Disseminating and Using Student Assessment Information in Chile

María-José Ramírez
Disseminating and Using Student Assessment Information in Chile

María-José Ramírez
Contents

About the Series .......................................................................................................................... v
About the Author ...................................................................................................................... vii
Abstract ......................................................................................................................................... ix
Executive Summary ................................................................................................................... xi
Introduction ...................................................................................................................................... 1
Chile’s Overall Assessment Program .......................................................................................... 3
Dissemination Purposes and Mechanisms .................................................................................... 5
Drivers for Change ....................................................................................................................... 13
Impact of the Assessment Program ............................................................................................. 14
Lessons Learned .......................................................................................................................... 15
References ...................................................................................................................................... 21

Figures
Figure 1. Trends in Students’ Reading Mean Scores: SIMCE, Grade 4 ................................. 5
Figure 2. School Report with Results for Teachers and Principals ....................................... 8
Figure 3. School Results Published in a Chilean Newspaper ................................................. 10
Figure 4. Parent Report ............................................................................................................. 11
Figure 5. Geo-referential System with School Results ......................................................... 12

Table
Table 1. SIMCE Dissemination Strategy: Mechanisms, Purpose, Audiences, and Content ................................................................................................................................. 6
About the Series

Building strong education systems that promote learning is fundamental to development and economic growth. Over the past few years, as developing countries have succeeded in building more classrooms, and getting millions more children into school, the education community has begun to actively embrace the vision of measurable learning for all children in school. However, learning depends not only on resources invested in the school system, but also on the quality of the policies and institutions that enable their use and on how well the policies are implemented.

In 2011, the World Bank Group launched Education Sector Strategy 2020: Learning for All, which outlines an agenda for achieving “Learning for All” in the developing world over the next decade. To support implementation of the strategy, the World Bank commenced a multi-year program to support countries in systematically examining and strengthening the performance of their education systems. This evidence-based initiative, called SABER (Systems Approach for Better Education Results), is building a toolkit of diagnostics for examining education systems and their component policy domains against global standards, best practices, and in comparison with the policies and practices of countries around the world. By leveraging this global knowledge, SABER fills a gap in the availability of data and evidence on what matters most to improve the quality of education and achievement of better results.

SABER-Student Assessment, one of the systems examined within the SABER program, has developed tools to analyze and benchmark student assessment policies and systems around the world, with the goal of promoting stronger assessment systems that contribute to improved education quality and learning for all. To help explore the state of knowledge in the area, the SABER-Student Assessment team invited leading academics, assessment experts, and practitioners from developing and industrialized countries to come together to discuss assessment issues relevant for improving education quality and learning outcomes. The papers and case studies on student assessment in this series are the result of those conversations and the underlying research. Prior to publication, all of the papers benefited from a rigorous review process, which included comments from World Bank staff, academics, development practitioners, and country assessment experts.

All SABER-Student Assessment papers in this series were made possible by support from the Russia Education Aid for Development Trust Fund (READ TF). READ TF is a collaboration between the Russian Federation and the World Bank that supports the improvement of student learning outcomes in low-income countries through the development of robust student assessment systems.
The SABER working paper series was produced under the general guidance of Elizabeth King, Education Director, and Robin Horn, Education Manager in the Human Development Network of the World Bank. The Student Assessment papers in the series were produced under the technical leadership of Marguerite Clarke, Senior Education Specialist and SABER-Student Assessment Team Coordinator in the Human Development Network of the World Bank. Papers in this series represent the independent views of the authors.
About the Author

María-José Ramírez is an Education Specialist in the Human Development Network at the World Bank. She has been working on the SABER-Student Assessment initiative, developing tools for evaluating the quality of assessment systems. Before joining the Bank, she was involved in key reforms of the assessment system in Chile, with responsibilities in both national and international assessments. In the Chilean Ministry of Education, she headed the data analysis unit of the assessment program (SIMCE) (2005–06) and worked as a national coordinator for TIMSS (1998–2000). She also was the College Director of Universidad Diego Portales, Chile (2007–09). She led university-level projects related to quality assurance, institutional analysis, accreditation, and academic evaluation. In the United States, she was a research assistant in the TIMSS and PIRLS International Study Center (2000–04). She received her PhD in educational research, measurement, and evaluation from Boston College (2004) and was the recipient of a Fulbright scholarship (2000) and the 2005 award for best empirical dissertation from the International Association for the Evaluation of Educational Achievement (IEA). Her work and publications focus on student assessments, education quality, and comparative education.
Abstract

Student assessment plays a key role in improving education. However, many countries do not disseminate and make effective use of their student assessment information. This paper discusses Chile’s effective use of its assessment information and draws lessons for other countries. A key lesson from the Chilean experience is that effective use of assessment information is a challenging task that takes time, effort, and resources. Another lesson is that the task requires leadership, long-term planning, and consensus building. Countries able to learn from Chile’s lessons will be in a better position to make effective use of their student assessments and, ultimately, to contribute to education quality and learning.
Executive Summary

The effective use of student assessment information remains a key challenge for developing countries. Countries that have managed to put in place their large-scale assessment may still face important barriers to effectively disseminate and use assessment information. How can countries make more effective use of assessment information? What can they learn from other countries that have more experience using assessment information?

Chile’s assessment program, SIMCE (Sistema de Medición de la Calidad de la Educación), offers important lessons to other countries aiming to better use their assessment information. SIMCE’s dissemination strategy is the product of more than 20 years of experience. Chile systematically uses assessment to inform decision making at different levels of the education system. The country has shown increases in student learning during the last decade. These increases are, at least in part, the consequence of good decision making informed by evidence from the assessment.

