PEL Pensamiento Educativo. Revista de Investigación Educacional Latinoamericana
2013, 50(2), 28-50

Pontificia Universidad Católica de Chile
Facultad de Educación

Adding or not Adding Contextual Value in Language: An Instrumental Case Study in Two Chilean Secondary Schools

Agregar o no agregar valor contextual en lenguaje: un estudio de caso instrumental en dos establecimientos chilenos de educación media

Bernardita Muñoz-Chereau

University of Bristol, UK

Abstract

This article reports a qualitative case study (Stake, 1995) carried out in two Chilean secondary schools, oriented to examine the school processes and/or dimensions that might explain observed differences in schools' Contextual Value Added (CVA) scores in language. After analysing a longitudinal sample of a cohort of 176,896 students nested within 2,283 schools that took SIMCE tests in grades 8/10 in 2004-2006 and their family questionnaires through the statistical technique of multilevel modelling (Goldstein, 1995), schools' CVA scores were used to choose one school with low and one with high performance. Complementing an analytical framework drawn from the literature (Scheerens, Glas, & Thomas, 2003), the study identified 11 dimensions that appeared to differentiate, as well as 5 dimensions that appeared similarly problematic in both schools. These dimensions were broadly in line -but also qualify and extended- the findings of a seminal Chilean Educational Effectiveness Research (EER) study (Belleï, Muñoz, Pérez, & Raczynski, 2004). This new evidence, although tentative and exploratory, locates the features of more and less effective Chilean secondary schools within the broader literature review on EER. Finally, the study provides timely contributions to the new assessment framework in place in Chile with the potential to inform and enhance educational policy and practice.

Keywords: value added, contextual value added, educational effectiveness research, Instrumental case study, Chilean secondary education

Post to:
Bernardita Muñoz-Chereau
University of Bristol, UK
8 Saint Michael's Hill, Kingsdown, Bristol, BS28DT, UK
Email: berniviajera@gmail.com
This research was financed by CONICYT, Programa Formación de Capital Humano Avanzado.

ISSN: 0719-0409 DDI: 203.262, Santiago, Chile
doi:10.7764/PEL.50.2.2013.3

Resumen

Este artículo reporta un estudio de caso cualitativo (Stake, 1995) en dos establecimientos secundarios chilenos orientado a examinar los procesos y/o dimensiones que pueden explicar las diferencias observadas en las puntuaciones de Valor Agregado Contextual (VAC) en el sector de lenguaje, ámbito poco estudiado en Chile y más ligado a la eficacia general que las matemáticas (Sammons, Thomas & Mortimore, 1997). Tras el análisis estadístico multinivel (Goldstein, 1995) de una muestra longitudinal de una cohorte de 176.896 estudiantes anidados en 2.283 establecimientos que rindieron el SIMCE en 8º básico a 2º medio en 2004-2006 y sus respectivos cuestionarios familiares, se utilizaron las puntuaciones del VAC para seleccionar un establecimiento que agregó valor y otro que no. Complementando un marco analítico-conceptual preexistente (Scheerens, Glas & Thomas, 2003), el estudio identificó 11 dimensiones diferentes y 5 similarmente problemáticas en los establecimientos. Estas dimensiones complementan y extienden las conclusiones de un estudio seminal de escuelas efectivas en Chile (Bellei Muñoz, Pérez & Raczynski, 2004). Esta nueva evidencia, tentativa y exploratoria, inscribe las características de los establecimientos secundarios chilenos dentro de la literatura más amplia sobre efectividad educacional, buscando contribuir al nuevo marco de evaluación vigente en Chile con el potencial de informar y mejorar las políticas y prácticas educativas.

Palabras clave: valor agregado, valor agregado contextual, efectividad educacional, estudio de caso instrumental, educación secundaria chilena

In the last decade Chile has showed outstanding progress by improving 40 points in PISA, positioning the country at the top of the reading performance ranking among Latin-American countries (OECD, 2011b). However, most of the previous EER studies carried out internationally, and particularly in the Chilean context (Carrasco & San Martín, 2012; Manzi, San Martín, & Van Bellegem, 2013; Ramirez, 2007; Thieme, Tortosa-Ausina, Prior, & Gempp, 2012), have focused exclusively on Mathematics. Perhaps this has been the case because it is well known that family background and home environment play a more significant role in determining attainment in Literacy than in Mathematics or Science, imposing an extra challenge when modelling data concerning language (Steele, Vignoles, & Jenkin, 2007). Given that previous research conducted at the secondary level has stressed the need to analyse the impact of the department explicitly as its size and extent tend to differ from the school effect (Ainley, 1994; Harris, Jamieson & Russ, 1995; Luyten, 1998; Sammons, Thomas, & Mortimore, 1997; Smyth, 1999) and has suggested that school effects in language are more closely related to overall effectiveness than Mathematics (Sammons et al., 1997), exploring school effectiveness in language in Chile clearly merits further research.

Literature review

Measures of SESI: Raw, VA and CVA

Different EER measures for judging school effectiveness based on student outcomes (Raw, VA and CVA) will be briefly reviewed in order to describe the approach taken by this study.

Raw attainment scores. In the international scenario there is considerable literature from western countries in general (Hoyle & Robinson, 2003; Ladd & Zelli, 2002; OECD, 2008, 2011a; Raudenbush & Willms 1995; Rosenkvist, 2010; Scherrer, 2011; Willms & Raudenbush, 1989) and the UK in particular (De Luca, as cited in Lambert & Lines, 2000; Goldstein et al., 1993; Leckie, 2008; Leckie & Goldstein, 2007; Sammons, 1999; Sammons et al., 1997; Slee, Weiner, & Tomlinson, 1998; Stoll & Mortimore, 1997; Thomas, Salim, Muñoz-Chereau, & Peng, 2011) supporting the fact that raw or unadjusted measures of pupil achievement in a given test provide a poor method for comparing schools' performance. Particularly in Chile, the decision to judge school performance based on their raw test results would be a way of rewarding social segregation (Treviño & Donoso, 2010). Even though concluding from raw results that schools are effective or ineffective would be misguided, the conclusion that can be derived is how big the achievement gap between schools is in a given educational system. The necessity to put raw results into context has supported the development of better measures for judging school performance. The claim that something more sophisticated than "league tables" is needed in order

to compare schools on a more equitable basis has been strongly associated with VA and CVA approaches (Schagen & Hutchison, 2003), because they offer the promise of a more rigorous approach for levelling the playing field (Braun, Chudowsky, & Koenig, 2010; Scherrer, 2011).

Value Added (VA). Within a bi-dimensional definition of SER/EER concerned with educational quality and quantity, Creemers (1996) defined VA as a way of measuring the quality of the school by considering the average score in a given measurement after correcting for input characteristics. The so-called VA approaches (Aitkin & Longford, 1986; Gray, Reynolds, Fitz-Gibbon, & Jesson, 1996) try to identify the individual school's contribution to students' relative progress over time (Creemers, 1996; Gray, Goldstein, & Thomas, 2003). This is the reason why the availability of baseline prior attainment as well as background factors is crucial. Under this approach, "the schools could see to what extent they have boosted pupils' progress" (Stoll & Mortimore, 1997, p. 16). A school that adds value has been defined as one in which pupils progress further than might be expected from consideration of the intake (Mortimore, 1991; Sammons, Cuttance, Nuttall, & Thomas, 1995). Other authors have taken a wider approach to VA approaches by defining them as necessary to put the raw school examination results in context (Thomas & Mortimore, 1996), while others have defined them as "methods that account for the differences in student intake among schools" (Timmermans, Doolaard, & De Wolf, 2011, p. 393). Teddlie and Reynolds (2000) defined VA as «a measure of the relative gain in achievement made by pupils. The rationale is that a school is not responsible for the absolute level of student achievement so much as for the progress made by pupils in its care» (p. 264).

Contextualized Value Added (CVA). It has been stressed that CVA is a way of estimating the effect of school policies upon student achievement by explicitly controlling for prior achievement and contextual factors (Ballou, Sanders, & Wright, 2004). So the difference between VA and CVA is that the latter is an extension of VA that explicitly controls for intake differences between schools in pupils' academic and background characteristics, as well as context factors (Leckie & Goldstein, 2011). Precisely because it takes into account contextual factors —and, critically, peer group effects that relate, for example, to the combined influence of SES over and above the individual effect— CVA is considered a methodologically more advanced or refined version of VA (Rosenkvist, 2010; Thomas & Mortimore, 1996; Timmermans et al., 2011). For this reason it is claimed that CVA models are better for identifying those schools with a greater proportion of students from disadvantaged backgrounds that are able to boost student performance (OECD, 2008).

