of teachers' competencies in classroom assessment, especially considering that this was the weakest area identified by the last evaluation of teachers' performance in Chile.

Having a census-based assessment system clearly creates the opportunity to contribute significantly to school practices by providing valuable information. However, this feedback does not occur spontaneously – special efforts must be made to facilitate the use of information and build a bridge of meaning between school and teacher practices and what the assessment system evaluates. An example

of this would be to combine, from the start, external assessment with materials for teacher assessment covering a wide range of curriculum areas for different grades, along with adequate training to foster the use of assessment for learning.

LESSON 8: Consider initial teacher training about assessment to increase impact on teachers' practices.

In the Chilean context, it is clearly crucial to work towards the transformation of the relationship between institutions which train teachers and SIMCE. These institutions are critical in shaping teachers' professional criteria, yet are often ideologically opposed to national assessment systems like SIMCE. The lesson here is that this requires ownership by teachers, which needs to begin with their initial training. A new evaluation culture among teachers must be created as part of a wider effort involving the Ministry of Education, teacher training institutions and continuing education.

LESSON 9: Communicating results is as important as generating them.

The communication of results must take into account that (a) access to information is crucial, (b) the presentation of data communicates implied messages and either facilitates or impedes comprehension, (c) communication plans must consider the media effect, (d) there are always trade-offs between transparency and confidentiality, and (e) it is important to monitor the impact of communication strategies.

(a) Providing reports to schools, parents and teachers does not guarantee that the information will reach its target audience. Access to information should be monitored and optimized by selecting an appropriate means or channel of communication. SIMCE sends reports for parents to the schools, which are in charge of distributing them. There is some evidence that schools provide this information when the results are sufficiently good and conceal it when they are poor. Thus, information for parents is also provided on a web page and no longer published in a newspaper with low circulation, but rather in a popular one, despite the increased costs this entails.

(b) The way in which results are communicated strongly influences their interpretation and suggested uses. Even choosing the level of disaggregation of data has an impact since it leads to a certain attribution of responsibility. National data suggests central responsibility, school-level data suggests school responsibility, and classroom-level data suggests teacher responsibility. Providing and highlighting contextual information such as socio-economic factors, both at the school and national levels, also has an

impact on the perception of the control of and responsibility for learning outcomes. However, it would be naive to assume that merely providing information will insure that it is understood and used appropriately, especially if it is not clear and explicit.

(c) The media must be considered one of the target audiences. They can stress some messages with great power while obliterating others. Although most media decisions are based on editors' policies, it is important to insure that no misinterpretations are caused by a lack of information. For example, even when SIMCE publishes school-level results in which schools are classified into five different socio-economic groups, the press still selects the highest-scoring ones when ranking the "best schools." Although it is not possible to prohibit such rankings, it is possible to influence the way the press presents them. Currently, SIMCE provides data on the best results according to the socio-economic background of each school's students.

Making available explicit interpretations of the results which consider contextual information leads to fair evaluations at all levels, and therefore, has an impact on the legitimacy of the assessment system and the information it provides. Schools will be much more willing to work with and use data from a system that takes into account the contextual variables that affect their results.

(d) Decisions about disseminating information must always take into account issues of confidentiality and must strike the right balance between the commitment to inform and the commitment to protect information that could be interpreted or used inappropriately. In the case of Chile, for example, the National Committee for the development of SIMCE (2003) recommended to maintain the confidentiality of individual scores until policies that effectively guarantee every student's right to attend a given school regardless of his or her academic aptitudes are implemented³².

(e) Periodically monitoring the way results are actually used and interpreted is an important factor in the development of the assessment system³³.

LESSON 10: Variables that contribute to sustainability and credibility must be taken into account.

The main variables that have contributed to the sustainability and credibility of SIMCE were initially legal and institutional (see the section on driving factors). These have been supplemented by more than fifteen years of consistently administering yearly tests and delivering information in a timely fashion.

³² In Chile, subsidized private schools (financed by the government) can select or reject potential students based on academic criteria, their parents' religion and other factors. They can also expel students due to poor academic performance and behavioral problems.

