Voice and voicelessness in the construction of assessment policies: Participation as a relevant dimension in the potential impact on teaching and learning[1]

María Teresa Flórez Petour

Department of Pedagogical Studies, Faculty of Philosophy and Humanities, University of Chile, Santiago, Chile
mtflorez@uchile.cl
https://orcid.org/0000-0003-3704-726X

Abstract

This article critically analyses the currently predominant processes of construction of large-scale assessment policies by applying Bowe et al.'s (1992) *policy cycle* framework in the analysis of the national curriculum assessment system in Chile. Based on qualitative evidence that includes media and policy texts analysis, ethnographic work, and interviews with key policy actors, it aims at illustrating the disparity of participation in policy design as a potential reason for the lack of impact of these policies on teaching and learning. The findings point to a need for reconceptualising our framework for the understanding of the processes of construction of large-scale assessment policies by overcoming the rationale of market-oriented accountability systems that predominate nowadays. Alternatives are explored through examples of large-scale assessment systems with a higher parity of participation of stakeholders in policy design and policy enactment.

Keywords: participation, large-scale assessment, accountability, policy

The conceptualisation of large-scale assessment policies in market-based accountability systems

The expansion of market-based and result-oriented accountability systems in education, in which large-scale standardised testing plays a significant role, has been widely documented in recent decades. Authors like Falabella and de la Vega (2016) have systematised this global trend and its main components, which they described as characterised by:

- A focus on the concept of quality as measured by standardised assessment systems with an aim of establishing hierarchies and comparisons in the context of a market-

---

oriented education system. This focus also involves an assumption of these systems as being able to generate an improvement in the quality and equity of teaching and learning.

- The responsibility for results and their improvement as being situated in school communities—thus sidelining broader social, cultural, and economic factors that might have an influence in the quality and equity of education.

- Competition and comparison as the main drivers of the education systems—under the assumption that increased competition through school choice has the potential of improving the quality of education.

- A managerial perspective in which results should be the basis on which informed decision-making is made, with a high reliance on assessment instruments as valid evidence of quality and as a good means to inform good teaching and learning. In this context, the links between the reports of results and the processes of pedagogical reflection that should derive from them are seldom explicit.

- A series of incentives and consequences that are connected to results, and which operate under the assumption that they will motivate schools towards improvement—thus sidelining other internal motivations of school communities and reducing the functioning of the system to extrinsic motivation through external control.

This summary of the characteristics that Falabella and de la Vega (2016) attributed to these systems reveals the way in which a connection is assumed between large-scale assessment systems and a positive impact on teaching and learning and, therefore, on the quality of education as a whole. Research evidence, however, has been consistent in highlighting that the effects of these systems in the context of market-oriented accountability policies are quite the contrary. Negative effects such as teaching to the test, narrowing of the curriculum, teacher burnout and stress for school communities, discrimination and gaming practices in institutions, stigmatisation of schools that serve more socioeconomically disadvantaged students, the lack of motivation towards pedagogical innovation and the search for success formulas, and the feelings of frustration and demotivation as a consequence of low results have been widely documented in recent decades (Berryhill et al., 2009; Darling-Hammond & Rustique-Forrester, 2005; Falabella 2016a, 2016b, 2020; Falabella & de la Vega, 2016; Shepard, 1992; UNESCO, 2017, among many others). In addition to this, Falabella and de la Vega (2016) concluded that research evidence about the effect of these assessment systems on the improvement of learning results is not conclusive and, therefore, the assumption of accountability mechanisms as a driver for increased learning is called into question. In the case of Chile, a context that is seen as an extreme example of the implementation of these mechanisms, policy reports tend to place in school actors the responsibility for these consequences, and for the lack of an informed use of results (Comisión Sistema de Medición de la Calidad de la Educación, 2003; Equipo de Tarea, 2015)—arguing for the need for increased support and capacity building in these communities. This perspective leaves as blind spots the need to revise the design of policy, as well as the pedagogical relevance of the instruments used by policy as a potential reason for the lack of positive impact on teaching and learning.