Even though theoretically such an analysis would not be possible with VA, given the evidence that prior attainment is already implicitly reflecting difference in intake, it is sound to expect strong correlations between VA and CVA estimates. As many studies have shown, the inclusion of individual student background/socio-economic characteristics has helped fine-tune VA measures (Sammons et al., 1997; Sammons, Thomas, Mortimore, Owen, & Pennell, 1993; Thomas & Mortimore, 1996). In any case, for a system of external school accountability, it would be more accurate to employ CVA as the main indicator of school performance. For example, when schools that serve large concentrations of disadvantaged students are considered —like public schools in Chile— if they do not have sufficient compensatory resources to offset the educational challenges that such students pose, they may wrongly appear as ineffective despite using their insufficient resources more productively and efficiently than other schools (Ladd & Walsh, 2002). These are the theoretical justifications why the qualitative study reported here is based on a CVA model.

Processes and/or Dimensions of Effective and Ineffective Schools

Effectiveness and ineffectiveness are opposed, relative concepts that can be defined from many approaches (Stoll & Myers, 1998). This study works with a particular definition of effectiveness: students' progress further than might be expected from a consideration of their intake (Mortimore, 1991), which locates this study not just within the Educational Effectiveness and School Improvement field, but more precisely within VA/CVA approaches. This study will explore processes, defined as aspects of the school environment associated with student achievement (D'Haenens, Van Damme, & Onghena, 2010), or school organizational and instructional variables resulting from an effort to open the "black box" of the school (Scheerens, 2000). Describing how well those processes help explain differences in school effects is central to EER in the international scenario, as well as in the context of Chile, where little research has been conducted.

In the last 30 years, a number of researchers have been devoted to trying to identify common features concerning the processes and characteristics of more and less effective schools in different contexts. Studies using different methodological approaches -from statistical analyses to case studies of outlier schools and a combination of other methods- have identified key characteristics of effective and ineffective schools, working as underlying explanatory mechanisms that have helped to illuminate understandings of what seems to make the difference at the school level. However, focusing on the SESI academic field developed mainly in the last 40 years, what makes an effective school and how schools improve still seems to be an elusive issue. Even though the question about the crucial factors in education that promote students' academic achievement has been debated for several decades, a consensual answer has not been produced. Country specificity (Opdenakker & Van Damme, 2000) and the presence of highly contested issues such as Differentiation (Duri-Bellat & Mengat, 1998; Klieme, 2012) have supported the claim that research in these disciplines could not elucidate clear recipes or guarantee a path for improving the performance of schools.

Although there are many lessons that can be learned from the efforts of the studies implemented in industrialised countries, it is also necessary to acknowledge the tension between the generalizability and the context specificity of the EER knowledge-base. It is not enough to be aware of how effective and ineffective schools have been operationalised in different countries, nor to choose one approach and replicate it in Chile. It is also necessary to challenge the generalizability of these approaches in order to search for dimensions and/or characteristics associated with students' outcomes that are particularly relevant in Chile. An overemphasis on "what works" may result in missing important context-specificity, whereas an exaggeration of context may avoid generalizations. A more realistic and humble aim for EER is to increase the understanding of school and classroom processes, and the way inputs and context seem to influence students' educational outcomes. That is also the aim of this study.

Processes and/or dimensions of effective schools. Given the overwhelming amount of literature devoted to describing different process factors linked with effective schools, reviewing this work in detail is beyond the scope of this article and has been done elsewhere (Hattie, 2008). Moreover, the fact that there seem to be no generally agreed theory of educational effectiveness, but rather common features of effective schools and effective teaching in a range of countries (Klieme, 2012; Sammons, 2007), means that a complete study could be conducted reviewing differences and similarities among different approaches, especially because consensus hides divergent conceptualisations (Scheerens, 2000).

In reference to Chile, the early qualitative case study carried by Belleï, Muñoz, Pérez and Raczynski (2004), with 14 (10 public and 4 subsidized private) primary schools located in economically disadvantaged areas, has been singled out as one of the most prominent EER studies in the region (Murillo, 2007). The study identified more than 30 dimensions of which the most noticeable ones seem to be 5 key effective dimensions: Pedagogic-centred leadership, Positive culture, High academic expectations, Home-school partnership, and Good work in the classroom (Belleï et al., 2004). However, the study has recognised methodological limitations (Carrasco, 2010), including, for example, focusing only on effective schools and using raw attainment scores to select the sample. Given that raw attainment scores have been consistently criticized for judging the effectiveness of a school, as described above, the question of whether the schools were effective or merely selective remained unanswered.

Nevertheless, as a brief summary Table 1 compares the effectiveness-enhancing processes/dimensions of schooling in 10 studies conducted in different countries.

Table 1
Effectiveness-enhancing processes/dimensions of schooling in different studies

14 effectiveness-enhancing factors (Scheerens, Glas, & Thomas, 2003)	10 features of school success (National Commission on Education, Mortimore, in Maden, 2001)	5 key effective dimensions (Belleï, Raczynski, Muñoz, & Pérez, 2004)	9 processes of effective schools (Teddlie & Reynolds, 2000)	9 processes factors (Sammons, Thomas, & Mortimore, 1997)	8 general effectiveness factors (Scheerens & Bosker, 1997)	11 factors of effective schools (Sammons et al., 1995)	4 key characteristics of secondary schools (Smith & Tomlinson, 1990)	5-factor model of (Edmonds, 1979)	8 school processes of effective schools (Rutter et al., 1979)
Educational leadership	Strong positive leadership by the head and senior staff	Pedagogic-centered leadership	Effective leadership Effective teaching	Clear leadership of HT, HoD Effective SMT	Professional leadership	Shared leadership	Effective leadership and management by senior and middle managers	Strong educational leadership	Teachers as positive role models
School climate as orderly atmospheres/ School climate in terms of effectiveness orientation and good internal relationships	A good atmosphere or spirit, generated by shared aims and values; and by a attractive and stimulating physical environment	Positive culture	Producing a positive school culture emphasizing responsibilities and rights	Strong academic emphasis	Productive climate and culture	Pupil rights and responsibilities	Climate of respect between all participants	Orderly and secure environment	School ethos
High expectations/ Achievement orientation	High and consistent expectations of all pupils	High academic expectations	Creating high and appropriate expectations Developing and maintaining a pervasive focus on learning	High expectations	High expectations	High expectations		High expectations of pupil attainment Emphasis on the acquiring of basic skills	High teacher expectations
Parental involvement	Parental involvement in children's education and in supporting the aims of the school.	Home-school partnership	Involving parents in productive and appropriate ways	Parental support/involvement	Parental involvement	Home-school partnership (parental involvement)			
Evaluative potential	Well-developed procedures for assessing how pupils are progressing		Monitoring progress at all levels		Appropriate monitoring	Monitoring progress		Frequent assessment of pupil progress	
Consensus and cohesion among staff				Shared vision/goals Consistency in approach		Shared vision and goals	Teacher involvement in decision-making		Shared staff-student activities
Reinforcement and feedback	Rewards and incentives to encourage pupils to succeed					Positive reinforcement	Positive feedback and treatment of students		Positive feedback and treatment of students
Structured instruction	Clear and continuing focus on teaching and learning			Quality of teaching		Purposeful teaching			Effective classroom management
Independent learning	Responsibility for learning shared by the pupils themselves			Student-centered approach					Students given responsibility
Curriculum quality/opportunity to learn				Focus on central learning skills		Learning environment concentration on teaching and learning			
Effective learning time Differentiation Classroom climate		Good work at the classroom			Effective instructional arrangements	Learning organization			
Factors not included in this model:	Participation by pupils in the life of the school Extra-curricular activities which broaden pupils' interests and experiences, expand their opportunities to succeed and help to build good relationships with the school				Practice-oriented staff development				Good working conditions for staff and students

After reviewing 10 studies oriented to outline processes of effective schools, it is clear that school processes or factors are not a unitary body of knowledge. They seem rather complex, multidimensional, and difficult to measure. However, at least half of the studies included in the current review identified (a) Educational leadership, (b) School climate, (c) High expectations/Achievement orientation, (d) Parental involvement, (e) Evaluative potential, (f) Consensus and cohesion among staff, (g) Reinforcement and feedback, and (h) Structured instruction as key processes present in effective schools.

Processes and/or dimensions of ineffective schools. Although it has been stressed in the literature that ineffective schools are not merely lacking the features of those that are more effective, there is no doubt that ineffective schools have been comparatively less researched nationally and internationally (Barber & Dann, as cited in Maden, 2001; Stoll & Myers, 1998). Four studies in this area are those reported by Sammons et al. (1995, as cited in Thomas & Mortimore, 1996), Stringfield (1998), Van de Grift & Houtveen (2007), and Sammons (2007).

