

Teacher evaluation in Chile: highlights and complexities in 13 years of experience

Beatrice Avalos-Bevan

To cite this article: Beatrice Avalos-Bevan (2018) Teacher evaluation in Chile: highlights and complexities in 13 years of experience, *Teachers and Teaching*, 24:3, 297-311, DOI: [10.1080/13540602.2017.1388228](https://doi.org/10.1080/13540602.2017.1388228)

To link to this article: <https://doi.org/10.1080/13540602.2017.1388228>



Published online: 12 Oct 2017.



Submit your article to this journal [↗](#)



Article views: 111



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)



Teacher evaluation in Chile: highlights and complexities in 13 years of experience

Beatrice Avalos-Bevan 

Centre for Advanced Research in Education, University of Chile, Santiago, Chile

ABSTRACT

The paper examines the process of establishing a teacher evaluation system in Chile and its acceptance by teachers over time. The conceptual base upon which the system was established is described. Evidence is also examined from a variety of data sources and research related to the evaluation system as well as teachers' use of its results. This evidence includes information from an external review of the system noting tensions between the formative versus accountability purposes of the evaluation system and that are being attended to in the new Teacher Professional Development Law (2016).

ARTICLE HISTORY

Received 11 April 2017
Accepted 2 October 2017

KEYWORDS

Teacher performance evaluation; impact of evaluation instruments; formative and accountability purposes; external review and policy responses

The evaluation of teacher performance has always been part of school educational practices, generally exercised in more or less formal ways by the school's leadership or by an inspectorate system such as in England. However, the gradual influx of quality assurance mechanisms affecting education systems and linked to public concerns about education quality, has led to a diversity of forms of teacher evaluation that increasingly have become external to the schools. These develop for the purposes of certifying new teachers, for promotion within a teacher career structure, and as a means of rewarding high levels of expertise (Millman & Darling-Hammond, 1990; Goodwyn, 2017).

Teacher performance evaluation also found its place in Chile in line with similar purposes in other country contexts such as many states in the US, and is now considered to be an established system. Throughout 13 years of implementation, it has provided evidence of how teachers in the public municipal system carry out their teaching duties as well as of the extent of their content and pedagogic content knowledge. The evaluation system began to operate in 2004 following a long process of discussions among main stakeholders beginning in 1999, but which had already been considered in 1994 during a national congress organised by the Chilean Teachers' Union (*Colegio de Profesores*). This article describes and analyses the evaluation system beginning with a historical overview of its development and the conceptual basis upon which it was built, dealing then with its implementation over time and with what are considered to be its strengths and shortcomings, and ending with a reference to the challenges ahead in the light of the new teachers' law which formally sets up a teacher career system (Ministerio de Educación de Chile, 2016).

Historical overview

The teacher performance evaluation system was established at a time when teachers became a focus of concern for Chilean policy-makers. This, because they were assumed to be responsible for unsatisfactory student achievement during the 1990s as evident in national standardised tests (at the 4th, 8th and 10th grades), and mainly because of reforms occurring in the school system that would require teacher involvement. Reforms of schooling included specific programmes addressing poor learning results of primary school students, changes in the primary and secondary school curriculum, and more importantly a two-hour extension of the school-day in order to provide for a wider offer of curricular activities (Avalos, 2010; Cox, 2003). The role of teachers in these reforms was considered important, and policy-makers believed that they would need assistance to handle the new demands. Thus, the Ministry of Education sponsored teacher professional development activities and teacher visits abroad to learn about teaching and school practices in different countries.

Within the reform context, the examination of the quality of teacher performance became a matter of policy discussion. Linked to this discussion was a clause in the 1991 teacher's Statute that mandated the implementation of a system of teacher evaluation. Aware of this clause, the Teachers' Union had begun discussions in 1993 on their role in furthering a system of teacher evaluation, which would have a formative emphasis and not be in the hands of the school principals. Teachers opposed any evaluation involving school principals as main evaluators as they did not trust that they would be sufficiently free of political prejudice. In fact many of those in office at the time had been appointed during the military dictatorship period (1973–1990) on account of their allegiance to the sitting government and had supported the widespread dismissal of teachers that took place during the 1980s. Thus, with the intention of contributing to an independent and formative evaluation system the Teachers' Union agreed to be part of a committee that would discuss and propose the basic elements needed to build such a system.¹ The committee included representatives from the Teachers' Union, the Ministry of Education and the Association of Municipalities as managers of the public municipal school system, but excluded representatives from private subsidised schools that at the time covered 50% of the publicly funded school system. The committee deliberations led to the proposal of a formative teacher evaluation scheme with two components: a compulsory process to take place every two years leading to salary increments, and a voluntary system embedded in a teacher career ladder with three or four promotion stages and additional salary increments. The system would serve to identify the range of teacher capacity and would provide professional development opportunities for teachers assessed as having an unsatisfactory level of performance. Over time and until 2003, discussions continued at different levels, but principally within the Ministry of Education in consultation with the Teachers' Union and in the midst of continued teacher mistrust and resistance (Avalos & Assael, 2006). Eventually a settlement was reached that included a mixed compulsory formative and summative performance evaluation system designed to assess teachers every four years. The original proposal of an evaluation process linked to a teacher career path was left aside.

After a pilot implementation of the system in 1993, the proposal became a law in 2004 and was applied only to teachers in municipal schools.