³³ There have been two evaluations about the understanding and use of SIMCE results by different target groups since 2000. In addition, the questionnaires administered to teachers and parents frequently ask about their use of SIMCE data.

The (probably unintended) incremental nature of the implementation of SIMCE has also contributed to its installation within the school system: SIMCE began by keeping school-level data confidential and testing two learning areas and two grades. Later, it expanded to include secondary education, added two additional subjects and started publishing school results.

The extensive use of the indicators provided by SIMCE at the school level for diverse educational policies has further reinforced the demand for such data and has consequently strengthened support for maintaining the assessment system. Finally, involving key stakeholders in a consultation process has also played a part in the sustainability of SIMCE.

In addition to the above-mentioned factors, the commitment to make results public every year – regardless of the news coverage this leads to – has contributed to the credibility of the system. Preserving this credibility, however, also requires time-consuming controls to insure the quality of data, especially if changes have been introduced. In the experience of SIMCE, the most important changes implemented in the recent past (changing the measurement methodology and including open-ended questions) caused months of delay in the publication of results for two years. The reliability of scores becomes more important as their visibility and the consequences tied to them increase. Resources for quality control and time for experimenting with changes before implementing them are reasonable investments in the most important aspect of a measurement system: its credibility.

In summary, different variables have played a part in the sustainability and credibility of SIMCE as an assessment system. The institutional framework has facilitated its continuity. Incremental and continual implementation has helped make it part of school culture. The extensive use of data in formulating educational policies has reinforced the necessity of the assessment program. Consulting stakeholders has helped to guide the development of the assessment system with respect to differing needs, and transparency in providing information and allocating the necessary resources and time for quality control have been crucial in maintaining the system's credibility.

LESSON 11: Participating in international surveys increases information and helps to build national assessment capabilities.

Participation in international studies has been crucial in placing Chilean learning outcomes in an international context and has also enriched policy analysis with comparative data about different social and policy contexts. Curriculum design and performance standard development have been informed by the curriculum analysis developed by TIMSS (Ministry of Education 2004b, p 28) and by data about the performance of Chilean and other countries' students in international surveys in general. These studies have shown that our "elite" students do not perform well in comparison to those from other countries (Ministry of Education 2004a) and have provided data about the gap between high and low performers in

an international context. In this way, the national assessment system is complemented by international data about student learning.

Besides providing more information, participating in these studies has contributed to the technical development of the teams involved and has helped improve the national assessment as well.

6. References

- Bellei, C. (2005) *The Private-Public School Controversy: The Case of Chile*. Kennedy School of Government, Harvard University.
- Benveniste, L. (2004): *La Evaluación Estudiantil como Reflejo de la Relación Estado – Sociedad*, en Carnoy M. (ed.) *Las Reformas Educativas en la Década de 1990: un estudio comparado de Argentina, Chile y Uruguay*, Proyecto alcance y resultados de las reformas educativas en Argentina, Chile y Uruguay, Banco Interamericano de Desarrollo, Ministerios de Educación de Chile, Argentina y Uruguay, Grupo Asesor de la Universidad de Stanford.
- Carnoy, M. Brodziak, I. Molina, A. , Socías, M. : *Do Teachers' Pay Incentive Programs Improve Student Achievement Gains? The case of Chile's SNED*, not published.
- CEP (2006) *Estudio Nacional de Opinión Pública N° 52*. Junio-julio 2006, Centro de Estudios Públicos.
- CIDE (2003) *IV Encuesta a actores del sistema educativo*. CIDE, Santiago.
- Cox, C. (2003) *Las Políticas Educativas de Chile en las Últimas dos Décadas del Siglo XX*, en Cox (ed.) *Políticas Educativas en el Cambio de Siglo*, Editorial Universitaria.
- Cox, C. (2006) *Policy Formation and Implementation in Secondary Education Reform: the Case of Chile at the Turn of the Century*. Education Working Paper Series, Number 3, The World Bank.
- Delannoy, F. (2000) *Education Reforms in Chile, 1980-98, A lesson in Pragmatism*, Country Studies, Education Reform and Management Publication Series, Vol I. N° 1, June.
- Elacqua G. and Fábrega R., (2004) *El consumidor de la educación: el actor olvidado de la libre elección de colegios en Chile*. Escuela de Gobierno, Universidad Adolfo Ibáñez.
- Ferrer, G (2006) *Sistemas de Evaluación de Aprendizajes en América Latina, Balance y Desafíos*, PREAL.
- Forster, M. (2001) *A Policymakers Guide to Systemwide Assessment Programs*, ACER, Melbourne.
- Eyzaguirre, B and Fontaine, L (1999) *¿Qué mide realmente el SIMCE?* Estudios Públicos 75, 107-161.
- García Huidobro, J. et al. (2002) *El programa de las 900 escuelas: Trayectoria de una década*, Gobierno de Chile, Ministerio de Educación.
- García Huidobro, J.E. (2002) *Usos y abusos del SIMCE*. Mimeo (not published).
- González, P., et al. (2002) *Recursos Diferenciados a la educación subvencionada en Chile*. Serie Economía N° 150. Centro de Economía Aplicada, Departamento de Ingeniería Industrial, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile.
- JUNAEB (2005) *SINAE: Sistema Nacional de asignación con equidad para becas JUNAEB*, Gobierno de Chile, JUNAEB.
- Kenneth, Y., McEwan P. and Urquiola, M. (2003) *The Central Role of Noise in Evaluating Interventions*