Sammons et al. (as cited in Thomas & Mortimore, 1996) described the relationship between negative VA results and the quality of teaching and learning. The authors highlighted disrupted teaching, low expectations, lack of appropriate feedback, inappropriate grading of school/homework, low departmental morale and lack of examination preparation.

Stringfield tried to clarify what he called the "anatomy of ineffectiveness" when describing common characteristics found in the Louisiana School Effectiveness Study (LSES) implemented in the United States in the 1990s (Teddlie & Stringfield, as cited in Stringfield, 1998). They carried 16 case studies in 8 pairs of demographically matched outlier schools claimed to serve identical communities. Although this study shares a similar methodological limitation with Belleï et al. (2004), they focused on Year 3 classrooms (9 years old) and reported qualitative differentiations at 3 levels. At the school level they highlighted seven dimensions: lack of academic focus, academic periods starting late and ending early, resources often worked at cross-purposes, bureaucratic leadership, head teachers passive in teacher recruitment, lack of teacher assessment and lack of public rewards for students' academic excellence. At the classroom level they found leisurely pace, minimal planning, preponderance of unengaging tasks, low rates of interactive teaching, teachers working in isolation, parts of the mandated curriculum not covered in teaching (less opportunity to learn), lack of any sense of academic push and finally, at the student level, low or uneven time on tasks and classes experienced as "intellectual anarchy" (lack of structure and sense making).

The study conducted by Van de Grift and Houtveen (2007) in the 284 underperforming elementary schools in the Netherlands reported six prominent weaknesses: (a) insufficient learning material offered at school to achieve core targets, (b) insufficient time allowed to achieving the minimum objectives of the curriculum, (c) poor instructional quality, (d) insufficient insight into students' performance levels, (e) insufficient or inappropriate special measures for struggling learners, and (f) dysfunctional organisation of the school prolonged over time.

Sammons (2007) also reviewed studies concerning the characteristics of ineffective schools and highlighted four main aspects: lack of vision, unfocused leadership, dysfunctional staff relationships, and ineffective classroom practices.

This brief review supports the fact that when specific features or issues that ineffective schools are claimed to share have been outlined, they inevitably compose a "deficit discourse". Given this misguided emphasis -in the sense that it is not clear that being aware of what is missing is enough to rectify the situation- is so predominant in the literature concerning ineffectiveness, and that some EER authors have claimed that these theories are capable of explaining effectiveness as well as ineffectiveness (Scheerens, 2012), this study decided to explore the same dimensions identified in the effective school when studying the less effective one. Special attention was paid to describe how each factor was operating, not just highlighting its deficits.

14 Effectiveness-enhancing theoretical and empirical SE factors. Even though only 8 school processes or dimensions were regularly identified on this literature review, the 14 effectiveness-enhancing theoretical and empirical SE factors identified by Scheerens, Glas and Thomas (2003) were chosen as an analytical framework to inform the design of the study, because, apart from being derived from a meta-analysis of 12 different EER projects carried out in western countries, they included dimensions not just

associated with leadership and management (i.e. educational leadership, consensus and cohesion among staff), but crucially dimensions related to the teaching and learning process that takes place mainly at the classroom level, such as classroom climate, effective learning time and independent learning. In this way they were wide enough to explore different aspects of the school processes and practices associated with school performance.

Table 2
14 Effectiveness-enhancing theoretical and empirical SE factors and its components

	Factors	Components
1	Achievement orientation, high expectations: clear focus on the mastering of basic subjects	-High expectations (School level) -High expectations (Teacher level) -Records on pupils' achievement
2	Educational leadership: indirect control and influence on the school primary process	-General leadership skills -School leader as information provider -Orchestrator of participative decision making -School leader as coordinator -Meta-controller as classroom processes -Time educational/administrative leadership -Counselor and quality controller of classroom teachers -Initiator and facilitator of staff professionalization
3	Consensus and cohesion among staff: Coherence, consistency and continuity among the staff	-Types and frequency of meetings and consultations -Contents of cooperation -Satisfaction about cooperation -Importance attributed to cooperation -Indicators of successful cooperation
4	Curriculum quality/opportunity to learn: The degree in which the implemented curriculum matches the achievement curriculum	-The way curricular priorities are set -Choice of methods and textbooks -Application of methods and textbooks -Opportunity to learn -Satisfaction with the curriculum
5	School climate as orderly atmospheres: School culture oriented to maintain orderliness	-The importance given to an orderly climate -Rules and regulations -Punishments and rewording -Absenteeism and drop out -Good conduct and behavior of pupils -Satisfaction with orderly school climate
6	School climate in terms of effectiveness orientation and good internal relationships:	-Priorities in an effectives-enhancing school climate -Perceptions on effectiveness-enhancing conditions -Relationships between pupils -Relationships between teachers
7	Evaluative potential: Aspirations and possibilities of a school to use evaluation as a basis for learning and feedback	-Evaluation emphasis -Monitoring pupils' progress -Use of pupil monitoring systems -School process evaluation -Use of evaluation results -Keeping records on pupils' performance -Satisfaction with evaluation activities
8	Parental involvement: Relationship between the actual degree of involvement of parents and the effort displayed by the school to engage them.	-Emphasis on parental involvement in school policy -Contact with parents -Satisfaction with parental involvement

9	Classroom climate: Good, ordered relationships, work attitude and satisfaction with the classroom interactions.	-Relationships within the classroom -Order -Work attitude -Satisfaction
10	Effective learning time: Importance and implementation of effective time on task at the class and school level	-Importance of effective learning -Time -Monitoring of absenteeism -Time at school -Time at classroom level -Classroom management -Homework
11	Structured instruction: Well organized and closely monitored lessons	-Importance of structured instruction -Structure of lessons -Preparation of lessons -Direct instruction -Monitoring
12	Independent Learning: Extent to which pupils are responsible and make decisions concerning their learning according to their own interests	- No sub-components
13	Differentiation: How the school deals with differences between pupils and takes care of pupils with learning and behavioral problems	-General orientation -Special attention for pupil at risk
14	Reinforcement and feedback: Basic requirement for learning that deals with the rapport to pupils in connection with their achievement	-No sub-components

Note: Adapted from Scheerens, Glas, & Thomas, 2003.

Methods

Research Questions (RQ)

Two RQ are addressed in this study: (a) Using an analytical framework drawn from the literature (Scheerens et al., 2003): How is a secondary school that adds CVA in language different or similar to a school that does not in terms of their processes/practices?, and (b) Are there new processes/practices that emerged from the data relevant to highlighting potentially new aspects of effective/ineffective schooling in order to differentiate between the more and less effective schools? If that is the case, what are they? In this way the impetus for carrying out an Instrumental Case Study (ICS) focusing on two schools was to generate a deeper understanding of how and why the differences in performance between these schools occurred.

How ICS schools were chosen: a key issue of this study

MLM was used to conduct a secondary data analysis of a working sample of 2,283 schools and 176,896 10th grade students (16 year olds) in language SIMCE tests. The student outcome data was matched to (a) Prior attainment at age 14 in 4 subjects: Language, Mathematics Social and Natural Sciences; (b) Pupil characteristics/background (Gender, Number of books in the house, Number of people in the house and Parental education) provided by questionnaires administered by MINEDUC to the students' parents in

2004 and 2006; and (c) School context (School mean prior attainment)[1]. After carrying out MLM it was not only possible to distinguish between schools where students' relative progress was significantly better, average and poorer than predicted on the basis of their intake characteristics, but also to track the schools' identifications. A purposive qualitative sampling took place, identifying two schools that differed in terms of their CVA language measures: School A had one of the highest CVA scores and school B one of the lowest, so were identified as performing better and worse than would be expected after accounting for differences in the student intake.

School A and B descriptive characteristics

According to the official information provided by SIMCE web site[2], school A and school B are classified in the "middle" SES group, which means that on average parents have 12-13 years of education, their monthly income between $325,000 and $550,000 and between 25% and 42% of the students are described as vulnerable. Although school A is a public school that provides education from year 7 to year 12 and school B is a subsidized private school that provides education for primary and secondary education, both enroll girls and boys, are mid-sized and are located in two urban municipalities within the Metropolitan region in the city of Santiago. In both schools the HT had been in their post for a significant period (>10 years). Although some differences in the context and profile of the two schools could have affected the overall results, this descriptive information supports the claim that the ICS is comparing "like with like".