The conceptual basis of the teacher evaluation system and its development over time

Clearly, during the discussion of the system and later in its implementation procedures, there was a tension between what are recognised as being two major approaches to teacher evaluation and its relation to education quality and student learning: a formative process directed to informing about teachers' performance quality and providing relevant professional development opportunities, or a summative process directed to the identifying inadequate teachers and removing them from teaching in the light of accountability purposes (Papay, 2012). From the perspective of policy-makers and educational authorities the accountability purpose is valued as a short cut to identifying and providing the education system with competent teachers. But it requires measures that usually link performance assessment to student results in standardised tests based on statistically identifying what a single teacher contributes to his or her students' learning. In this respect, there continue to be uncertainties about the value-added measures that need to be calculated (Goodwin & Miller, 2012). Papay (2012) reviewed a number of studies that point to sources of bias in value-added measures affecting reliability that are not easily resolved. Among these are non-random assignment of students to teachers, the difficulty of separating the contribution to whole class learning of a single teacher from other influences, and the fact of teacher variations in performance from year to year. Also noted are problems with validity given the limited range of student learning that is measured by subject-specific standardised tests in terms of what are broader educational objectives (Papay, 2012).

On the other hand, if teacher evaluation procedures are considered not so much from the point of view of the policy-makers' desire to weed out the incompetent but from the perspective of teacher performance improvement in itself, then a formative non/accountability-based system makes sense and can be valued by the profession. In this respect, as Danielson (2011) puts it, teacher evaluation would have as main its purposes to insure teacher quality in the light of a consistent and shared good definition of teaching and on this basis to promote continuous professional learning.

Within the context of these contrasting approaches, the Chilean system opted to include elements of both, though with an emphasis on the formative focus. On the one hand, the system was defined in its purpose as contributing to the professional development of teachers and to the improvement of their teaching capacities in line with student learning needs. As a system it would also be a source of information for the education system and for teacher education to use in decision-making. On the other hand, the system also incorporated an accountability emphasis associated with dismissal of teachers showing an unsatisfactory level of teaching competence over three consecutive evaluations. Overall, and in line with Danielson (1996, 2010) and others such as Marzano (2012) the system would rest on a shared definition of good teaching. This definition was reached through a wide consultation process with teachers and embedded in a framework of teaching and professional competences agreed upon through Chilean teacher consultations (Ministerio de Educación de Chile, 2003). In what follows the teacher evaluation system is described in more detail.

The framework of good teaching

The key document guiding the teacher evaluation processes sets out the main domains and competences that all teachers are able to recognise as part of their daily work in and outside the classroom. In its formulation it drew from work carried out by Danielson (1996) as well as from standards for initial teacher education developed by the Ministry of Education (Ministerio de Educación de Chile, 2000).

The Framework is organised in four dimensions each with 20 competences. The domains refer to (a) activities teachers do prior to actual classroom teaching with a focus on planning, (b) the climate teachers manage to build in the classroom including the stimulation of students to do their best, (c) actual classroom teaching activities and (d) out-of-classroom professional duties such as collaboration with colleagues and interaction with parents as well as reflective analysis related to their classroom teaching. Thus, rather than specific behaviours the Framework consists of statements describing what teachers do or can do in relation to each major domain. The Framework applies to all teachers and does not specify how teachers at different school levels or with respect to the various subjects should teach. The viewpoint is that the standards offer a perspective of what is to be understood as quality teaching and that rubrics can be used to assess such quality.

The evaluation process

The Chilean Ministry of Education entrusted the handling of the evaluation process to a university measurement centre, which was responsible for setting it up on the basis of the directives included in the Teacher Evaluation Law (Manzi, González, & Sun, 2011). This necessitated the development of a system of rubrics to evaluate each teacher according to the competency framework, the design and piloting of the evaluation instruments to be used, supervision of the actual process whereby teachers would be evaluated and the processing of information for every teacher evaluated. The final decision regarding the evaluation of each teacher was to be taken at municipal level by a group of highly accomplished teachers prepared for the task by the ministerial body responsible for the overall management of the system.

The Chilean evaluation system operates on the basis of four sources of evidence: a portfolio, an interview by a peer teacher with at least five years of experience, a written report of two school authorities on the basis of a set framework and a self-evaluation report by the teacher following a given structure. The most significant of these instruments is the portfolio that every teacher must prepare on the basis of precise instructions provided by the Ministry about how it is to be presented. The portfolio has two sections: (1) a teaching unit involving eight 45-min lessons, the sample of an assessment instrument actually used for the unit taught and responses to a series of questions that indicate reflection about what was achieved after teaching the unit and (2) a video recording of one lesson taught, along with written information about its nature and objectives. An external team performs the filming of this lesson on a date agreed with the teacher.

Among the instruments the portfolio has the highest weighting in the process of establishing the competence level of the teacher being evaluated (60%), followed by the peer interview with 20% and the other two sources of evidence with 10% each.