- that Use Test Scores to Rank Schools*, NBER Working Paper No. 10118, November 2003, JEL No. I2.
- Ministry of Education (2005a) *Todos Pueden*, Santiago, Chile.
- Ministry of Education (2005b) *Hacia la Excelencia Académica, SNED 2004-2005*. Santiago, Chile.
- Ministry of Education (2004a) *Competencias para la Vida: resultados de los estudiantes chilenos en PISA 2000*, Unidad de Currículum y Evaluación, Santiago, Chile.
- Ministry of Education (2004b) *Chile y el Aprendizaje de Matemáticas y Ciencias según TIMSS 2003*, Unidad de Currículum y Evaluación, Santiago, Chile.
- Ministry of Education (2003) *Evaluación de Aprendizajes para una Educación de Calidad*, Informe de la Comisión para el desarrollo y uso del Sistema de Medición de la Calidad de la Educación convocada por el Ministro de Educación.
- Ministry of Education (2003) *Estadísticas de la Educación 2002*, Departamento de Estudios y Estadísticas, División de Planificación y Presupuesto.
- Ministry of Education (2000) *Evaluación del Programa de las 900 Escuelas*.
- Ministry of Finance (2005a) *Proyecto de ley de Presupuestos del sector público para el año 2006*, Dirección de Presupuestos.
- Ministry of Finance, (2005b) *Estudio de Finanzas Públicas*, Marcel, M. and Tokman, C., Dirección de Presupuesto.
- Ministry of Finance, (2001) *Programa del las 900 escuelas, informe final*, Dirección de Presupuesto, Programa de Evaluación de proyectos gubernamentales.
- Mizala, A. and Romaguera, P. (2002) *Evaluación del Desempeño e Incentivos en la Educación Chilena, Cuadernos de Economía*, año 39, n° 118, (353).
- Molina, S. (2005) *La Equidad de la Educación en Chile*, en García Huidobro, J.E. (ed.) *Políticas Educativas y Equidad*, p.205.
- National Research Council (2003) *Assessment in support of instruction and learning: Bridging the gap between large-scale and classroom assessment*. Washington, DC: The National Academy Press.
- OECD (2004) *Chile: Reviews of National Policies for Education*, Centre for Co-operation with non members, Paris.
- Raczynski, D. and Muñoz, G. (2006) *Reforma Educacional Chilena: el difícil equilibrio entre la macro y la micro política*. CIEPLAN.
- Raczynski, D. (2004) *Una experiencia de acción afirmativa: el programa de las 900 escuelas (Chile)*. en García Huidobro, J.E. (ed.) *Políticas Educativas y Equidad*, p.121.
- Schiefelbein, E. (1992) 'Análisis del SIMCE y sugerencias para mejorar su impacto en la Calidad' en Gómez, S. (Ed.) *La Realidad en Cifras, Estadísticas Sociales*, Santiago, FLACSO, INE, UNRISD, pp. 241-280.