Participants

Table 3
Participants, data collection techniques and aims

Participants	Data collection techniques	Aim
2 Head-teachers (1 each school)	Semi-structured interviews (2 hours each, 4 hours in total)	Explore organizational structures and leadership activities
2 Curriculum Developers ("Jefe UTP") (1 each school)	Semi-structured interviews (2 hours each, 4 hours in total) and field notes	Explore organizational structures and instructional strategies
2 Year 10 language teachers (each school)	Semi-structured interviews (3 hours each, 6 hours in total) and field notes 10 Year 10 (5 each school) classroom observations	Explore classroom climate and instructional strategies

Three different key stakeholders were included in the ICS carried out in school A and B, in order to integrate different perspectives: two Head-Teachers (HT), two Technical-Pedagogic teachers (UTP) and two language teachers within each school. The reason to include UTPs (the pedagogical head who provides instructional support and control) is given by the fact that they share the leadership with the HT. While HTs play predominantly an administrative or managerial role, UTPs play a pedagogical one. Although this dual leadership at the school level has recently been recognised as a salient problematic feature of the Chilean school system (Anderson, Marfán, Sánchez, & Horn, 2011), the effects of this division of labor on school effectiveness has received little attention. Therefore the ICS included both participants. This decision of including different participants had also the aim of avoiding the well-known tendency within EER of over-relying on the perception of the HT as the main data provider (Fertig & Schmidt 2002; Scheerens et al., 2003). Finally, qualitative research literature has stressed the importance of including at least three perspectives in order to triangulate and generalise data (Glaser & Strauss, 1967).

[1] Details of the data modelling can be found in Muñoz-Chereau (2013).
[2] In order to maintain the anonymity of participants, the link to school A and B SIMCE results is not reported.

The ICS was focused on language in general and in Year 10 specifically. This was the case because even though the quantitative data referred to a previous cohort of students, it was the same year group. So the ICS was carried out in 2010 with a special focus on Year 10.

Data collection and instruments

Different instruments were used to collect data from the ICS participants, as described in Table 4. The semi-structured interviews with HTs, UTPs and Language Teachers were the primary data-gathering instrument. This type of interview was chosen in order to provide adequate coverage for all the dimensions studied. The questions were developed in the form of a general statement, which was then followed by sub-questions for further data collection. The interviews were tape-recorded to ensure a precise account of the conversations.

The researcher also observed classrooms taught by the teachers interviewed. These observations were focused on language in general and in Year 10, specifically. This was the case because even though the quantitative data referred to a previous cohort of students, it was the same year group. So, classroom observations were made as a way of obtaining an in-depth understanding of the classroom climate and the instructional strategies that Language Teachers developed with Year 10. Audio recordings and field notes were carried out during the classroom observations.

Analysis

The method of analysis chosen for this study was a hybrid approach of qualitative methods of thematic analysis (Fereday & Muir-Cochrane, 2006) that first incorporated an approach of deductive a priori codes. The analysis was then followed by a data-driven inductive approach allowing for themes to emerge directly from the data. In this way the researcher also developed new codes by directly examining the data.

Deductive a-priori codes

The analysis was carried out step-by-step. First, the data collected was transcribed and translated from Spanish to English in a word processing document. Then the researcher segmented the data into meaningful analytical units using the theory-driven a priori codes. This procedure was done through reading, listening to and summarizing the raw data.

Data-driven categories

In order to make sense of the data that was not included in the a priori codes, the researcher used Memos, that is, short phrases written in the margins of transcripts about what she found puzzling, insightful, and interesting in the data, taking note of potential new themes identified in the text. In this way, two new categories needed to be created in order to describe the local circumstances. Without paying attention to Agency and Trust, the description of the ICS located in challenging contexts would have been incomplete. In this way, instead of coming up with a new list of effective school dimensions or factors, this study first used and then expanded the 14 enhancing SE factors as a way to resolve the tension between context-specificity and generalizations described before.

Results

Access, researcher position and reflexivity

It is important to mention that gaining access to school A and B was a difficult task for various reasons. School A was actively engaged in multiple activities (other research, collaborating with other local schools, etc.) and it was necessary to persevere through e-mails and phone calls before those responsible agreed to participate in the study. This was perhaps the case because this study is by no means alone in the

assessment of the school as an effective one. In fact, a recent interview in an influential local newspaper identified the school as an example of an effective educational institution working in disadvantaged contexts. Although the initial contact was straightforward with School B, collecting data turned to be more elusive as, unlike School A, this was the first time that the school had participated in SER/EER research, and school authorities found it difficult to see the potential benefits that could arise from it. For example, the HT did not allow pictures to be taken or give access to the teachers' room, arguing that it was against the school rules.

So no matter how in-depth the descriptions presented in the ICS attempted to be, they portrayed things mainly from an outside perspective, as I was not a member of the ICS schools. As the HT from school A explained: "[…] *when you get into the teachers' room you will realise that nobody will make a big fuss because we are used to being observed by outsiders*", and not a neutral, but a critical one: "*maybe you as an educational researcher are going to destroy me (laughs).*" In this sense, it was evident that I was perceived as an outsider with a judgemental attitude (reinforced through the activities of observing, asking questions, taking notes, etc.), which generated some discomfort. The advantage of this position was that it gave me a certain perspective and privileged access to the data collected, as I related the experience to my previous role as an educational psychologist working with schools for more than 10 years. I was "embedded" in the Chilean school culture, a fact that allowed me to compare schools A and B with the hundreds of Chilean schools I have visited as a practitioner. Despite the fact that I knew the field, I am aware that, to some extent, being a newcomer to the ICS schools negatively impacted the data collected, in the sense that it is unlikely that from my position I could fully understand the complexities operating in the schools in general, and all the meanings that the participants attached to their practices in particular. In order to do so, I would have needed to choose schools that I had already worked with, contradicting the design of this study, oriented to study schools according to their CVA performance.

The 16 processes or dimensions will be summarized for each school, providing some exemplary quotes from the participants in Table 5.

Processes and practices in school A that added value in language

(a) Achievement orientation and high expectations upon students and teachers, with a clear focus on raising and maintaining high attainment.
(b) Leadership was a mix of human and educational style, which fitted well with the challenging school context.
(c) Consensus among staff was found to be mainly related to the academic process. It was clear what was expected of each participant (for example, prepare handouts, but not lesson plans) and support was provided accordingly (coaching from the HT, autonomy to try different ways of action, resources to choose method and materials). This consensus led to cohesion, expressed in the familiarity and relaxed interactions between the school members.
(d) Good internal relationships were seen as an aim in itself that characterized the school climate.
(e) The evaluative culture in school A was oriented to meet external accountability requirements.
(f) Classroom climate was characterized by a work attitude in most of the students and the teacher made fun of himself, thus enhancing the fun factor.
(g) They gave predominance to effective learning time, which was expressed in an extended length of the school day, homework, buses to carry students from remote areas to and from the school every day and an efficient use of time on tasks within the classroom.
(h) Differentiation was an implicit policy implemented through tracking students in ability groups organized in parallel classes.
(i) Orderly atmosphere was seen as a means to academic work with a set of explicit rules that were known by everybody.

However, not everything was trouble-free in this school. The ICS also showed how the following five factors were problematic/absent in the school:

(j) Teaching to the test undermined opportunity to learn.
(k) Parents were seen as resource providers undermining their involvement.
(l) The lessons followed a structure that excluded independent practice.

(m) The teaching style was centered on the teacher without independent learning.
(n) The way feedback was provided supported more competition than the students' learning process.

Focusing on the processes and practices that were included in the analysis in order to complement the ICS descriptions:

(o) Agency: Although school A did not consider the immediate local community as a valuable key player, they felt that they worked in a toxic/challenging environment that they could cope/manage in order to bring about change.
(p) Trust: The interaction between staff and students was based on trust. They believed in the honesty, goodness, skill and safety of the school members, and acted accordingly, specifically referring to teachers.

Sixteen processes and practices in school B that did not add value in language

The findings pertaining to school B are as follows:

(a) Low expectations upon students and the orientation were not on achievement, but on enacting moral values.
(b) Leadership was mainly symbolic and technical, focused on maintaining discipline and control, which generated fear, distrust and isolation between staff.
(c) Despite the consensus among staff concerning the importance of maintaining discipline, they lacked cohesion related to the academic process.
(d) Teachers felt like islands, without significant links to support or improve their work. School climate had an overemphasis placed on order and discipline to the detriment of internal relationships.
(e) Evaluative practices were implemented mainly as a gatekeeper: to fail or pass students.
(f) In the classroom the work attitude was interrupted by conflicts, there were "invisible" children, and jokes were made at the expense of the pupils.
(g) There was little importance given to effective learning time, apparent in the fact that they had a short school day, did not send homework to students in secondary school, and spent less time on tasks during lessons due to frequent interruptions, as well as the inclusion of non-academic activities, such as daily religious celebrations.
(h) There was no differentiation as a school policy or practice.
(i) Regarding orderly atmosphere, although they defined order and discipline as one of their main features, it seemed that the rules were unknown, created as they went along. Rather than being a means to academic work, order was seen as an end in itself, to the detriment of academic work (i.e. the interruption of the class asking for pony-tails for P.E.).
(j) Lack of planning undermined opportunity to learn.
(k) Parents were seen as resource providers, which undermined their involvement.
(l) The lessons followed a structure that excluded independent practice.
(m) The teaching style was centred on the teacher without independent learning.
(n) The way feedback was provided supported more competition than the students' learning process.