SIMCE is a highly institutionalized assessment program. It meets most of the technical standards for educational assessments, and provides the most credible indicator of educational quality in Chile. SIMCE has a comprehensive dissemination strategy that responds to three main purposes: to inform policy, to provide pedagogical support to educators, and to hold schools accountable. For each one of these purposes, SIMCE has in place one or more mechanisms to disseminate information. Each mechanism targets a specific audience with tailored information. The most important mechanisms are the website, the national report, the school report, and the parent report.

Several contextual drivers allowed for development of the SIMCE dissemination strategy. These drivers include the political context in which the assessment took place, the institutional conditions of the assessment office, political pressures to improve the assessment, and international influences. The single most important driver to further develop the dissemination strategy was the SIMCE Commission. This technical-political commission was appointed by the Minister of Education to review the assessment program. The commission built consensus around the purposes of the assessment and the mechanisms to disseminate assessment information. SIMCE has been gradually implementing the recommendations of this commission.

The purpose of this paper is twofold. First, it aims to describe the uses of student assessment information in Chile, highlighting the purposes, audiences, and mechanisms to disseminate information, as well as the drivers that allowed for these uses to take place. Second, the paper aims to draw lessons from the Chilean experience in order to inform other countries seeking to improve their assessment programs. The Chilean experience shows that making effective use of assessment information is a very challenging task that takes time, effort, and resources. A sustainable dissemination strategy needs to be built gradually, and the conditions for the disseminated information to be used needs to be ensured.
Disseminating and Using Student Assessment Information in Chile

María-José Ramírez

Introduction

The effective use of assessment information remains a key challenge for developing countries. Those that have managed to put in place large-scale assessment program may still face important barriers to effectively disseminating and using assessment information. For instance, stakeholders may have negative attitudes toward the assessment; information may be collected but not disseminated; or information may be disseminated, but not understood or used by stakeholders. In all these cases, the assessment would fail to contribute to its ultimate purpose of improving learning and education. How can developing countries make more effective use of assessment information and what can they learn from other countries’ experiences?

Chile’s SIMCE assessment program offers important lessons to other countries aiming to better use their assessment information. SIMCE was the first national assessment in Latin America, and it has been operating on a yearly basis since 1988. The program is highly institutionalized and has been mandated by law since 1990. SIMCE has a comprehensive dissemination strategy aimed at informing policy, providing pedagogical support to educators, and holding schools accountable. Chile systematically uses assessment to inform decision making at different levels of the education system.

Student achievement has been gradually improving in Chile. SIMCE shows that students have been reaching significantly higher mean scores since year 2000, and the OECD Programme for International Student Assessment (PISA) has singled out Chile as the country that most improved its reading results between 2000 and 2009. The 2010 McKinsey report upheld Chile as a “promising star” because of its sustained increase in student learning. This increase is, at least in part, the consequence of good decision making informed by evidence from SIMCE.¹

In considering the lessons from this case study, countries should acknowledge the context from which these lessons come from. Chile is a political and economically stable upper-middle-income country that still needs to reduce poverty and inequality rates. The country’s return to democracy in 1990

¹ All data gathered from OECD’s Education at a Glance report and PISA program (http://www.oecd.org/edu/), Chile’s national assessment program (www.simce.cl), and McKinsey & Co. (2010).
provided the necessary political stability to allow for social and economic growth. Chile is a leader in Latin America for its economic development (gross national income per capita of US$9,950), economic growth (4.1 percent annual growth per capita between 1991 and 2005), and its “very high” human development index. Poverty was reduced from 20 percent to 15 percent between 2000 and 2009. However, Chile is lagging in the distribution of the wealth: its Gini coefficient (52) is the worst among OECD countries.

The education system in Chile is relatively strong when compared to Latin America, but still weak when compared to richer countries. Coverage is universal in primary education and is widely available in secondary education (gross enrollment rate 96 percent). There are 3.5 million students distributed in nearly 9,000 schools, most of which (57 percent) are private. The schools are funded through a national voucher policy, where the state pays a fixed amount of money to the school (no matter if public or private) based on the number of students attending classes. Investment in education is very high when considering public and private expenditures (7.1 percent of GDP), but it is low when considering public investment only (4.2 percent of GDP). The results of national and international assessment programs show that Chilean students are slowly reaching higher performance levels. However, the overall performance remains low compared to richer OECD countries, and is very unevenly distributed along social classes.

The purpose of this paper is twofold. First, it aims to describe the uses of student assessment information in Chile; highlighting the purposes, audiences, and mechanisms to disseminate information; and describe the drivers that allowed for these uses to take place. Second, the paper aims to draw lessons from the Chilean experience in order to inform other countries seeking to improve their assessment programs. The Chilean experience shows that making effective use of assessment information is a very challenging task that takes time, effort, and resources. To begin, a sustainable dissemination strategy needs gradually to be built. Then, an assessment culture and the right institutional conditions need to be developed for the assessment information to be used.

The paper is structured as follows. After the introduction, the next section provides some historical background and describes the current status of Chile’s assessment program. Subsequent sections outline the dissemination purposes and mechanisms currently in place; analyze the drivers that allowed the current dissemination strategy to take place; and discuss the evidence regarding the

---


3 All data gathered from OECD’s Education at a Glance report and PISA program (www.oecd.org/edu), Chile’s Ministry of Education (www.mineduc.cl), and Chile’s national assessment program (www.simce.cl).
impact of the assessment. The last section presents the main lessons to be extracted from the Chilean experience.

**Chile’s Overall Assessment Program**

**History**

Chile’s political context has shaped the purposes, uses, and consequences of the assessment program. Therefore, it is important to highlight that context in order to understand how the assessment information has been used to improve quality, and how these uses have grown in complexity across the years.