Authors like Bowe et al. (1992) have provided a framework on how policy is conceptualised in neoliberal and market-oriented education systems, which could explain why the assumptions of result-oriented accountability policies are not met. According to these authors, with the spread of managerial logics in education, a linear and top-down understanding of policy construction became predominant (Bowe et al., 1992). From this perspective, the stages of policy design and policy implementation are seen as separate spaces, where the former is seen as the realm of "experts" who develop guidelines for action, and the latter is seen as the responsibility of school actors who are conceptualised as mere executors of someone else's design (Bowe et al., 1992). When policy is understood from this logic, schools become an easy target in the responsibilisation for the failure of policy, leaving the design made by experts unscathed. This could also be among the reasons why schools experience these policies as external, imposed, and scarcely relevant to pedagogical practice, with the lack of use of results as a consequence (Flórez et al., 2018).

Given the predominance of market-based accountability systems, it is not strange to forget that other approaches to accountability are possible, such as bureaucratic and professional accountability (Falabella & de la Vega, 2016). Authors like Onora O'Neil (2013) and Pasi Sahlberg (2010) have argued for the need to question current approaches to accountability and to promote new ways of understanding collective responsibility in education. These authors advocate for the development of what they call *intelligent accountability*, where principles such as trust and mutual responsibility replace the top-down and control-oriented approaches that currently dominate policy (O'Neil, 2013; Sahlberg, 2010). They have also criticised the narrowness of the information that is collected and provided through large-scale assessment systems, given that it does not point at relevant and contextualised learning and it does not benefit pedagogical practice from a formative perspective (O'Neil, 2013; Sahlberg, 2010). Quite the contrary, the consequences attached to these systems generate an impoverishment of pedagogical practice due to its reduction to what external tests promote as valuable learning.

Long-standing critical voices in the field of assessment policy have provided evidence about the disconnection between test-based accountability systems and the enhancement of pedagogy—despite this being an assumption in these systems. This article aims at questioning current modes of understanding the construction of assessment policies and situates this understanding among the reasons for the lack of impact on teaching and learning. It assumes a more complex, dynamic, multi-voiced, and contested conceptualisation of these processes and applies it to the critical analysis of the national curriculum assessment system in Chile as a policy case in order to illustrate how lack of parity of participation in policy design might be one of the reasons why market-based accountability policies are unable to fulfil their promises.

Theoretical framework

In order to gain a more complex understanding of policy construction processes, this study draws on the tradition of critical policy scholarship. This involves going beyond the strand of

research for policy, which is limited to provide guidelines for policy improvement, and assuming an approach that considers a critical analysis of how policy is constructed in the context of a struggle for meaning in its definition. In particular, the *policy cycle* model by Bowe et al. (1992) was considered appropriate for the kind of problems that were addressed by this research. This approach seeks to overcome linear conceptualisations of policy as something that "'gets done' to people by a chain of implementors whose roles are clearly defined by legislation" (Bowe et al., 1992, p. 7) and where the dissociation between those who generate policy and those who are expected to implement it is naturalised. Exclusion of the latter from policy design is taken for granted in these conceptualisations. As Bowe et al. (1992) stated, the governmental position is assumed as unequivocal and problems are, thus, seen as errors of implementation. Bowe et al. (1992) instead developed a dynamic understanding of policy as a process characterised by power struggles in which meanings are negotiated, contested, reinterpreted, and translated through a continual recontextualisation of policy texts. This approach involves, on the one hand, an understanding of *policy as text* in terms of the diverse contexts where it is used and the procedures through which different actors get involved in what the authors saw as "continual political struggles over access to the policy process" (Bowe et al., 1992, p. 10). From this perspective, policy texts are understood as open spaces where contradictions can be traced as "the outcome of struggle and compromise" (Bowe et al., p. 21). On the other hand, this also involves understanding *policy as discourse*, that is, as a series of claims about the values through which education should be understood and taken into practice, with multiple effects and possibilities for reinterpretation across the policy process (Bowe et al., 1992). Later developments by Stephen Ball and his colleagues have added new layers to the understanding of policy processes and these were also considered relevant in the conceptualisation of the study.