Focusing on the processes and practices that were included in the analysis in order to complement the ICS descriptions, the following was found regarding school B:

(o) Agency: they passively accepted that the social context determined their work.
(p) Trust: they did not rely on the honesty, goodness, skill or safety of the school members (i.e. "*the child is always trying to find a way of doing what is forbidden*").

Table 4
Illustrative quotes of the 16 effectiveness enhancing dimensions used in this study

N	Effectiveness-enhancing dimensions	Illustrative quotes	
		School A	School B
1	Achievement orientation, high expectations: clear focus on the mastering of basic subjects	'I feel that my challenge is to get good results with my students…now and then the HT asks me 'How are YOUR results going? He doesn't ask me 'How are THE results going, but MINE, and I do feel that obtaining high academic results is a personal challenge…' (Interview, Teacher).	'Even though we do give importance to the academic aspect, we focus on providing a moral education…an education oriented to help the children become good people…because when someone asks us where the school is located, and we answer… they say, 'Ah! In front of that slum'…they inevitably raise an eyebrow like saying 'You can't be a good school if you are working in that area'…if you know what I mean' (Interview, HT).
2	Educational leadership: indirect control and influence on the school primary process	'A strength we need as heads is the ability to manage the uncertainty we face every day at the school…when I talk to my teachers and they complain that they are tired and worried, I reply that that is a good sign…if you worry, you will try to find solutions and work harder!…So we end up laughing…I don't know…it seems so obvious to me that they [teachers] need to be empowered…give them room, give them the opportunity to play this game in their own way, with their own resources…they need to realize that the challenge of getting good results is their own…but in order for that to happen, I exempted them from administrative tasks and bureaucracy…' (Interview, HT).	'She [the HT] is concerned with administrative, not educational tasks… it is not the same peeping into a classroom generating fear, than knowing and understanding the work that is taking place…' (Interview, UTP).
3	Consensus and cohesion among staff: Coherence, consistency and continuity among the staff	'When I realized that the children needed to read more primary sources, I worked with my colleagues from Science and History. We chose texts according to the relevant content that they were working on, and then I taught them strategies to improve their comprehension' (Interview, Teacher).	'I feel like an island…I try my best with my students, but I don't feel supported in any way by the rest of the staff' (Interview, Teacher).
4	Curriculum quality/opportunity to learn: The degree in which the implemented curriculum matches the achievement curriculum	'We have increased the time for teaching and learning… but neglecting contents and sectors that are not included in the tests' (Interview, Teacher). They called this 'modifications': the last 2 months of the school year 'the levels that are taking the national tests, focus exclusively on language, maths and science' (Interview, HT).	'I am always a bit worried that I won't be able to cover the whole programme' (Interview, Teacher).

5	School climate as orderly atmospheres: School culture oriented to maintain orderliness	In school A, at the beginning of the school year each student signed '12 non-negotiable rules and regulations' that stressed that they must 'Come to school every day and on time. There is a punishment if you are late/ Respect each member of the school through your actions and attitudes/Come to school wearing the uniform/Be on time in the classroom in an orderly manner, sit in your assigned place and start working immediately/Come prepared to work every day, etc.	When the lesson is finishing, the PE teacher came inside the classroom and said in a loud voice 'Those that are not wearing a pony-tail can't join the lesson' generating a worrying atmosphere among the girls' (Classroom observation, Year 10)
6	School climate in terms of effectiveness orientation and good internal relationships	'In our mission we declared ourselves pluralists in three dimensions: first, religion: we cannot bring in our approaches to religion, because if I am Catholic and you are Mormon and we stress our differences, we are not going to enjoy working together… the same applies for our political and sexual orientation… these issues cannot divide us, so I stress that no matter your personal preferences, if you are going to work in this school, you have to be loyal to the institutional mission and vision in terms of embracing pluralism in order to work together, not letting our differences be more important than our shared views' (Interview, HT).	'…So we go into the classroom at random and open every bag, I even look in their pockets and the students cannot refuse, because we make this rule explicit when they enroll register at the school, so if they don't like it, they have to go to a different school…' (Interview, HT)
7	Evaluative potential: Aspirations and possibilities of a school to use evaluation as a basis for learning and feedback	'We are continuously evaluated from outside…apart from the ones done by the teachers. For example, yesterday the students took a test from DEMRE, University of Chile, and before that we took a test from Andres Bello, Cepech…we take advantage of any institution dealing with evaluation, like the Catholic University, that can give us 'free overviews' of our school […]' (Interview, HT).	'It seems obvious that after a certain period of time we should audit our own assessment system, but we don't…in the last couple of years we improved our internal assessment rules, because it had lots of deficiencies…and we do have at the end of each semester an assessment meeting in order to discuss student attainment and other issues, but then, we don't translate that into priority areas…so one thing is to have an assessment procedure, and a different one is to transform it into a tool' (Interview, UTP).
8	Parental involvement: Relationship between the actual degree of involvement of parents and the effort displayed by the school to engage them.	'We work with parents through our shared financing system…every month each parent pays the school 3,000 pesos and with this money the teacher pays for the materials provided, as well as 1,500 per class to maintain the garden and the building…in this way parents see how we invest their money' (Interview, HT).	'I find it really difficult dealing with parents…I am quite happy with them up to year 4, but at higher levels just a few come to parents meetings to complain or talk about superficial things, like the school party at the end of the year…I can see that students are left alone in the sense that they don't spend time together, mainly because mothers are working…so there is no control, they don't sign their children's homework and if you ask them if their child studies at home, they don't have a clue.' (Interview, HT).

9	Classroom climate: Good, ordered relationships, work attitude and satisfaction with the classroom interactions.	'Now you might think I am mad about PSU [the national test to get into University]…well…you are right! (Students laugh).' Later on, when explaining a poetic type of text he declaimed: 'poet, don't talk about a rose, make the rose bloom in the poem…ah, now I turned into a poet!' (Classroom Observation, Year 10).	The teacher is giving an example related to non-verbal communication and says that just because two students are sitting together, like Tom and Mary [two actual students], it doesn't mean they are having a more serious relationship. She gives this example pointing to the pupils who are sitting together. The boy blushes and the girl also seems embarrassed. The teacher expands the example by saying that just because a boy is wearing a pink bracelet [showing a boy's actual bracelet to the whole class] it doesn't mean that he is gay. All the class laughs, except for the boy' (Classroom Observation, Year 10).
10	Effective learning time: Importance and implementation of effective time on task at the class and school level	School A gave predominance to effective learning time, which was expressed in an extended length of the school day, homework, buses to carry students from remote areas to and from the school every day and efficient use of time on tasks within the classroom.	School B gave little importance to effective learning time, which was expressed in the fact that they had a shorter school day, did not send homework to students in secondary school and spent less time on tasks during lessons due to frequent interruptions, as well as the inclusion of non-academic activities, such as daily religious celebrations.
11	Structured instruction: Well organized and closely monitored lessons	Lessons' observed in both schools followed a similar structure that excluded independent practice and could be summarized in four parts: (a) looking back (b) presenting new subject matter (c) guided practice and (d) giving feedback and/or correction. Although both schools showed successful class management strategies, none of the observed lessons incorporated independent practice, revealing a type of instruction heavily based on direct instruction from the teacher, conceived as a content provider and content mediator.	
12	Independent Learning: Extent to which pupils are responsible and make decisions concerning their learning according to their own interests	'Students are not made responsible for their own work, let alone choose their own assignments' (Classroom observation, School A). Students didn't have enough opportunities to develop these skills. The main difference between the case study schools in reference to this dimension was the amount of time given to each part of the class: in School A much more importance was given to guided practice, whereas in School B Priority was given to looking back.	
13	Differentiation: How the school deals with differences between pupils and takes care of pupils with learning and behavioral problems	'If you look at the SIMCE results from last year, they were good as an average, but in terms of effectiveness we have managed to support the progress of the lower group, and I think this year the results will be similar.' (Interview, Teacher).	There was no differentiation as a school policy or practice. In School B most of the teachers were confronted with the challenge of teaching a wide range of student abilities. There was only special monitoring after school hours, dedicated exclusively to pupils at risk of repeating the year. But there was no strategy oriented to setting or tracking students during the school day.