Once the information is processed teachers are provisionally classified in four performance categories: ‘outstanding’, ‘competent’, ‘basic’ or ‘unsatisfactory’. The Municipal Committee of Teachers referred to above agrees on the definitive classification, after discussion on the evidence provided. Teachers who are deemed to be ‘outstanding’ and ‘competent’ have a four-year period before they are next evaluated, receive various non-monetary incentives and may apply for monetary ones on the basis of a content knowledge examination. Teachers categorised as showing ‘basic’ or ‘unsatisfactory’ teaching performance are provided with specific professional development opportunities and must submit to re-evaluation: after two years for ‘basic’ teachers and after one year for ‘unsatisfactory’ ones. Altogether ‘basic’ teachers have up to three consecutive evaluations to improve and ‘unsatisfactory’ teachers have up to two consecutive evaluations to improve before they are dismissed from the municipal education system (See Figure 1):

Individual teachers receive a report on their performance level for each category as examined, but without specific reference to the rubrics used to classify them. Reports are also provided to the school leadership and to the relevant municipal authority.

Teacher performance over time

The system began to operate with only primary level teachers (first–eighth grade) and was gradually extended to teachers at all levels, including those in special education schools. While it is compulsory for municipal teachers, over time there always has been a number who resist being evaluated, though diminishing, as shown in the Figure 2.

As far as results are concerned, data illustrated in Figure 3, show that over time 67% of teachers consistently tend to perform at the ‘competent’ level (Manzi et al., 2011). However, educational authorities do not consider these results as desirable. They believe that there should be a greater increase of highly accomplished or ‘outstanding’ teachers than what has been the case since 2011, and a lowering in numbers of ‘basic’ and ‘unsatisfactory’ teachers.

There are clear differences in how the sources of evidence reflect the quality of teacher performance as noted by the external evaluation of the system (OECD, 2013). The following Figure 4 illustrates the classification of teachers according to the instrument used.

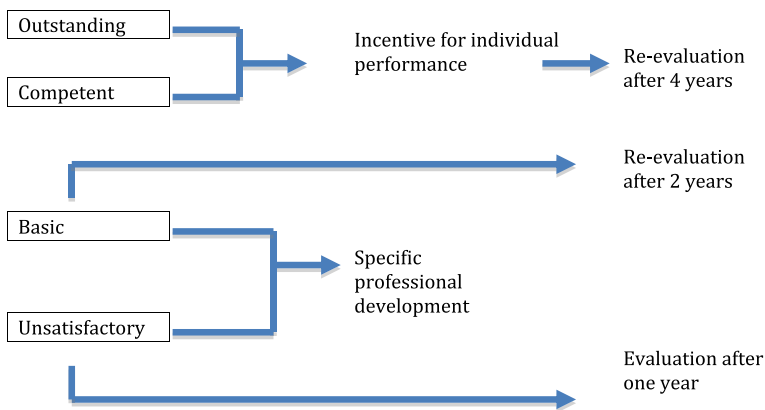


Figure 1. Consequences of evaluation for each performance category.

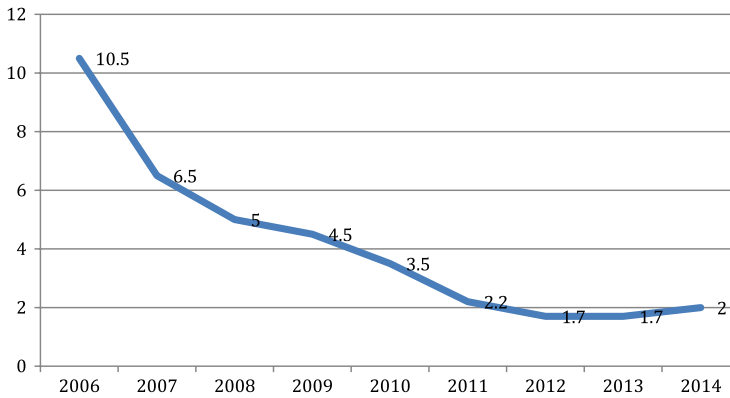


Figure 2. Percentage of teachers resisting evaluation. Source: MIDE UC, Anuarios Estadísticos de la Evaluación Docente: http://www.docentemas.cl/dm06_evaluacion.php.

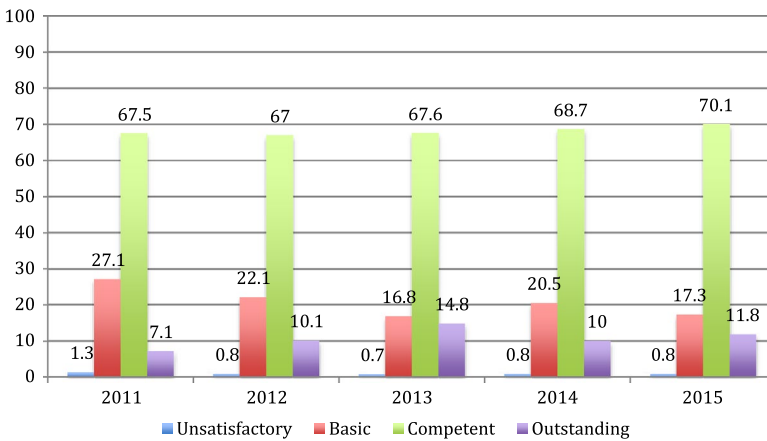


Figure 3. Competence levels of teachers evaluated 2011–2015. Source: Ministerio de Educación: Resultados de la Evaluación Docente (2015) http://www.docentemas.cl/docs/Resultados_Evaluacion_Docente_2015.pdf.