Chile’s large-scale assessment program was created under a right-wing dictatorship that ran from 1973 to 1990. The assessment was piloted during the 1980s with the purpose of holding schools accountable by informing parents about school quality. Parent would choose the best school for their children based on school results, no matter if the school was public or private. Therefore, the assessment was developed to serve as an information mechanism in the context of a national voucher policy. According to this policy, the state would pay a subsidy to public and private schools based on student attendance. The assessment program was tasked with improving efficiency, to be achieved by pushing the schools to compete for higher test scores, parent choice, and funding (Majluf 1987).

The assessment program was also created with the aim of monitoring effectiveness of educational policies and supporting pedagogy. Effectiveness would be addressed by monitoring how many students were reaching the performance standards set by the national curriculum. Pedagogical support would be provided by disseminating pedagogical guidelines and reports to school supervisors and teachers (Himmel 1996; Majluf 1987).

With the return of democracy in 1990, the education priorities shifted towards quality and equity, and an ambitious education reform was launched. A new curriculum was introduced, putting a greater emphasis on the learning of higher-order skills and on the capacity of the students to apply their knowledge and skills to real life situations. The introduction of this new curriculum pushed for the alignment of other components of the school system, including the assessment. This alignment involved adjusting the tests so that they measured what students were expected to learn according to the curriculum. It also involved reporting results to teachers in a more meaningful way. SIMCE started reporting the percentage of students reaching different performance standards, and provided descriptions of what students knew and were able to do at each standard level.
Current Status

SIMCE is a technically sophisticated assessment program. The core subject areas of the assessment are mathematics, Spanish language, and social and natural sciences. In recent years, new areas of the curriculum have been added on a sample base (writing, English language, computer skills, and physical education). The test questions are mainly in multiple-choice format, but open-ended questions and essays have gradually been introduced. All schools and students in the target grades (grades 4, 8, and 10) participate in the assessment (census). In 2010, nearly 500,000 students from near 9,000 schools answered the SIMCE tests. Students, their parents, and teachers provide background information (for example, socioeconomic status, opportunities to learn at school, teaching qualifications) through questionnaires. Around 22,500 test supervisors and external administrators are directly involved in the field operation, the success of which requires high levels of expertise.

SIMCE meets most of the technical standards for educational assessments, hence ensuring the validity and reliability of the assessment. The tests are constructed following rigorous procedures to ensure that test questions are an accurate representation of the curriculum. Tests from consecutive years are equated to ensure the comparability of results. Test administration is standardized to ensure the same testing conditions for all the students. Accommodations for special needs students have been gradually introduced. Test results and related information are widely disseminated to reach different audiences and stakeholders. SIMCE has high credibility and has never suffered from a major leakage.

SIMCE has centered attention on student learning, and its results are systematically used to inform decision making. The program’s results are widely covered by the media, usually on the front pages of newspapers or as the main story of the news on television. The assessment results are regularly used for policy design, analysis, and evaluation. Incentives are assigned to teachers and schools based on assessment results. To a lesser extent, schools and teachers analyze their students’ results and make pedagogical decisions accordingly.

Despite important improvements, SIMCE still needs to move forward to meet some important standards of educational assessment. First, SIMCE does not have permanent mechanisms to monitor the consequences of the assessment. That is, the assessment program does not conduct regular studies to monitor how test results are being understood, valued, and used by educators, parents, and policy makers, and how they ultimately impact student learning. The information available about the consequences of the assessment mainly comes from ad hoc studies. Second, SIMCE does not publish a comprehensive technical report and does not provide data files that would allow independent researchers to replicate results.
Dissemination Purposes and Mechanisms

The following sections describe how Chile has used the assessment information to meet three main purposes: to inform policy, to provide pedagogical support, and to hold schools accountable.

Inform Policy

The primary use of the assessment information has been to inform policy. Political pressures to improve education have lead to more evidence-based decision making, especially at the Ministry of Education. SIMCE has been used to monitor quality and equity of education, to design and evaluate intervention programs, and to target resources toward the lowest-performing schools. For example, SIMCE data is used to monitor education quality at the country level (figure 1). The data was also used to evaluate the impact of the SNED accountability program discussed later in this paper (Gallego 2008; Rau and Contreras 2008).

Figure 1. Trends in Students’ Reading Mean Scores: SIMCE, Grade 4

![Graph showing trends in students' reading mean scores](image)

Source: SIMCE 2010 national report.

Both quality and equity indicators are used to inform policy and decision making. The most widely used indicator to inform policy (and the one with the strongest impact in the media) is the change in national mean scores across years. Changes in mean scores for public and private schools and for schools that serve students from different socioeconomic backgrounds also receive considerable attention. The impact of policies in equity is mainly measured by the changes in the achievement gap between lower- and higher-income students.
SIMCE information is disseminated within the Ministry of Education to ensure its use for decision making. Previous to the release of the SIMCE results, the dissemination strategy is widely discussed with the minister and his/her advisors, therefore ensuring that they have firsthand understanding of the results. After the results release, the main findings are presented to the broader staff at the Ministry of Education. The mechanisms used to inform policy are the SIMCE national report, the press kit, the data files, and website (table 1).