14	Reinforcement and feedback: Basic requirement for learning that deals with the rapport to pupils in connection with their achievement	'The teacher sticks the results for each student in the previous tests on the wall. A student rushes to have a look at the ranking, checking his own results, but also the results of his classmates' (Classroom Observation, Year 10).	'The teacher started the session by giving their marked tests back to the students. The tests were ordered from the upper to the lower grade, so even though she did not make it explicit, one could figure out the relative position of each student' (Classroom Observation, Year 10).
15	Agency: Sense of efficacy, ability or determination to pursue the school goals.	'We manage and cope daily with toxic environments and uncertainty…for example the teacher gets to the classroom and a girl might be crying because something happened [at home]…they take problems and worries to school …so how does the teacher engage them with the learning process? That is the main challenge!' (Interview, HT).	'How to ask a child to do more than he or she possibly can? … I tell the parents that we do have some tools, but not all of them, and if the parents don't support their child's education, there is nothing we can do […]' (Interview, HT).
16	Trust: to have a belief or confidence in the honesty, goodness, skill or safety of a person, organization or thing" (Cambridge Advanced Learner's Dictionary).	Our case is showing that young teachers, just finishing university, have the skills and abilities to generate high quality learning…if you don't oppress them or stand in their way with pointless administrative tasks' (Interview, HT).	The child is always trying to find a way of doing what is forbidden' (Interview, HT).

In summary, this study showed how a generic model of school effectiveness was useful to describe the similarities and differences in the processes and practices of the ICS schools. Nine factors that were found to be different in the practices or processes in the secondary school that added contextual value (school A) when compared with the school that did not (school B) were identified. These factors were: (a) Achievement orientation, (b) Educational leadership, (c) Consensus and cohesion among staff, (d) School climate in terms of effectiveness orientation and good internal relationships, (e) Evaluative potential, (f) Classroom climate, (g) Effective learning time, (h) Differentiation, and (i) School climate as orderly atmospheres. Then, five factors that were found to be similarly problematic or absent in both case study schools were also outlined: (j) Curriculum quality/opportunity to learn, (k) Parental involvement, (l) Structured instruction (m) Independent learning, and (n) Reinforcement and feedback.

However, the generic model of school effectiveness (14 enhancing factors) was found to be incomplete when trying to describe the quality of education provided by the ICS schools. It appeared that schools A and B could not be effective without a supportive professional context. Along this line, it was not the educational leadership or the climate alone that seemed to make the difference. More precisely, what emerged from the qualitative analysis was that two new dimensions needed to be included: "Agency", defined as the ability or determination to pursue the school goals, and "Trust", understood as the belief or confidence in the honesty, goodness, skill or safety of a person, organization or thing, which were also playing a key role. Despite the fact that both schools worked in similar contexts, school A showed strong Agency and Trust, which implied playing a transformational and progressive role, in the sense that their stakeholders believed they could successfully challenge the limitations imposed by the background. School B found it much harder to cope and not despair, precisely because they reproduced what the students brought to the school in a more reactionary way.

Five of the dimensions explored in the ICS blurred instead of sharpened the differences between the two schools, showing that not everything was unproblematic in the most effective school. After using CVA measures as a screening instrument to carry out the ICS, it was possible to conclude that being effective in one academic subject was by no means everything that mattered. This research posed the need to take a broader approach to assessment by more evenly distributing incentives to schools, in order to amplify their focus on more areas to increase performance within Chilean education.

A matter of concern that emerged when carrying out the ICS was its actual implications for the schools. Regarding school B in particular, the question is the extent to which the school had the capacity to improve. Given that the school processes described in the ICS can be seen as a barrier to improvement, the complexity of implementing change within this school seemed very high. Overcoming the negative organisational barriers described above is essential to effectively implementing sound improvement initiatives. Though previous studies with low-performing schools have consistently found that these schools lack the capacity to improve on their own (O'Day, 2002, cited in Rosenkvist, 2010), it is necessary to be careful not to generalize, due to the fact that nothing was said in terms of other subject areas or among different group of pupils. The fact that school B has been described as being backward in the analysed dimensions does not mean that it could not be advanced in other aspects. School B was praised as a safe and orderly institution. This could have been serving other purposes not stated in the ICS.

Another matter of concern was that school A taught to the test to an extent that undermined students' opportunities to learn. This provided evidence that the assessment emphasis operating in Chile places such a strong burden upon schools that it produces unintended consequences or perverse incentives. This conclusion resonates with many of the issues and limitations that SER has been dealing with in recent decades (Angus, 1993; Ball, 1995; Elliot, 1996; Fielding, 1997; Scherrer, 2011; Slee et al., 1998; Wrigley, 2004). The question that emerged from the analysis is "effective for what?"

Discussion

First of all, it is clear that although SER/EER focusing on processes linked with outcomes has accumulated a large amount of evidence about what has distinguished school and classroom practices in effective and ineffective schools in western countries, there has not been much empirical research conducted in Latin America, let alone Chile. The one previous relevant study in Chile by Belleï and colleagues (2004) has significant methodological limitations, although their findings are nevertheless broadly supported and extended by the current study. Therefore, this study has started filling this gap by conducting an ICS in an effective and an ineffective Chilean secondary school.

Does this study support or weaken the use of VA/CVA to identify schools as highly effective or ineffective? This is a validity question regarding the use of the approach taken by this study for purposes of classifying schools and will be taken as the main angle to discuss the results of the ICS.

Given that measures at the system level can assist decision-making such as accountability, resource allocation, evaluation, improvement or the provision of information to the public, it is important to bear in mind that the classification of schools to be implemented by the Quality Agency of Education in Chile will vary if raw, VA, CVA or other measures are used. This study advocates for the use of CVA measures, by identifying and exploring verifiable examples of both unusually effective and ineffective schools.

This issue of searching for fairer and more accurate approaches of comparing Chilean secondary school performance —such as CVA— is paramount in Chile, given the well-known alarming levels of selectivity, segregation and unfair competition in the educational system (Carnoy, 2007; Carnoy & McKewen, 2003; Hsieh & Urquiola, 2003, 2006; San Martín & Carrasco, 2012; Valenzuela et al., 2009). It is too easy to mistake high or low raw results as indicators of school effectiveness, as raw test scores too often hide "dubious" practices that schools in Chile routinely implement, such as selecting their students on the basis of academic and socioeconomic criteria. I believe that VA in general, and CVA in particular, offer a better answer than raw results to this dilemma.

In line with previous research that has recognized the usefulness of identifying factors at the school level associated with positive VA/CVA results in order to assist all schools to develop plans and targets for improvement (Thomas & Mortimore, 1996), this 16 enhancing factors approach could be used in the future as an evaluation checklist to identify areas of strengths and weaknesses, or as a coding scheme to capture the diversity of dimensions operating in schools, but not as a final blueprint or recipe for school effectiveness, which is obviously beyond the scope of this study.

It is necessary to go a step forward in the "either-or" debate on raw estimates versus CVA estimates. Raw estimates are useful to describe equity and quality issues in the school system as a whole, while CVA estimates allow conclusions concerning the relative quality and consistency of each school's performance. Rather than consider each of these SESI measures in isolation, both of them are needed to look at Chilean schools in a fresh light. However, CVA or raw estimates should not be used as the sole indicators of school effectiveness, but as pieces of a larger accountability system like that proposed by the LGE, which is planning to use not just academic outcomes to evaluate schools' quality. Inspections and other ways of feedback based on other dimensions of schooling could be a promising development to support school improvement initiatives. CVA should be presented alongside other information, especially because any indicator by itself provides a limited picture of schools' performance. As with every empirically-based indicator of school performance, CVA is fallible, depending on variability and bias (OECD, 2008). But when properly implemented, CVA generates a school-level indicator that, in conjunction with other indicators, can provide an informative portrayal of school functioning.

If SIMCE under the new framework imposed by Agencia Calidad de la Educación (2011) will have even higher consequences for schools, then the conclusions drawn from the results need to be justified. So even though it is clear that no single indicator should be the sole basis for high-stakes decisions (National Research Council, 1999, as cited in Braun et al., 2010), this study strongly recommends CVA as a more meaningful measure of school effectiveness than raw scores, due to the fact that it works as a handicap: levelling the playing field.

The focus upon CVA would describe a more favourable picture of these schools (mainly public schools)

than would have been the case had the focus been on other approaches, reducing the sense of failure that has been reinforced in recent decades through the routine publication of SIMCE raw scores. Moreover, the use of CVA models can help to increase the confidence of stakeholders who are concerned about the management of schools and teachers educating socially and economically disadvantaged students. The use of a CVA approach will not only produce more accurate models, but will also be more politically powerful, in the sense that it will send a message of fairness to the stakeholders.