These differences reflect the potential bias of one of the sources of evidence. This is the case of the teacher’s self-assessment, which is based on self-ratings as opposed to actual evidence. As could be predicted the self-assessment tends to provide a much more positive evaluation of a teacher’s performance than does the portfolio, which is based on the actual evidence the teacher provides regarding what he or she is able to do as far as teaching is concerned.

Validity checks

The university institution in charge of implementing the system has been careful to monitor the process and carry out different checks of its quality. Starting in 2005, it developed a research programme around a set of key questions. The first one referred to content validity, that is, the degree to which the system as implemented was in line with the teacher

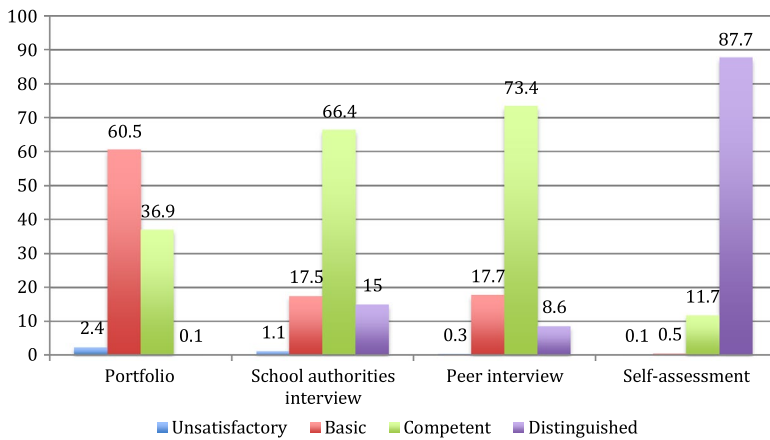


Figure 4. Nature of evidence provided by each evaluation instrument in relation to competence categories. Source: Ministerio de Educación: Resultados de la Evaluación Docente (2015): http://www.docentemas.cl/docs/Resultados_Evaluacion_Docente_2015.pdf, http://www.docentemas.cl/dm06_evaluacion.php.

competences set out in the *Framework of Good Teaching* (Ministerio de Educación de Chile, 2003). Results of the study indicated that the rubrics used to evaluate the evidence adequately covered two of the four domains: classroom climate and classroom teaching practices. However, these rubrics were less adequate for assessing the level of teacher's content knowledge as evident in lesson plans, or for dealing with competences in the fourth domain referred to professional out-of-classroom responsibilities (Taut, Santelices, & Manzi, 2011). In addition to the initial check for content validity these studies also examined the internal consistency of the instruments used from 2005 to 2008 and the trustworthiness of the ratings given (Taut et al., 2011). Results showed that the rubrics used to evaluate the evidence provided by a teacher produce highly consistent and stable results over time and that variance in ratings could be attributed to real performance differences (48%), to corrector or assessor ratings (31%) and to error (17%). Equally important have been studies comparing teachers classified in the evaluation system as 'outstanding' or 'unsatisfactory' with their actual classroom performance and student achievement. In general, these studies provide evidence of a good fit between how teachers are classified and how they actually perform in schools (Taut et al., 2011). The fact that student results were not linked to the evaluation system, has avoided having to face concerns about the reliability of the system as a measure of teacher performance but also raises questions regarding decisions taken to dismiss 'unsatisfactory' teachers after two continuous evaluations on the basis mainly of their portfolio evidence and without 'live' observations of their everyday performance. In other words, analysis of the system raises questions about the validity of its accountability emphasis.

Other related evaluations

Both the Ministry of Education authorities and the general public have been consistently concerned about teacher competency in subject-matter knowledge. However, content knowledge assessment is not specifically a part of the evaluation system, though the portfolio does provide evidence of the uses of knowledge in teaching. Thus, to complement the

formal teacher evaluation, and also to provide incentives for performance improvement, a voluntary evaluation of subject-matter knowledge was established in 2004 for ‘competent’ and ‘distinguished’ teachers. The examination is known as the Voluntary Individual Grant Assignment (AVDI) and consists of a test measuring content and pedagogical knowledge. Depending on their results teachers are re-categorised in terms of their prior evaluation categories and receive a monetary bonus over four years, which varies in its amount depending on how teachers are finally rated. However, teachers do not consider this voluntary examination as a legitimate option and so over time only around a third of eligible teachers have taken the tests. Apart from their own disregard for this type of performance testing another factor contributing to teacher disinterest is the low amount of the bonus, in that it represents only 15–25% of their basic salaries depending on performance level (Manzi et al., 2011).

Since 2002, a different evaluation has been in place rewarding teachers who show evidence of good levels of subject-matter and pedagogic knowledge. It is known as the Assignment for Pedagogical Excellence (AEP) and was inspired in the National Board for Professional Teaching Standards Certification in the United States (www.nbpts.org). Teachers in any publicly funded school (municipal and private subsidised) are eligible for the evaluation. Depending on results, a teacher receives a salary bonus lasting over a period of 10 years. Two instruments are used to gather evidence of teachers’ knowledge: a portfolio and a written subject-matter examination. As a result of the assessment teachers may be categorised as ‘outstanding’, ‘competent’ or ‘sufficient’ in their knowledge performance. Over time the number of teachers actually taking the examination has been considerably lower than those who apply for it (no more than 30% and as low as 13%). As reported by teachers, this low examination rate is due to lack of time to prepare the required portfolio (Falck, Kluttig, & Riberi, 2014). Also over time, a larger proportion of younger teachers with one to 10 years of experience have been successful compared to teachers with more experience, mostly because younger teachers take the examination. Finally, the number of teachers from municipal schools who receive the certification has been consistently lower over time than those who belong to subsidised private schools (Figure 5).