In particular, the policy cycle model suggests considering three main contexts, namely, the context of influence, the context of policy text production, and the context of practice, which are not seen as rational stages of a linear process (Ball, 1993) but, rather, as spaces that overlap and interact with each other in the struggle for meaning. The *context of influence* is understood as the domain where the definition of the politics of policies is disputed by antagonist actors. In the words of Bowe et al., "it is here that interested parties struggle to influence the definition and social purposes of education, what it means to be educated" (1992, p. 19). The context of influence demands an answer to questions around those ideologies, power relations, and interests that take place within a broader sociopolitical context—aspects that frame who can speak, and what can be said about policy.

The *context of policy text production* refers to the space in which ideologies, values, and hegemonic discourses disputing the context of influence are represented in texts, usually "in the language of the general public good," as an attempt to control the meaning of policy (Bowe et al., p. 20). It entails the process in which policy is encoded in the form of official documents, instructions, public speeches, and press communications that seek to guide the roles, goals, and practices that the different policy subjects/agents should follow. Policy text production is not coherent, and the resulting texts tend to express diverse tensions, contradictions, and discontinuities. Policy texts, indeed, often result in a fabric "heteroglossic

in character," which interweaves different interests "to achieve apparent consensus and legitimacy" (Lingard & Ozga, 2007, p. 2).

Finally, the *context of practice* is the main realm to which policy text production is addressed, and specific responses are expected from its actors and institutions. But policy texts involve both constraints and possibilities and, given that practitioners "come with histories, with experience, with values and purposes of their own," their responses are never passive because they always entail interpretation and recreation (Bowe et al., p. 22). Hence, questions regarding this dimension focus on how recontextualisation takes place and is disputed by the different actors involved in the context of practice, and the ways in which policy shapes subjectivities, values, and power relations at a micro level. These processes of recreation, interpretation, translation, and resistance are understood in later developments by Ball et al. (2012) as *policy enactment*.

This framework has the potential to offer a more complex and dynamic approach to the understanding of market-based accountability policies that strongly rely on external assessment systems. From this perspective, the lack of impact of these systems on teaching and learning is not merely seen as a failure of practitioners who have not been able to implement policy as intended by policy makers but, rather, as the result of a way of constructing assessment policies that has systematically excluded practitioners from the formulation of policy. This exclusion, as discussed later in this paper, leads to a perception of these policies as distant and imposed, with the lack of use of its results in teaching and learning as a natural consequence.

Methodological design

The Sistema de Medición de la Calidad de la Educación [System for the Measurement of Quality in Education], known as SIMCE, was selected as a case for this study. This system was created in 1988 as a means to measure learning around the curriculum at a national level in Chile. It was legally instituted in 1990 by the Organic Law for Education promulgated before the end of Pinochet's dictatorship, with the specific aim of informing the school market with an indicator of quality that facilitated parental choice and competition for student enrolment between schools, along with the purpose of informing policy through its results. It was administered by the Ministry of Education from 1992 and, after 2012, its administration was transferred to the Agency for Quality in Education, a governmental agency independent from the Ministry. It consists of a battery of tests in language and mathematics (and history and science, in some levels) where items predominantly correspond to multiple-choice questions. Its purposes have accumulated in time and its frequency has also grown. More importantly, its consequences have increased, with an impact that, in the present, includes not only the public image of the school through open publication of results but also the salaries of teachers, a part of schools' financial scheme, and the potential closure of educational institutions with repeated low results (Flórez, 2018). All this turns SIMCE into a high-stakes assessment system connected to a wide variety of policies, and also into a crucial piece in the operation of the market-oriented education system in Chile. This has led to increased

criticism towards the policy among different actors; demands for its reform have emerged throughout the different waves of the social movement for education in 2006 and 2011, with support from different actors (critical researchers, students, teachers), culminating in a call from students to boycott the test in 2013. Criticism continues in the present and an intense debate is now taking place in Chile around the continuity of SIMCE. It is, therefore, an interesting case in the exploration of large-scale assessment policies given that it is situated in a country that is known as a paradigmatic case in the implementation of market-oriented education policies (Verger et al., 2016).