For the given reasons, this study does advocate the suitability of implementing a CVA model as a more equitable measure of school effectiveness, bearing in mind that the decision to focus on specific measures should be aligned with broader policy objectives oriented to improve quality and equity through education. VA/CVA scores can also be used as information to monitor policy initiatives of this kind.

Finally, by the end of 2013, the Chilean Quality Agency (QA) will not only classify schools in four performance groups, but also high-stake consequences (from organizational interventions and monetary sanctions up to school closure) will follow for those schools that are not meeting the standards (i.e. classified as having poor performance). Given that the law states that the procedure for sorting schools could eventually include VA models depending on the availability of data (Ley General de Educación, 2010), this study is timely.

The effort derived from the new assessment framework to be implemented by QA referring to moving away from the current system based on SIMCE raw results is promising, but the imposition of high stakes seems problematic, especially because effective policies are not only those that are clear to the stakeholders, but also those that take into account their ability to implement them (Creemers, 2012). Inevitably negative consequences will follow for those schools classified as having poor performance. Apart from the well-known consequences of failure (i.e. labelling effects and school stigma), making this classification public could be counterproductive, particularly because, as the ICS showed, what seemed to make a difference between school A and school B were not just the well researched dimensions within the EER, but also the far less known Agency and Trust: two dimensions that helped school A to overcome obstacles, to sustain their commitment, resilience and effectiveness over time. This key sense of efficacy is precisely the belief that could be seriously damaged by the way the information is going to be used. It is doubtful whether the failure discourse can help schools recover or improve their performance. What seems clearer is that their improvement will depend to a great extent on the support provided. Along this line, this study helped to identify 16 enhancing dimensions that can support improvement at the school level.

The original article was received on May 2nd, 2013
The revised article was received on August 20th, 2013
The article was accepted on September 2nd, 2013

References

Agencia Calidad de la Educación (2011). *¿Qué es la Agencia de Calidad de la Educación?* Retrieved from http://www.mineduc.cl/contenido_int.php?id_contenido=22276&id_portal=1&id_seccion=4220

Ainley, P. (1994). *Degrees of difference: Higher education in the 1990s*. London: Lawrence & Wishart.

Aitkin, M, & Longford, N. (1986). Statistical modelling issues in school effectiveness studies. *Journal of the Royal Statistical Society, 149*(1), 1-43.

Anderson, S., Marfán, J., & Horn, A. (2011). *Leadership efficacy, job satisfaction and educational quality in Chilean elementary schools. 25th International Congress for School Effectiveness and School Improvement*. Retrieved on February 3, 2013 from www.icsei.net/icsei2011/Full%20Papers/0095.pdf

Angus, L. (1993). The sociology of school effectiveness. *British Journal of Sociology of Education*, *14*(3), 333-45.

Ball, S. (1995). Intellectual or technicians? The urgent role of theory in educational studies. *British Journal of Educational Studies*, *43*(3), 255-71.

Ballou, D., Sanders, W., & Wright, P. (2004). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioural Statistics*, *29*(1), 29-37.

Belleï, C., Muñoz, G., Pérez, L., & Raczynski, D. (2004). *¿Quién dijo que no se puede? Escuelas efectivas en sectores de pobreza en Chile*. Chile: UNICEF.

Braun, H., Chudowsky, N., & Koenig, J. (2010). *Getting value out of value-added.* Retrieved on January 24, 2013 from http://www.nap.edu/catalog/12820.html

Carnoy, J., & McKewan, P. (2003). Does privatization improve education? The case of Chile's national voucher plan. Chapter 1. In D. Planck, & G. Sykes (Eds.), *Choosing choice. School choice in the international perspective* (pp. 24-44). New York: Columbia Press.

Carnoy, M. (2007). Improving quality and equity in Latin American education: A realistic assessment. *Pensamiento Educativo*, *40*(1), 103-130.

Carrasco, A. (2010). *A case-study of the Chilean policy agenda for disadvantaged primary schools: meeting their challenges?* (Ph.D. Thesis). Cambridge: University of Cambridge.

Carrasco, A., & San Martin, E. (2012). Voucher system and school effectiveness: reassessing school performance differences and parental choice decision-making. *Estudios de Economía*, 39(2), 123-141. Retrieved from www.estudiosdeeconomia.cl

Creemers, B. (1996). The goals of school effectiveness and school improvement. In D. Reynolds, R. Bollen, B. Creemers, D. Hopkins, L. Stoll, & N. Lagerweij (Eds.), *Making good schools: Linking school effectiveness and school improvement* (pp. 36-58). London: Routledge.

Creemers, B. (2012). *Developing, testing and using theoretical models for promoting quality and equity in education*. Keynote at the 3rd meeting EARLI SIG. Retrieved on April 2, 2013 from http://www.earli.uzh.ch/programme/keynote-speeches

D'Haenens, E., Van Damme, J., & Onghena, P. (2010). Constructing measures for school process variables: The potential of multilevel confirmatory factor analysis. *Qual Quant, 46*, 155-188.

Duri-Bellat, M., & Mengat, A. (1998). Importance of ability grouping in French "colleges" and its impact upon pupils' academic achievement. *Educational Research and Evaluation*, *4*, 348-368.

Elliot, J. (1996). School effectiveness research and its critiques: Alternative visions of schooling. *Cambridge Journal of Education*, *26*, 199-223.

Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International Journal of Qualitative Methods*, *5*(1), 80-92.

Fertig, M., & Schmidt, C. (2002). *The role of background factors for reading literacy: Straight national scores in the PISA 2000 study*. Retrieved on February 23, 2013 from http://ssrn.com/abstract=323599

Fielding, M. (1997). Beyond school effectiveness and school improvement: Lighting the slow fuse of possibility. *The Curriculum Journal*, *8*(1), 7-27.

Glaser, B., & Strauss, A. (1967). *Discovery of grounded theory. Strategies for qualitative research*. London: Sociology Press.

Goldstein, H. (1995). *Multilevel statistical models*. London: Edward Arnold.

Goldstein, H., Rasbash, J., Yang, M., Woodhouse, G., Pan, H., Nuttall, D., & Thomas, S. (1993). A Multilevel analysis of school examination results. *Oxford Review of Education*, *19*(4), 425-33.

Gray, J, Goldstein, H., & Thomas, S. (2003). Of trends and trajectories: Searching for patterns in school improvement. *British Educational Research Journal*, *29*(19), 83-88.

Gray, J., Reynolds, D., Fitz-Gibbon, C., & Jesson, D. (Eds.) (1996). *Merging traditions: The future of research on school effectiveness and school improvement*. London: Cassell.

Harris, A., Jamieson, M., & Russ, J. (1995). A study of "effective" departments in Secondary Schools. *School Organisation*, *15*(3), 283-299.

Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* London: Routledge.

Hoyle, R., & Robinson, J. (2003). League tables and school effectiveness: A mathematical model. *Proceedings of the Royal Society of London* B, *270*, 113-199.

Hsieh, Ch-T., & Urquiola, M. (2003). *When schools compete: How do they compete? An assessment of Chile's nationwide school voucher program*. (Working Paper series 1008). Retrieved on May 21, 2012 from http://www.nber.org/papers/w10008

Hsieh, Ch-T., & Urquiola, M. (2006). The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program. *Journal of Public Economics, 90*(8-9), 1477-1503.

Klieme, E. (2012). *Qualities and effects of teaching. Integrating findings across subjects and cultures.* (Keynote at the 3rd meeting EARLI SIG). Retrieved on March 19, 2013 from http://www.earli.uzh.ch/programme/keynote-speeches

Ladd, H., & Walsh, R. (2002). Implementing value-added measures of school effectiveness: getting the incentives right. *Economics of Education Review,* 21, 1-17.

Ladd, H., & Zelli, A. (2002). School-based accountability in North Carolina: The responses of school principals. *Educational Administration Quarterly*, *38*, 494-529.

Lambert, D., & Lines, D. (2000). *Understanding assessment. Purposes, perceptions, practice.* London: Routledge.

Leckie, G. (2008). *Multilevel modelling of school differences in educational achievement*. (Ph. D. Thesis). Bristol: University of Bristol.

Leckie, G., & Goldstein, H. (2007). *The limitations of using school league tables to inform school choice*. Retrieved on March 23, 2010 from http://209.85.229.132/search?q=cache:bR0DoCO0iaUJ:eprints.ncrm.ac.uk/265/1/School_league_tables.pdf+unfair+league+tables+filetype:pdfandcd=5andhl=enandct=clnkandgl=uk

Leckie, G., & Goldstein, H. (2011). Understanding uncertainty in school league tables. *Fiscal Studies*, *32*(2), 207-224.