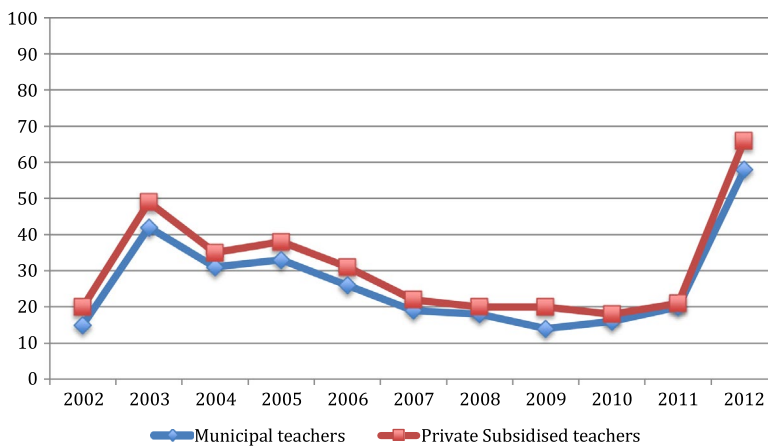


Figure 5. Percentage of teachers certified as being pedagogically excellent. Source: MIDE UC Statistics 2013.

The large increase of successful teachers in 2012 might be due to a change in the assessment procedures. Prior to 2012 the weighting given to content knowledge was low (30%) but in 2012, this weighting was raised to 50% with the other 50% given to pedagogic content knowledge thus lowering its contribution to the overall scores. This would be in line with greater teacher difficulty in adequately preparing portfolios, the main instrument assessing pedagogic content knowledge.

Strengths and shortcomings of the system

As referred to earlier the implementing university institution carried out a series of studies to assess the validity of the system and its instruments (Manzi et al., 2011; Rodríguez, Manzi, Peirano, González, & Bravo, 2014) with largely positive results. On the other hand, only a few independent studies have focused on the evaluation process itself and on teacher views about it (Avalos & Assael, 2006; Catalán & González, 2009). It was, thus, appropriate on the part of the Ministry of Education to participate in the OECD's *Review on Evaluation and Assessment Frameworks for Improving School Outcomes* programme (see: www.oecd.org/edu/evaluationpolicy) and invite an OECD team to review the teacher evaluation system in November 2011. The team examined all the documents related to the system and conducted a number of interviews with relevant stakeholders including practicing teachers (OECD, 2013). This Review together with other sources including teacher opinions gathered in survey data (Avalos, 2013) help to discuss the strengths and weaknesses of the Chilean evaluation system.

The Evaluation Framework

The Framework, which is the anchor that holds the evaluation system together and the key reference point throughout the years of teacher evaluation, is recognised as valid by teachers. It is similar in its content to those in place in other systems and in particular to the Praxis III evaluation developed by the Educational Testing Service in the USA (OECD, 2013). Nevertheless, it has shortcomings that have long been noted by observers and that were highlighted by the OECD review team. An important weakness is that the generic description of the Framework competencies is not followed by a specific description of the relevant performance levels for each competency. That is, there is no description of the kind of performance that would be typical for *each* competency level. This limits teachers' understanding of what are the specific strengths and weaknesses that feed into their evaluation categories, more so as they also are not given information on the rubrics used to assess them. Recognising these shortcomings of the Framework, the Ministry worked on a revised version, which was sent out for consultation to teachers' around the country. The new version includes the needed performance descriptions for the four assessment levels (outstanding, competent, sufficient and unsatisfactory) relating these in turn to the competences that make up the four key domains as illustrated below (Table 1):

Quality of the evidence provided by some of the instruments used

As noted earlier the instruments used for the assessment of teacher performance are not equal in what they offer and in how they contribute to the final evaluation of a teacher. In fact it is now being recognised that the teacher's self-assessment by its very nature cannot

Table 1. Modified framework for teacher assessment.

Domains:	Competences	Performance levels (for each criteria)			
		Unsatisfactory	Sufficient	Competent	Outstanding
Preparation for teaching	4	Description	Description	Description	Description
Classroom climate	3	Description	Description	Description	Description
Classroom teaching	4	Description	Description	Description	Description
Professional duties	4	Description	Description	Description	Description

provide a valid input to the overall evaluation of a teacher and should really not form part of the system. On the other hand, there is limited weight given to the school principal's assessment. This was a concession to the teachers' union mistrust of the fairness of principals' assessment given that an important number of them appointed by the military government were still in place in schools when the Teacher Evaluation system was agreed. This situation has now changed and will probably affect future decisions on the role of school principal assessment. The portfolios is clearly the most important instrument insofar as it provides evidence of teachers' classroom practice and their reflective capacities, as well as contribute substantially to the teachers' final assessment. Nevertheless, the OECD review team observed some unclear or weak elements in how the portfolio is configured and in how its different pieces contribute to the overall assessment of a teacher (OECD, 2013). They were also critical of the extremely precise instructions that teachers must follow to complete their portfolio (specific preparation of an eight-period unit and the filming of an exemplary lesson) as these left little space for assessing the range of teacher practices over a semester or a year. These narrow directives have caused teachers to be unnecessarily anxious about the process and mistrustful as to whether it really shows what they know and how they perform in their everyday work (OECD, 2013). The OECD team also manifested some concern for the lack of teacher live classroom observations considering that they serve to provide good evidence of how teachers perform. On the other hand, from the perspective of teachers working in highly prestigious public schools, the accountability emphasis of the education system (including evaluation) is meaningless and its formative aspect is not what drives their professional convictions as they are already committed to professional improvement in the traditional sense (Quaresma & Orellana, 2016)