Table 1. SIMCE Dissemination Strategy: Mechanisms, Purpose, Audiences, and Content

<table>
<thead>
<tr>
<th>Assessment guidelines (since 1988)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Purpose:</strong> Provide pedagogical support.</td>
</tr>
<tr>
<td><strong>Audience:</strong> School principal, pedagogical coordinators and teachers.</td>
</tr>
<tr>
<td><strong>Content:</strong> (a) Assessment framework and its relationship to the national curriculum. (b) Examples of test questions with an analysis of the contents and skills required to answer them correctly.</td>
</tr>
<tr>
<td><strong>Others:</strong> Distributed to all schools before the assessment (usually in the middle of the school year). Also available online. Publication highly valued by teachers.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>School report (since 1988)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Purpose:</strong> Provide pedagogical support.</td>
</tr>
<tr>
<td><strong>Audience:</strong> School principal, pedagogical coordinators and teachers.</td>
</tr>
<tr>
<td><strong>Content:</strong> (a) National-, school-, and class-level mean scores by subject areas and grades tested. (b) Differences between school mean scores and mean scores from the previous assessment, from the national mean, and from schools of the same socioeconomic group. (c) Percent of students by performance level—advanced, intermediate, beginner. (d) Examples of test questions with an analysis of the contents and skills required to answer them correctly. (e) Workshop guidelines for the schools to analyze assessment results and set improvement plan. (See figure 2.)</td>
</tr>
<tr>
<td><strong>Others:</strong> Distributed to all schools that participated in the assessment once the SIMCE results are released (usually at the beginning of the next school year).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>National report (since 2006)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Purpose:</strong> Inform policy.</td>
</tr>
<tr>
<td><strong>Audience:</strong> Decision makers, general public.</td>
</tr>
<tr>
<td><strong>Content:</strong> (a) National and regional mean scores in subject areas and grades tested. (b) Percent of students by performance level—advanced, intermediate, beginner. (c) Mean scores by socioeconomic background, gender, public/private school. (d) Trends in mean scores across years.</td>
</tr>
<tr>
<td><strong>Others:</strong> Distributed at the central, regional, and provincial offices of the Ministry of Education. Distributed to persons likely to be interviewed by the media (e.g., university professors).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Newspaper supplement (since 1995)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Purpose:</strong> Hold schools accountable.</td>
</tr>
<tr>
<td><strong>Audience:</strong> Parents, general public.</td>
</tr>
<tr>
<td><strong>Content:</strong> (a) School mean scores, and mean scores by subject areas and grades tested. (b) Differences between school mean scores and mean scores from the previous assessment, from the national mean, and from the mean of schools from the same socioeconomic group.</td>
</tr>
<tr>
<td><strong>Others:</strong> Published in a newspaper with national and regional coverage. Usually accompanied by rankings of schools.</td>
</tr>
</tbody>
</table>

(Table continues on next page)
### Table 1 (continued)

<table>
<thead>
<tr>
<th><strong>Parent report (since 2005)</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Purpose</strong>: Hold schools accountable and involve parents in school.</td>
<td></td>
</tr>
<tr>
<td><strong>Audience</strong>: Parents.</td>
<td></td>
</tr>
<tr>
<td><strong>Content</strong>: (a) School mean scores, and mean scores by subject areas and grades tested. (b) Differences between school mean scores, and between subject area/grade mean scores of schools from the same socioeconomic group. (c) Percent of students reaching different performance standards. (d) Recommendations to support student learning. (See figure 4.)</td>
<td></td>
</tr>
<tr>
<td><strong>Others</strong>: Distributed to parents through the schools once the assessment results are released (usually at the beginning of the school year). Also available online.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Online item bank (since 2007)</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Purpose</strong>: Provide pedagogical support.</td>
<td></td>
</tr>
<tr>
<td><strong>Audience</strong>: Teachers.</td>
<td></td>
</tr>
<tr>
<td><strong>Content</strong>: Offers released test questions from all subject areas and target grades. Includes questions from both the national and international assessments.</td>
<td></td>
</tr>
<tr>
<td><strong>Others</strong>: Teachers can search test questions based on subject area, school cycle, and questions format (multiple choice or open-ended).</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Press kit (since 2006)</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Purpose</strong>: Inform policy.</td>
<td></td>
</tr>
<tr>
<td><strong>Audience</strong>: Journalists and regional offices of education.</td>
<td></td>
</tr>
<tr>
<td><strong>Content</strong>: PowerPoint presentation with main results.</td>
<td></td>
</tr>
<tr>
<td><strong>Others</strong>: Distributed to journalists during the press release or before with embargo.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Data files (since 2005)</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Purpose</strong>: Inform policy, provide pedagogical support, and hold schools accountable, depending on research topic.</td>
<td></td>
</tr>
<tr>
<td><strong>Audience</strong>: Researchers.</td>
<td></td>
</tr>
<tr>
<td><strong>Content</strong>: Data files with school-level results.</td>
<td></td>
</tr>
<tr>
<td><strong>Others</strong>: Data files with student-level results are provided upon request after justifying the research project and committing not to use the results to identify students or teachers.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Data analysis tool (since 2007)</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Purpose</strong>: Inform policy, provide pedagogical support, and hold schools accountable, depending on the type of analysis.</td>
<td></td>
</tr>
<tr>
<td><strong>Audience</strong>: Researchers and decision makers.</td>
<td></td>
</tr>
<tr>
<td><strong>Content</strong>: Computes mean scores, differences in mean scores, and percent of students reaching different performance standards. Computes results for different years, grades, public and private schools, subject areas, gender, and socioeconomic level, among others.</td>
<td></td>
</tr>
<tr>
<td><strong>Others</strong>: Data file with student-level results are provided upon request after justifying the research project and committing not to use the results to identify students or teachers.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Geo-referential system (since 2010)</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Purpose</strong>: Hold schools accountable.</td>
<td></td>
</tr>
<tr>
<td><strong>Audience</strong>: Parents.</td>
<td></td>
</tr>
<tr>
<td><strong>Content</strong>: Google maps with the geographical location of schools and their mean scores (figure 5).</td>
<td></td>
</tr>
<tr>
<td><strong>Others</strong>: Parents can compare the scores of schools that are in the same geographic area.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Website <a href="http://www.simce.cl">www.simce.cl</a> (since 2001)</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Purpose</strong>: Inform policy, provide pedagogical support, and hold schools accountable.</td>
<td></td>
</tr>
<tr>
<td><strong>Audience</strong>: General public.</td>
<td></td>
</tr>
<tr>
<td><strong>Content</strong>: All the mechanisms described above. Also purposes of the assessment, organizational structure of SIMCE, technical documents, and references to publications using SIMCE data.</td>
<td></td>
</tr>
</tbody>
</table>