Ley General de Educación (2010). *Establece la Ley General de Educación.* Retrieved on February 1, 2013 from http://www.leychile.cl/Navegar?idNorma=1006043

Luyten, H. (1998). School effectiveness and student achievement, consistent across subjects? Evidence from Dutch elementary and secondary education. *Educational Research and Evaluation*, *4*(4), 281-306.

Maden, M. (2001). *Success against the odds. Five years on. Revisiting effective schools in disadvantaged areas*. London: Routledge.

Manzi, J., San Martín, E., & Van Bellegem, S. (2013). School system evaluation by value-added analysis under endogeneity. *Psychometrika*. Retrieved on September 11, 2013 from http://link.springer.com/article/10.1007%2Fs11336-013-9338-0/fulltext.html. doi: 10.1007/S11336-013-9338-0

Mortimore, P. (1991). School effectiveness research: Which way at the crossroads? *School Effectiveness and School Improvement*, *2*(3), 213-229.

Muñoz-Chereau, B. (2013). *Searching for fairer ways of comparing Chilean secondary school performance: a mixed methods study investigating contextual value added approaches.* (Ph. D. Thesis). Bristol: University of Bristol.

Murillo, J. (2007). School effectiveness research in Latin America. In T. Townsend (Ed.), *International handbook of school effectiveness and improvement* (pp. 75-92). The Netherlands: Springer.

OECD (2008). *Measuring improvements in learning outcomes. Best practices to assess the value-added of schools*. Retrieved on April 12, 2011 from www.oecd.org/publishing/corrigenda

OECD (2011a). *La medición del aprendizaje de los alumnos. Mejores prácticas para evaluar el valor agregado de las escuelas*. Retrieved on September 11, 2013 http://www.oecd.org/edu/school/47871357.pdf

OECD (2011b). *Society at a glance – OECD social indicators, key findings: Chile*. Retrieved on December 6 from http://www.oecd.org/dataoecd/39/23/47572883.pdf

Opdenakker, M., & Van Damme, J. (2000). Effects of schools, teaching staff and classes on achievement and well-being in secondary education: Similarities and differences between school outcomes. *School Effectiveness and School Improvement*, *11*(2), 165-196.

Ramírez, M. J. (2007). Diferencias dentro de las salas de clases. *Estudios Públicos*, *106*, 5-22.

Raudenbush, S., & Willms, J. (1995). The estimation of school effects. *Journal of Educational and Behavioural Statistics, 20*(4), 307-335.

Rosenkvist, M. A. (2010). Using student test results for accountability and improvement: A literature review. *OECD Education Working Papers, 54*. OECD Publishing. Retrieved on May 19, 2011 from http://dx.doi.org/10.1787/5km4htwzbv30-en

Rutter, M., Maughan, B., Mortimore, P., Ouston, J., & Smith, A. (1979). *Fifteen thousand hours: secondary schools and their effects on children.* Cambridge, Mass.: Harvard University Press.

Sammons, P. (1999). *School effectiveness coming of age in the twenty-first century.* The Netherlands: Swets and Zetlinger.

Sammons, P. (2007). *School effectiveness and equity: Making connections. A review of school effectiveness and improvement research. Its implications for practitioners and policy makers*. Retrieved from http://cdn.cfbt.com/~/media/cfbtcorporate/files/research/2007/r-school-effectiveness-and-equity-full-2007.pdf

Sammons, P., Cuttance, P., Nuttall, D., & Thomas, S. (1995). Continuity of school effects: A longitudinal study on primary and secondary school effects on GCSE performance. *School Effects and School Improvement, 6*(4), 285-307.

Sammons, P., Thomas, S., Mortimore, P., Owen, Ch., & Pennell, H. (1993). *Assessing school effectiveness: Developing measures to put school performance in context*. London: Institute of Education, International School Effectiveness and Improvement Centre.

Sammons, P., Thomas, S., & Mortimore, P. (1997). *Forging links: effective schools and effective departments.* London: Paul Chapman Publishing.

San Martín, E., & Carrasco, A. (2012). Clasificación de escuelas en la nueva institucionalidad educativa: contribución de modelos de valor agregado para una responsabilización justa. *Temas de Agenda Pública*, *53*, 1-18. Santiago, Chile: Centro de Políticas Públicas.

Scheerens, J. (2000). Improving school effectiveness. Fundamentals of educational Planning. *International Institute of Educational Planning*, *68*. Paris: UNESCO.

Scheerens, J. (2012). *Theories of educational effectiveness and ineffectiveness.* Keynote at the 3rd meeting EARLI SIG. Retrieved on April 2, 2013 from http://www.earli.uzh.ch/programme/keynote-speeches

Scheerens, J., Glas, C., & Thomas, S. (2003). *Educational evaluation, assessment and monitoring. A systemic approach.* The Netherlands: Swets and Zeitlinger Publishers.

Scherrer, J. (2011). Measuring teaching using value-added modelling: The imperfect panacea. *NASSP Bulletin, 95*(2), 122-140.

Shagen, I., & Hutchison, D. (2003). Adding value to educational research: The marriage of data and analytical power. *British Educational Research Journal*, 29(5), 749-765.

Slee, R., Weiner, G., & Tomlinson, S. (1998). *School effectiveness for whom? Challenges to the school effectiveness and school improvement movements*. London: Falmer Press.

Smyth, E. (1999). Pupil performance, absenteeism and school drop-out: A multidimensional analysis. *School Effectiveness and School Improvement*, *10*, 480-502.

Stake, R. (1995). *The art of case study research.* London: Sage.

Steele, F., Vignoles, A., & Jenkin, A. (2007). The effect of school resources on pupil attainment: a multilevel simultaneous equation modelling approach. *J. R. Statist. Soc., 170*(3), 801–824.

Stoll, L., & Mortimore, P. (1997). School effectiveness and school improvement. In J. White, & M. Barber (Eds.), *Perspectives on school effectiveness and school improvement*. (pp. 14-25). London: Institute of Education.

Stoll, L., & Myers. K. (1998). *No quick fixes. Perspectives on schools on difficulty.* London: Falmer Press.

Stringfield, S. (1998). Anatomy of ineffectiveness. In L. Stoll, & K. Myers (Eds.), *Perspectives on schools on difficulty* (pp. 209-221). London: Falmer Press.

Teddlie, C., & Reynolds, D. (2000). *The international handbook of school effectiveness research.* London: Falmer Press.

Thieme, C., Tortosa-Ausina, E., Prior, D., & Gempp. R. (2012). *Valor agregado multinivel y factores contextuales en educación: una comparación no paramétrica robusta*. Retrieved on February 4, 2013 from https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=xveep&paper_id=53

Thomas, S., Salim, M., Muñoz-Chereau, B., & Peng, W-J. (2011). Educational quality, effectiveness and evaluation: Perspectives from China, South America and Africa. In Ch. Chapman, D. Muijs, A. Harris, D. Reynolds, & P. Sammons (Eds.), *Challenging the orthodoxy of school effectiveness and school improvement: Towards new theoretical perspectives* (pp. 125-145). London: Routledge.

Thomas, S., & Mortimore, P. (1996). Comparison of value added models for secondary school effectiveness. *Research Papers in Education*, *11*(1), 5-33.

Timmermans, A., Doolaard, S., & de Wolf, I. (2011). Conceptual and empirical differences among various value-added models for accountability, *School Effectiveness and School Improvement, 22*(4), 393-413.

Treviño, E., & Donoso, F. (2010). *Agrupación de escuelas para intervenciones de política: análisis del caso chileno*. Retrieved on January 30, from http://mt.educarchile.cl/MT/jjbrunner/archives/Trevino_ValorAgregado2010.pdf

Valenzuela, J., Bellei, C., Osses, A., & Sevilla, A. (2009). *Causas que explican el mejoramiento de los resultados obtenidos por los estudiantes chilenos en PISA 2006 respecto a PISA 2001. Aprendizajes y Políticas.* Santiago, Chile: Centro de Investigación Avanzada en Educación, Universidad de Chile.

Van de Grift, W., & Houtveen, T. (2007). Weaknesses in underperforming schools, *Journal of Education for Students Placed at Risk (JESPAR*), *12*(4), 383-403.

Willms, D., & Raudenbush, S. (1989). A longitudinal hierarchical linear model for estimating school effects and their stability. *Journal of Educational Measurement*, *26*(3), 209-232.

Wrigley, T. (2004). School effectiveness: The problem of reductionism. *British Educational Research Journal, 30*(2), 227-244.