Professional development opportunities

Teachers deemed to be 'unsatisfactory' or 'basic' are provided with relevant professional development opportunities in order to improve in the next evaluation. However, not much feedback from teachers is available on how useful these 'mandatory' opportunities have been. In a way, it could be inferred that given the low number of teachers who have actually been dismissed professional development has been effective, but there is no hard evidence about this. It could be argued that the concept of remedial professional development reinforces an instrumental perspective related to correcting deficits rather than promoting growth and diversity in teaching. Equally, the fact that well-evaluated teachers are rewarded with free professional development opportunities such as internships in other countries and options to take part in formal courses could be seen as favouring an 'accountability' concept of professional development rather than a formative one. The OECD team suggested that more should be done to provide conditions for teachers to improve through developing

communities of practice in their schools, assisted by a more formative role of the school leadership.

Teacher reaction to the evaluation system

Over time teachers' reactions to being evaluated have been mixed. On the positive side they have appreciated the fact that it was built *with* teacher participation rather than *for* teachers, as one of them stated during a group meeting:

Let's take the Teacher Evaluation: Beyond what anybody might legitimately think, it is a reform that involved the participation of the Teachers' Union. Institutionally, the union participated and held a relevant role ... obviously one can still question how it has been implemented in practice. ... (Avalos, 2013, p. 169)

However, teachers differ in the degree of importance they attach to the routine compulsory evaluation in relation to the two complementary 'reward-based' voluntary assessments known as AVDI and AEP. Asked in a national teacher survey how important they considered the different examinations to be, AVDI and AEP were rated as important by 52% of teachers compared to only 38% who rated the routine evaluation as important (Avalos, 2013).

Overtime it has been a given to hear teachers complain about the demands on their time of building the portfolio as required, as well as their misgivings about the objectiveness of their teaching performance ratings given its reliance on only one highly structured filmed lesson. These complaints were also voiced during interviews with the OECD team members:

We visited schools where some of the teachers had completed the portfolio whereas others had not; the ones who had not yet done it were enormously anxious as to whether they would be able to compile it accurately and completely. ...

Many teachers felt that completing the portfolio was far too time-consuming and that they were not given release time in school to complete it. (OECD, 2013, pp. 14–115)

Accountability policies and the teacher evaluation system

There is no doubt that the establishment of a teacher evaluation system responded to the growing importance of accountability concerns among policy-makers in the early 2000s prompted by student unsatisfactory learning results in national and international standardised tests. Opinions about teachers' responsibility in these results ranged from cautiously defending them and attributing their failure to unsatisfactory teacher initial preparation to blaming them overtly for not moving students to learn. It was, thus, felt that a teacher evaluation system could provide information on how professionally teachers were carrying out their work and reward those who were doing a good job, while conversely removing from the classrooms those who clearly were not. The teachers' union on the other hand tended to defend teachers as victims of their poor working conditions and of inadequate initial training and professional development opportunities, while also recognising that it was the union's responsibility to promote quality among their members and be open to observation and assessment (Avalos & Assael, 2006).

The resulting teacher evaluation system reflects these crossed views. Thus, it has a summative focus expressed in the use of narrow quantitative indicators and insufficient attention given to the range of possibilities of how teachers may conduct their teaching practices.

But it is also true that despite some pressures on the developers of the system as well as on those who approved it, the temptation has been resisted to use student results as a reference point for deciding on teacher performance.

The OECD review team also noted tensions between a formative and a summative intention, both in the nature of the instruments used and in the processing of the information. In their view, the system does not fully recognise the range of teaching practices in which teachers engage, due to the structured nature of the evidence requested in the portfolio; but also on account of the limited feedback given to teachers on their strengths and weaknesses. Equally, the system does not generate space for professional dialogue nor foster to a sufficient degree teacher reflection on their teaching. In this sense, the conclusion is that the system represents ‘a missed opportunity for strengthening professional development’ (OECD, 2013, p. 158).

From a different theoretical perspective Fardella (2012) also reaches similar conclusions about the orientation of the teacher evaluation system. Using discourse analysis Fardella (2012), retrieves from the precepts and codes of the written documents a view of teaching, which, though not coercive, in fact excludes other views about how teaching can be carried out, this being embedded in evaluation instruments. In other words, the written documents and precepts on how to complete instruments favour a certain view of teaching while making others ‘invisible’.

Conclusion: the challenges ahead

The condition of teachers especially their low salary levels and workload reflected in a much higher ratio of teaching/non-teaching hours (70/30) as compared to conditions in countries participating in the TALIS 2013 survey (OECD, 2014), the lack of an appropriate career structure and of sufficient quality of teacher initial preparation were all factors leading the Chilean government to the enactment in early 2016 of a new law for teachers, appropriately named the *Teacher Professional Development Law* (Ministerio de Educación de Chile, 2016). The Law deals with all these issues and poses challenges to the teacher evaluation system, which in turn will have to be reviewed and substantially modified in some of its aspects.