**Source**: SIMCE.
Provide Pedagogical Support

The assessment information is also used to provide pedagogical support to the schools. However, this use has been second in importance compared to informing policy. Two main factors that have contributed to this are a weak assessment culture in the schools, and the publication of assessment results that have little pedagogical meaning for teachers.

Several mechanisms allow using the assessment to provide pedagogical support to the schools. The two most important ones are the assessment guidelines and the school report. Complimentary mechanisms are the online item bank, the assessment data files, the data analysis tool, and the website (table 1).

The school report (figure 2) is distributed to all participating schools, and includes detailed information about the school results compared to the previous year, the national mean, and the mean of schools that serve students from a similar socioeconomic background. The report provides examples of the test questions, distribution of students’ responses in multiple-choice questions, examples of responses to open-ended questions, and scoring criteria. The test questions and student responses are analyzed in terms of contents and skills. Since 2000, results indicate the percentage of students reaching different performance levels; this was an important step to provide pedagogical meaning to the results.

However, there are weak institutional conditions and incentives for educators to use SIMCE results at the school level. The Ministry of Education has addressed this weakness with seminars and workshops for the schools to analyze their results. Guidelines for the workshops are provided in the school report. From 2004 to 2008, the Ministry of Education mandated that all schools implement a full-day “SIMCE workshop” to analyze their results and plan for

Figure 2. School Report with Results for Teachers and Principals

Source: SIMCE.
improvement. However, doing so implied cancelling classes for all the students. Now schools are expected to run the workshops without cancelling classes, during non-teaching hours. For many schools it is just not feasible to gather all its teachers to review the assessment. Another problem is that the schools are not accountable to an external body (for example, municipalities) regarding their improvement plans.

A new law approved in 2011 (yet to be implemented), should create better institutional conditions and incentives for educators to analyze and use SIMCE results (Government of Chile 2011). It is promising that this new law will require schools create improvement plans (based on SIMCE results among other indicators) and be accountable to an external body for their improvements. Providing feedback to the schools and teachers with useful assessment information does not seem to be enough to improve higher student learning. The right institutional arrangements and incentives should also be in place for the assessment information to have an effect.⁴

Moreover, SIMCE does not systematically inform teacher training institutions to ensure that future educators are “assessment literate.” In fact, pre-service and in-service teacher training programs barely cover assessment topics. As a consequence, educators know little about how to analyze student responses and how to use this information to improve pedagogy.

Hold Schools Accountable

A third important use of assessment information is holding schools accountable for student performance. This use of assessment is reached by a combination of strategies aimed to inform the general public about the school results and to distribute incentives (economic and symbolic) to the schools. These strategies are: (i) the publication of school results, (ii) the dissemination of information to the parents, and (iii) the accountability programs. Each of these strategies is described in the following paragraphs.

Publication of schools results. The publication of school results has been mandated by law since 1990 (Government of Chile 1990, 2009, 2011). However, the law was not implemented until 1995, when school results were published for the first time in a newspaper supplement (since 2001, school results have also been available online). The publication includes the school mean scores, as well as differences between school means and school mean results from the previous assessment, from the national mean, and from schools from the same socioeconomic group.⁵ The publication of school results has a wide impact, with

---

⁴ Evidence supporting this statement can be found in an experiment comparing the impact of assessment information without and without incentives to the schools (see Muralidharan and Sundararaman 2010).

⁵ Schools are classified in five socioeconomic groups, depending on the student population they serve.
both positive and negative consequences. SIMCE results are on the front pages of the newspapers, and education is on top of the public discussion (figure 3). On the negative side, the information allows for the stigmatization of low-performing schools serving the poorest students. In fact, around 80 percent of the variance in school mean scores can be predicted by the socioeconomic status of the student population served by the schools (Mizala, Romaguera, and Urqueola 2007).

**Figure 3. School Results Published in a Chilean Newspaper**

![Image](image.png)

*Source: Diario La Tercera 2011: 11.*

**Dissemination of information to parents.** Informing parents plays a double function: holding schools accountable and parent involvement. SIMCE has two main mechanisms to inform parents: the parent report and a geo-referential system. The parent report informs about school results (the school mean, and differences between the school mean and the national results, the school results from the previous assessment, and from schools from the same socioeconomic group). The report also provides recommendations on how parents can support their children’s education at home (figure 4).
Figure 4. Parent Report

Source: SIMCE.
A low-cost and sustainable strategy was used to introduce the parent report. First, the SIMCE Commission built consensus around the importance of disseminating this report. Second, SIMCE used the cheapest mechanism, which was a black and white flyer that could be easily reproduced and distributed to schools. Third, SIMCE distributed a few printed reports to the schools, and asked them to photocopy and distribute the report to parents. In the following years, SIMCE gradually absorbed the costs of making the reports available to all parents. Parents value the information provided, but still have difficulties understanding and using it (Taut et al. 2009).