With the policy cycle approach as a conceptual framework, this study aimed at exploring the following research questions:

- What are the main actors and their interactions in the definition of policy in SIMCE and the discourses they sustain in connection to the test?
- Who speaks and who is spoken for or (re)presented in policy texts in SIMCE?
- What processes of interpretation, translation, and resistance occur in school communities when enacting the SIMCE policy?
- How is the policy around SIMCE constructed and produced, considering the network of actors involved in the whole process, their interactions, and the discourses that circulate between them?

To respond to the complex nature of these questions, a multidisciplinary approach guided the study, facilitated by a team whose composition involved multiple fields of expertise (media analysis, discourse analysis, sociology of education, and ethnography). Using the varied expertise of the team to enrich our understanding of policy, multiple sources were considered, including media texts, policy documents, interviews with key actors, and ethnographic data collection. The strategies for data collection and analysis are described below.

Media texts

The selection in this case was restricted to editorials and news in written press, specifically in two major circulation newspapers in Chile (*El Mercurio* and *La Tercera*). A first round of selection included all the editorials and news that were published in connection to SIMCE in 2014, which were analysed using critical discourse analysis (CDA). In order to broaden our findings in terms of the historical trajectory of SIMCE, a second phase of the project involved the selection of the editorials and news that were published in both newspapers during the years 2000, 2003, 2008, and 2010, specifically in those months when results were released and when the test was taken. These years were selected because they constituted crucial points in the development of SIMCE, according to our previous knowledge about its historical trajectory. The second sample was first analysed using the concept of *framing* in media studies (Entman, 2007), which allowed for a general mapping of actors and sources and for a general view of the problems and solutions attached to the (re)presentation of SIMCE in media. These news and editorials were later analysed using CDA, keeping some of the categories used in the 2014 sample, in order to facilitate some comparisons.

Policy texts and interviews with key actors

A sample of 272 policy texts (33 texts from experts, 59 legal documents, 180 policy-user-oriented documents) was selected considering different periods in the historical trajectory of SIMCE (1988–1995, 1996–2000, 2001–2005, 2006–2010, 2011–2015). An initial scanning and analysis of these documents was carried out. On the basis of this first round of analysis, seven documents were selected for detailed CDA, considering the degree to which they represented the tensions that characterised each period. Simultaneously, the initial scanning and analysis served as a basis to select a group of 12 actors who held key roles in the trajectory of construction of policy texts around SIMCE. Some of these actors were also the authors of the policy texts that were selected for detailed analysis. Participants included government officials, heads of technical units for the development of the test, and two members of the Teachers Union who had been actively involved in public discussions around this policy. Semi-structured interviews were carried out with these actors with the aim of enriching the information provided by documents.

Ethnographic data

A qualitative ethnographic methodological approach was used to understand how schools translated and interpreted the SIMCE policy. This involved sustained presence in schools with a special emphasis on daily social relations (Geertz, 2005; Rockwell, 2009). Cases were selected considering three main criteria: school categorisation, geographic location, willingness to participate in the study. Two schools categorised as "emergent,"[2] which were previously categorised as "recovering," were selected (Sur School and Rosa School). A third school categorised as "autonomous,"[3] and from the same geographic area was also included in the study (Novo School). Data collection focused on the period when SIMCE was taken, and involved observation registers, interviews, as well as more informal observations. For ethical reasons, the names of the schools correspond to pseudonyms to maintain anonymity.