The key element of the new law is a career ladder structure for teachers, which provides the umbrella for improved salary and working conditions as well as requirements for initial teacher and continuing professional development opportunities. The career system comprises three compulsory stages and two voluntary ones. The compulsory stages are labelled as ‘initial’ for beginning teachers, ‘early’ for teachers with four to eight years of service and ‘advanced’ for teachers with more than eight years. Moving from one stage to another will require changes in the performance evaluation system. Thus, teachers moving from the ‘initial’ to the ‘early’ stage of their career will be assessed on subject-matter knowledge as well as on a portfolio assessment. No further subject-matter test will be required of teachers who are declared ‘competent’ or ‘outstanding’ in this test. All other evaluations will be mainly based on portfolio assessment. Portfolio assessment will now include information on collaborative work in schools and out-of-classroom responsibilities besides the usual teaching ones assessed on the basis of the Evaluation Framework. Teachers who are declared ‘competent’ or ‘outstanding’ in two consecutive portfolio assessments will not require further evaluations. The two voluntary stages of the career will offer opportunities for good performing teachers to move ahead, taking on diverse responsibilities along with less teaching

load and salary incentives. It is interesting that these two stages have been labelled as Expert 1 and Expert 2 in line, perhaps with the concept of expertise that lies behind assessments such as the Highly Accomplished Teacher evaluations in the US, Australia as well as England (Goodwyn, 2017), although not essentially inspired in those models.

Besides these provisions, the Law establishes the right of beginning teachers to be supported in their first year of teaching by mentors who will be prepared for the task. This is new for the Chilean education system. The career system will also cover teachers in private subsidised schools, meaning that they will also be subject to performance evaluation and hold the right to professional development opportunities.

These changes in the teacher legislation in many ways take into consideration observations made by the OECD team to the evaluation system and the legal context in which it has functioned. It is the case of the new career system, its stages and links to the performance evaluation of teachers, all of which were recommended by the OECD. The fact that the portfolio will become a key source for performance assessment and career advancement recognises the role it played in the original evaluation system by comparison to the other sources of evidence. More pertinent, however, is the Law's recognition of the importance of school-based teacher collaborative activities over their improvement and learning and its role as a source of information for the portfolio. In this, the law reflects international research on teacher collaboration (Vangrieken, Dochy, Raes, & Kyndt, 2015) and the fact that collaboration is now an important teacher feature in the TALIS studies (OECD, 2014). Also important is the embedment of other knowledge-related assessments such as the AVDI and AEP in the new system. It does, however, maintain traces of the emphasis on accountability in its merging of a career ladder system with compulsory evaluation, rather than separating the two purposes as originally proposed by the Teachers' Union in the 1990s, in terms of a routine school-based or other external evaluation, and a voluntary assessment linked to career progress.

All in all, there will be important challenges for those responsible of framing and implementing the new teacher performance evaluation system. These include further development of instruments such as the portfolio so as to adequately reflect the career teacher stages, including the mode of valuing school-based teacher collaboration for learning. More importantly, in framing the new structure and contents of the Portfolio it will be necessary to resist the temptation of defining these so specifically that they cease to recognise the variety of ways in which teachers carry out their teaching duties. In this respect, rather than externally prescribing what evidence should be included in the portfolio, it would be preferable to request teachers to build their portfolios over each year's work, selecting the pieces of evidence they consider relevant to the standards of the Framework.

As noted above there continue to be pressures to highlight the accountability aspects of teacher evaluation reflected in the importance given to testing subject-matter knowledge (derived from continued mistrust of the quality of initial teacher education), and which now becomes part of the mainstream evaluation.² To lessen these 'summative' effects it will be necessary to take on the task of preparing improved subject-matter tests that reflect, in the quality of their items, both content and pedagogic content knowledge, and so in ways that allow for application more than recalling of facts.

All in all, the Chilean teacher performance evaluation system has a solid experience from which to build an improved operation for the challenging demands of the new teacher career system. It will be important, therefore, that the tensions between the summative

and accountability purposes and those formative ones that, though diminished, are still present in the Law be lessened through solid dialogue with teachers and between teachers and managers of the educational system.

Notes

1. For details on the process leading to the agreement of a teacher evaluation system see Avalos, 2004; Avalos and Assael (2006).
2. It should be noted that to an extent the concerns about the quality of initial teacher education would be met, according to the Law, by a new content-based external examination that future teachers will take in the year prior to completing their preparation.

Disclosure statement

No potential conflict of interest was reported by the author.

Funding

This work was supported by Basal Funds for Centres of Excellence, Project FB 0003 from the Associative Research Program of CONICYT, Chile.

Notes on contributor

Beatrice Avalos-Bevan is currently associate researcher at the Centre for Advanced Research in Education, University of Chile. She has a long career in research and teaching on teachers, teacher policy and teacher education in Chilean and international contexts and has publications relating to the teaching profession in Chile as well as to other Latin American countries, as well as contributions to international reports on teachers, teaching and educational policy.