The geo-referential system (SIMCE’s Maps) is available online and displays Google maps with schools and their mean scores all over the country (figure 5). This system was launched in 2010, and triggered an immediate and strong debate in education. Initially, the geo-referential system displayed the school mean scores together with a green, yellow, or red light (stoplights), depending on if the school mean score was above, at, or below the national mean. After a strong debate that reached the congress, stoplights were eliminated to avoid the stigmatization of low-performing schools. It is not known to what extent parents are using the geo-referential system.

Figure 5. Geo-referential System with School Results

Source: SIMCE.
Accountability programs. Chile uses a combination of accountability programs to put pressure on and distribute incentive to the schools. The first of these programs, the National Performance Assessment System (Sistema Nacional de Evaluación de Desempeño—SNED), was passed into law in 1995. It provides monetary incentives to teachers from high-performing schools. The second program, the Preferential Subsidy (Subvención Escolar Preferencial—SEP), was passed into law in 2008. It provides economic incentives and pedagogical support to schools serving low-income students that reach agreed objectives. Finally, the Quality Assurance Law of 2011 (yet to be implemented) will set up a stronger regulatory framework, will monitor school results, and will support them with technical assistance. The law will also allow for closing schools that do not show improvements.

The specific mechanisms through which schools are hold accountable are the newspaper supplement, the parent report, the geo-referential system, the data files, and the website (table 1).

Negative Uses of the Assessment

Beyond the “official” uses of the assessment, there are unintended negative uses that are a concern for the assessment program. The SIMCE Commission warned against the following: teachers teaching the subject areas covered in the assessment only; teachers overusing multiple-choice questions in classroom assessment; schools focusing resources on the tested grades only; schools excluding low-performing students from the assessment; the stigmatization of low-performing schools; and the further segregation of the school system along academic performance and socioeconomic level (Comisión SIMCE 2003). It is not know to what extent these negative uses are affecting the school system in Chile.

Drivers for Change

Several contextual drivers allowed for the development of the SIMCE dissemination strategy. These included the political context, legal mandate, institutional conditions, leadership, and international influences.

The political context provided a strong support for developing the assessment. When education became a national priority in the 1990s, SIMCE provided a much-needed, credible indicator of education quality.

The legal mandate to publish school results led to important innovations. SIMCE started publishing school results in a newspaper supplement and on its website, which boosted the coverage of SIMCE results in the media. SIMCE has to classify schools in socioeconomic groups to make school comparisons fairer.

Important institutional conditions allowed for strengthening SIMCE’s dissemination strategy. First, SIMCE benefited from strong and stable technical and political leadership. This leadership was key to setting long-term goals.
Second, the yearly budget for the assessment increased from US$2.5 million in 2001 to US$23 million in 2011. This allowed for strengthening the assessment teams and, most notably, for creating in 2005 a unit fully devoted to the analysis and dissemination of assessment information. The provision of a stable budget was a key ingredient for the implementation of the current dissemination strategy. In 2010/11, around 10 percent of the SIMCE budget was spent on dissemination.

The leadership provided by a technical-political commission allowed planning for future development of the dissemination strategy. Criticism of the government was growing because the reform efforts of the 1990s did not translate into increased test scores. This raised concern about the capacity of SIMCE to detect changes in student performance. In 2003, the Minister of Education asked a technical-political commission to review the assessment program. The members of the commission included representatives of the teachers union, parents, school administrators, universities, think tanks, and the Ministry of Education. Over five months, the SIMCE Commission went through a strategic planning process. The commission backed the assessment, and made recommendations for its further development. The recommendations included (i) publishing the parent report, (ii) reporting performance standards, (iii) reporting results to school administrators, (iv) increasing the participation of teachers in the assessment, and (v) providing teachers with greater access to information. These recommendations became the roadmap for SIMCE for the coming years (Comisión SIMCE 2003).

Finally, international influences also shaped SIMCE’s dissemination strategy. In the decade of the 2000s, there was a push for greater transparency that translated into publishing more and better information about the assessment. A critical step was releasing the SIMCE databases. This allowed for an increase in the number of researchers and policy makers using the assessment to inform decision making and evaluate policies. Another international influence that shaped SIMCE’s dissemination strategy was the accountability movement. This boosted the development of incentive schemes based on SIMCE results.

Impact of the Assessment Program

Has the Chilean assessment program contributed to education quality through improved student learning? While this is the ultimate criteria with which the effectiveness of an assessment program should be judged, this is a very hard question to answer. Assessment programs operate in combination with other policies, and therefore it is not possible to disentangle their unique contributions to improved education. Their effectiveness is conditioned not only by their technical quality, but also by external factors such as teachers’ capacity to

---

6 Data based on Chile’s national budget (http://www.dipres.gob.cl).
analyze results, set learning goals, and teach their students in order to reach those goals.

Despite these limitations, there is some evidence that supports the contribution of SIMCE to improved student performance. Chile’s gains in student performance in the last decade are documented both by SIMCE and PISA. These gains seem to be, at least in part, the consequence of stable and effective policies. Since these policies have been directly informed by SIMCE, therefore it is possible to argue that SIMCE has had an indirect effect on student performance. This rationale is backed by the high value ministers of education and policy makers give to the SIMCE information. This rationale is also backed by the McKinsey report (2010), which showed that the most improved school systems (including Chile’s system) systematically use assessment information to inform policy.

Nevertheless, at the school level, SIMCE information seems to be underused. While educators value the SIMCE information, they have difficulties understanding and using it with pedagogical purposes (Taut et al. 2009). Nevertheless, schools report making important decisions based on the assessment results, including decisions such as implementing remedial courses, focusing instruction on topics covered by SIMCE, and assigning the best teachers to the “SIMCE grades.”