During and after the process of data collection and analysis, the interdisciplinary team held extensive meetings both to increase consistency in the process of analysis (when multiple analysts were working with a single set of data) and to discuss emerging findings and trends or patterns in data. These meetings allowed for an increasingly integrated approach in the reconstruction of the process around the SIMCE policy, where the three contexts were portrayed as interacting in a single dynamic whole.

---

2   Since the promulgation of Law 20.248 in 2008, which establishes a system for preferential subvention for schools who receive higher levels of vulnerable students, Chilean state schools are categorised in three types according to their results in SIMCE and other indicators—although SIMCE constitutes the main means of categorisation (70% against 30% of other indicators). In this context, "recovering" schools are those who have had repeated low results in these tests, "emergent" schools are the ones who have not shown consistently increasing results over time, and "autonomous" schools are those whose results have improved systematically. The consequences of this categorisation are related to a higher or lower level of supervision and intervention according to results.

3   See Footnote 2.

Voice and voicelessness in the construction of assessment policies: Why would an impact on teaching and learning be expected?

As can be imagined, findings from a study of this magnitude are multiple and open different avenues for interpretation. In this paper, only those findings that are relevant to the topic under scrutiny are addressed, namely:

- Whose voice predominates in the policy landscape and whose voice is spoken for.

- The discursive strategies that are deployed in public discourse in order to exclude or silence some voices who hold differing views about the use and meaningfulness of SIMCE in practice.

- The way in which actors see each other as a source of distance that hinders mutual understanding and the inclusion of different groups of interest in policy design.

These findings are presented through examples that are illustrative of features that emerged as relevant patterns in the process of analysis.

With regards to the voices that predominate in the policy landscape, the analysis of media and policy texts showed similar patterns. Part of the analysis involved categorising the types of actors that appeared in these texts and also distinguishing between those who were (re)presented as mere actors and those who were given direct voice through quotations. As Figure 1 illustrates, in media texts "Executive power and bureaucracy" and "Experts" were the voices with the highest number of direct quotations; members of school communities showed a high presence as actors and were mainly presented as receivers of policy actions and as a sort of background where they were spoken through predominant voices. Direct quotations for these actors were lower compared to authorities and experts, and they appeared in the material predominantly in the shape of testimonies of teachers, head teachers, and students who, despite their socioeconomically disadvantaged context, were able to perform well in the test. Some examples of this (re)presentation of schools can be found in headlines such as "The case of the poorest school with the best score" (Pavez, 2008), "The strategy of children in Melipilla who overcame the limits of SIMCE" (Gutiérrez, 2003), "Vulnerable schools narrate new methods: Strategies to triumph in SIMCE" (Dalgalarrando, 2008), "Exceptions in SIMCE: New strategies demolish the myth of the socioeconomic factor. Private [consultants] raise performance of poor schools" (Zúñiga & Aravena, 2003). In that sense, the direct voice of school communities is used to legitimise the discourse sustained by authorities and experts.