ORCID

Beatrice Avalos-Bevan  <http://orcid.org/0000-0001-6269-291X>

References

- Avalos, B. (2004). Regulatory forces and accountability processes in Chile. *Research Papers in Education*, 19(1), 67–85.
- Avalos, B. (2010). Educational change in Chile: Reform or improvements. In A. Hargreaves, A. Lieberman, M. Fullan, & D. Hopkins (Eds.), *Second international handbook of educational change* (pp. 383–396). Dordrecht: Springer.
- Avalos, B. (Ed.). (2013). *Héroes o Villanos? La Profesión Docente en Chile* [Heroes or Villains? The Chilean Teaching Profession]. Santiago: Editorial Universitaria.
- Avalos, B., & Assael, J. (2006). Moving from resistance to agreement: The case of Chilean teachers performance evaluation. *International Journal of Educational Research*, 45(4–5), 254–266.
- Catalán, J., & González, M. (2009). Attitude towards the evaluation of teaching performance and its relationship to one's own self-performance in teachers of elementary schools in Copiapó. *La Serena and Coquimbo. Psyche*, 18(2), 97–112.
- Cox, C. (Ed.). (2003). *Políticas Educativas en el Cambio de Siglo. La Reforma del Sistema Escolar de Chile* [Educational Policies at the Turn of the Century. Reform of the Chilean School System]. Santiago, Chile: Editorial Universitaria.

- Danielson, C. (1996). *Enhancing professional practice: A framework for teaching*. Alexandria, VA: ASCD.
- Danielson, C. (2010). *Competencias docentes: Desarrollo, apoyo y evaluación* [Teaching competences: Development, support and evaluation]. Unpublished document for PREAL.
- Danielson, C. (2011). Evaluations that help teachers learn. *Educational Leadership*, 68(4), 35–39.
- Falck, D., Kluttig, M., & Riberi, V. (2014). Resultados generales 2002–2012 [General results 2002–2012]. In B. Rodríguez, J. Manzi, C. Peirano, González, & D. Bravo. (Eds.), *Reconociendo el Mérito Docente: Programa de Asignación de Excelencia Pedagógica 2002–2014* [Recognising Teacher Merit. Voluntary Individual Grant Assignment Programme 2002–2014]. Santiago: Centro UC Medición – MIDE.
- Fardella, C. (2012). Truths about teaching, effects and subjective consequences of teachers' evaluation in Chile. *Revista de Psicología*, 21(1), 209–227. <http://www.avdi.mineduc.cl/Home/resultados>
- Goodwin, B., & Miller, K. (2012). What research says / use caution with value-added measures. *Educational Leadership*, 70(3), 80–81.
- Goodwyn, A. (2017). *Expert teachers. An international perspective*. London: Routledge.
- Manzi, J., González, R., & Sun, Y. (2011). *La Evaluación Docente en Chile* [Teacher Evaluation in Chile]. Santiago: Centro UC Medición - MIDE. Retrieved September 2, 2016, http://www.mideuc.cl/libroed/pdf/La_Evaluacion_Docente_en_Chile.pdf
- Marzano, R. J. (2012). The two purposes of teacher evaluation. *Educational Leadership*, 70(3), 14–19.
- Millman, J., & Darling-Hammond, L. (Eds.). (1990). *The new handbook of teacher evaluation: Assessing elementary and secondary school teachers*. Newbury Park, CA: Corwin Press.
- Ministerio de Educación de Chile. (2000). *Estándares para la Formación Inicial Docente* [Initial Teacher Education Standards]. Santiago: MINEDUC.
- Ministerio de Educación de Chile. (2003). *Marco para la Buena Enseñanza* [Framework for Good Teaching]. Santiago: MINEDUC.
- Ministerio de Educación de Chile. (2016). *Ley 20.903 El Sistema de Desarrollo Profesional Docente* [Law 20.903 The System of Teacher Professional Development]. Santiago: Chile. <https://www.leychile.cl/N?i=1087343&f=2016-04-01&p=>
- OECD. (2013). *OECD reviews of evaluation and assessment in education: Teacher evaluation in Chile*. Paris: OCDE.
- OECD. (2014). *TALIS 2013 results: An international perspective on teaching and learning*. Paris: TALIS, OECD Publishing.
- Papay, J. N. (2012). Refocusing the debate: Assessing the purposes and tools of teacher evaluation. *Harvard Educational Review*, 82(1), 123–141.
- Quaresma, M. L., & Orellana, V. (2016). El accountability y su impacto en la labor docente: Percepciones de los profesores de liceos públicos de alto rendimiento académico de Chile [Accountability and its impact on teachers' work: Perceptions of teachers in high achieving Chilean public schools]. *Curriculo sem Fronteiras*, 26(2), 316–338.
- Rodríguez, B., Manzi, J., Peirano, C., González, R., & Bravo, D. (2014). *Reconociendo el Mérito Docente. Programa de Asignación de Excelencia Pedagógica 2002–2014* [Recognising teacher merit: The Individual Grant Assignment Programme 2002–2014]. Santiago: Centro UC Medición-MIDE.
- Taut, S., Santelices, V., & Manzi, J. (2011). Estudios de validez de la Evaluación Docente [Studies on the validity of the teacher evaluation system]. In J. Manzi, R. González, & Y. Sun. (Eds.), *La Evaluación Docente en Chile* [The Chilean Teacher Evaluation System] (pp. 157–175). Santiago: Centro UC Medición - MIDE.
- Vangrieken, L., Dochy, F., Raes, E., & Kyndt, E. (2015). Teacher collaboration: A systematic review. *Educational Research Review*, 15, 17–40.