Regarding the use of SIMCE for accountability purposes, parents do not seem to use assessment information to “shop for schools.” While parents declare choosing schools based on quality, most do not use SIMCE results to inform their decisions. School demographics have more weight than SIMCE when choosing a school. This makes sense, considering the difficulties parents have had in accessing and understanding SIMCE results (Elacqua, Schneider, and Buckley 2006; Taut et al. 2009). It is not known to what extent SIMCE information promotes parent involvement in school.

So far, only the oldest accountability program, SNED, has been subjected to impact evaluation. Research shows that SNED had small but significant effects on student performance, ranging from 0.1 to 0.2 standard deviation, but that it has not affected demand-side indicators like increased student enrollment. However, SNED seems to be a very cost effective program when compared to other reforms (Cabezas, Cuesta, and Gallego 2011; Gallego 2008; Rau and Contreras 2008; Mizala and Urquiola 2007). No studies were found about the impact of SNED on teacher motivation and pedagogical practices.

**Lessons Learned**

Countries aiming to strengthen the uses of their national large-scale assessment program may extract some important lessons from this case study. Chile systematically uses assessment information to inform policies, to distribute
incentives to the schools, and to support pedagogy. The McKinsey report (2010) points to the systematic use of assessment information as a key feature of the most improved school systems, including Chile’s system.

Several drivers allowed for the systematic use of assessment information in Chile. Probably the most important drivers were the political pressure to improve education quality, and the need to have a valid, fair, and credible indicator of quality. The legal support of the assessment pushed for a strong dissemination strategy, including the publication of school results. The SIMCE Commission built consensus around the purposes and mechanisms of the assessments, and pushed for providing more and better information to teachers and parents. Finally, the injection of additional resources to the assessment office was a necessary condition to improve SIMCE’s dissemination strategy.

Countries aiming to improve their assessment systems can draw important lessons from Chile’s use of assessment information. These lessons are grouped into those that refer to the dissemination strategy and those that refer to the conditions for using the assessment information.

Lessons Regarding the Dissemination Strategy

**Having a clear vision.** Countries need to have a clear vision about assessment. This requires agreeing upon the purposes of the assessment, the target audiences to be reached, the messages to be delivered and the uses to be promoted. Strong technical and political leadership need to converge for this vision to take shape. In Chile, this vision came first from the leadership of key officials from the Ministry of Education. Later, it came from the consensus reached by the SIMCE Commission. The report elaborated by this commission became the roadmap for the strategic development of the assessment program.

**Building consensus.** This is important to ensure the political viability of the assessment program. Stakeholders need to agree on a vision for the assessment. That is, they need to reach consensus regarding the purposes, audiences, and uses of the assessment. In Chile, this consensus was reached through a high-level technical-political commission whose members included all stakeholders. Thanks to a strategic planning process, this commission and its report allowed for strengthening the legitimacy of the assessment, and for identifying and prioritizing strategic goals. The commission agreed on the importance of strengthening the dissemination of information to parents and school administrators.

**Strengthening the dissemination strategy.** The dissemination strategy should be the “leading horse” that orients all the assessment activities. The starting question should always be “What information do we want to put out there?” Based on the response to this question, decisions have to be made regarding what to measure, who to measure, and how to measure. All the assessment activities should be planned in order to provide the required information to be disseminated, no more and no less. The dissemination strategy should clearly
identify the purposes and intended use of the assessment information, the target audience, and the dissemination mechanism.

Countries usually make the mistake of collecting much more data than they are able to communicate. Chile was not an exception. For years, SIMCE collected background information through questionnaires, but this information was not published. This was expensive and inefficient. Another common mistake is that countries run out of money before publishing the results. Thanks to good planning, this was never the case in Chile.

Implementing gradual changes. Countries should not attempt to develop in two years a dissemination strategy that took other countries decades to develop. Implementing a comprehensive dissemination strategy takes years, even decades. The dissemination strategy should be implemented gradually, should move from simple to complex, and should be built so that the first steps serve as a foundation for the following ones. Gradualism should go hand-in-hand with the capacity of the assessment office to implement changes. It should also go hand-in-hand with the country capacity to “digest” assessment information.

Whereas nowadays Chile has a comprehensive dissemination strategy, it is important to note that at the beginning the strategy only included a school report and assessment guidelines for teachers. New dissemination mechanisms had been introduced gradually. These changes responded to new political priorities (for example, involving parents), and changes in technology (such as the Internet), among others. They also responded to changes in assessment techniques. For instance, in the 1980s SIMCE reported student results as the percentage of correct responses only; in the late 1990s, it started reporting comparable scores using an Item Response Theory (IRT) scale; and in the 2000s it added the percentage of students that reached different performance standards to its reports.

Ensuring sustainable changes. Countries should be cautioned against launching reports that are not sustainable in time, and against experimenting with different types of reports in each assessment cycle. These actions create frustration and confusion. Once a report is published, stakeholders expect it to come on a regular base. Before introducing a modification in the dissemination strategy, the assessment office should make sure that the changes are sustainable in time. Many times, this means looking for the simplest and most cost-effective solution. In Chile, a good example of a sustainable change was the introduction of SIMCE’s parent report. This is a one-page black-and-white flyer that can easily be photocopied.

Moving from lower to higher stakes. When a new assessment program is introduced, it is important to ensure its legitimacy. This requires that stakeholders, especially educators, perceive the assessment results as valid, fair, and useful for improving education quality. If educators perceive the results as a threat, or if they do not trust them, then the assessment will not contribute to improving quality. Attaching consequences to an assessment program when it is
first launched is not good for building trust. Stakeholders need first to know the program in a “protected setting,” without the pressure of incentives. Once the program is recognized as an important piece of the school system, then it is a more appropriate time for introducing incentives. When introduced, incentives should be balanced with other uses of information (for example, an assessment component without consequences for the school).

In Chile, SIMCE gradually attached higher stakes to the assessment results. In the 1980s, results were distributed to the schools only, without consequences. In the mid-1990s, school results were published, and in the late 1990s and 2000s, accountability programs with higher stakes for teachers and schools were introduced.

*Introducing comprehensive accountability programs.* School and teacher quality are complex concepts and no single indicator can fully capture this complexity. If an accountability program were to rely on student assessment results only, other important aspects of quality would not be considered when distributing incentives. This would contribute to a perception of unfairness and lack of legitimacy of the accountability program. Therefore, countries aiming to distribute incentives to schools should avoid relying on a single indicator (such as mean test scores) for doing so. Multiple indicators that cover different facets of quality are preferable (for example, dropout rates, gain scores, and meeting yearly goals).

Different accountability programs have been gradually being introduced in Chile. In all of them, student scores are the main indicator. Other indicators include school repetition rates, school policies, and administration.

**Lessons Regarding the Conditions for Using the Assessment Information**

*Building an assessment culture.* This means having a workforce that is assessment literate, with educators that can value, understand, and use assessment information to improve education. If educators have a prejudice against the assessment, they will ignore it. If they have difficulties understanding assessment results, they will never use the information to improve pedagogy.

The assessment program has to make all possible efforts to provide educators with opportunities to learn about the assessment. This can be done in several ways. First, expose educators to several assessment experiences. Second, involve educators in the assessment, for instance by providing teachers with tests to administer to their own students, and scoring criteria and guidelines with which to analyze and use assessment results. Finally, build bridges between the assessment program and pre- and in-service teacher training programs.

After more than two decades of assessments results, Chilean schools still have a poor assessment culture and too many educators are not assessment literate. This is in part the consequence of teachers having little opportunity to learn about SIMCE. This is a weakness that other countries aiming to develop their assessment program should try to avoid.
Creating the right institutional conditions. Disseminating assessment information is a necessary but not sufficient condition for the information to be used. The right institutional conditions need to be in place for the information to be used. Chile has been gradually introducing institutional conditions for the schools to use the assessment results. For example, schools organize meetings with parents to distribute and discuss the parent reports, and they also organize workshops to analyze SIMCE results. In addition, schools receive monetary incentives based on their assessment results. Countries aiming to ensure the effective use of assessment information should plan for these or other conditions.

Ensuring that assessment results are valued and are perceived as fair, valid, and useful. The ultimate goal of the assessment is to contribute to improve education quality. For this to happen, stakeholders should value the assessment, and perceive the data it provides as fair, valid, and useful. Results are more likely to be perceived as fair if the socio-economic status of the students served by the schools is recognized. Results are more likely to be perceived as valid if teachers can see links between the curriculum and the assessment. Results are more likely to be perceived as useful if teachers can understand them, can analyze student responses, and can plan pedagogy accordingly.

In Chile, fairness in results was threatened in 2010 when a stoplight system was introduced to report school results. Stoplights were strongly criticized because they did not recognize the socio-economic status of the students served by the schools. This was perceived as unfair because socio-economic status was a strong predictor of the school results (for instance, poorer schools got red lights, richer schools got green lights). After a strong debate, SIMCE dropped the stoplights.

To contribute to the validity of the assessment program, SIMCE has been publishing assessment guidelines since the 1980s. These guidelines present the assessment framework of the tests, and explain the links between the framework and the curriculum.

Regarding the utility of the assessment, this is an area that still needs to be strengthened in Chile. While teachers value the SIMCE information, they have difficulties understanding and using it. To address this issue, SIMCE introduced workshops to train teachers in the analysis of assessment results.

Monitoring the consequences of the assessment. It is not enough just to disseminate information; the assessment program needs to make sure that information is used in the right way, while avoiding negative uses. Yearly interviews and surveys are important for determining that stakeholders are getting the assessment information, they value it, and they understand and use it. SIMCE has not yet systematically incorporated this feature into its design.

Countries aiming to improve the dissemination, use, and impact of their assessment programs should critically review these lessons. Their utility will depend on the countries’ capacity to select and adapt these lessons according to their cultural background, assessment development, and resources, among
others. Ultimately, the contribution an assessment program can make to improving education quality will depend on the effectiveness of its dissemination strategy.
Acknowledgments

References


Diario La Tercera. 2011. “School Results.” April 7, p. 11.


———. 2011 Quality Assurance Law (Sistema Nacional de Aseguramiento de la Calidad de la Educación Parvularia, Básica y Media y su fiscalización). Santiago, Chile.


**Websites**

The Central Bank of Chile (www.bcentral.cl)
Ministry of Social Development, Chile (www.ministeriodesarrollosocial.gob.cl)
PISA program (http://www.pisa.oecd.org)
SIMCE, Chile’s national assessment program (www.simce.cl)
World Bank (www.worldbank.org)
Student assessment plays a key role in improving education. However, many countries do not disseminate and make effective use of their student assessment information. This paper discusses Chile’s effective use of its assessment information and draws lessons for other countries. A key lesson from the Chilean experience is that effective use of assessment information is a challenging task that takes time, effort, and resources. Another lesson is that the task requires leadership, long-term planning, and consensus building. Countries able to learn from Chile’s lessons will be in a better position to make effective use of their student assessments and, ultimately, to contribute to education quality and learning.

María-José Ramírez, Education Specialist, Human Development Network, the World Bank

The Russia Education Aid for Development Trust Fund is a collaboration between the Russian Federation and the World Bank that supports the improvement of student learning outcomes in low-income countries through the development of robust student assessment systems. Visit the READ website at www.worldbank.org/readtf for additional information.