Figure 1
Total of frequency actors and sources 2000–2014
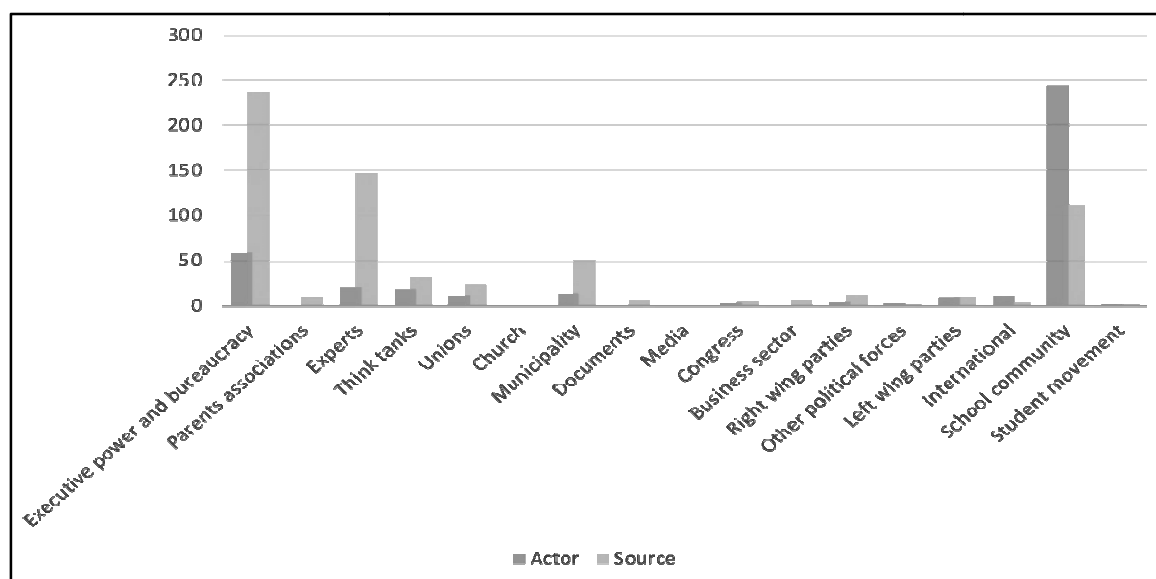
Consistent with the results from media analysis, "Executive power and bureaucracy" emerged as one of the predominant sources in the analysis of policy texts. Additionally, other relevant sources that emerged from this analysis were other policy "Documents," "Experts," and "International organisations," with "Think tanks" and "Private foundations" acquiring a more prominent presence after 2003. The direct voice of actors from school communities is completely absent from these documents, appearing again as a background and as the passive receivers of policy. Some documents produced by expert committees whose work involved audiences with school actors present a sort of simulation of their direct voices by grouping their perspectives in a single whole, thus generating a consensus effect where they all appear to agree. The excerpt below illustrates this discursive strategy:

> Summary of audiences with head teachers, teachers and school administrators: On the basis of the audiences it is possible to identify positive and negative elements in connection to SIMCE from the perspective of schools and administrators.

> Positive aspects: [SIMCE] is considered a technically solid measurement system, whose results can be relied upon. Practically all the actors invited to the audiences manifested that they believed that a national assessment of learning is necessary, particularly to monitor achievement and to generate support when weaknesses are identified. (Author's translation of Equipo de Tarea, 2015, p. 4)

This quotation illustrates the way in which potential differences between these actors are hidden under an appearance of consensus. No differences in trends were detected between the perspectives of head teachers and administrators compared to those of teachers. Similarly, in a highly segregated education system like Chile, it is hard to believe that the views of actors from state-administered schools would be similar to those of privately owned schools. This difference was not made explicit in the text. Our analysis of rhetorical elements of discourse in the context of CDA also detected a fallacious construction of the argument in this example

in that the necessity and relevance of having a national assessment system is artificially connected to the need for maintaining SIMCE—a system that is criticised by actors from school communities, critical research, and social movements. This simulation of direct voice, therefore, shows the way in which the perspectives of actors outside the elite of policymakers and experts are appropriated in order to support and legitimise predominant voices. This is reinforced in our analysis of policy texts by the detection of a higher level of individualisation and nomination for actors from executive power and bureaucracy and experts—in contrast to the presentation of actors from school communities mainly grouped in categories related to their role in the system. The effect is an invisibilisation of their experiences with SIMCE in public discourse, as illustrated bellow.

A second pattern in our findings is related to a series of discursive strategies commonly deployed in media and policy texts, all of which contribute to the silencing and exclusion of voices from the context of influence and the context of practice. These strategies are defined below:

- *Delegitimisation*: refers to the representation of opposing views in discourse as non-legitimate (van Dijk, 1999).
- *Affiliation*: involves those strategies in which actors do not speak for themselves but are rather spoken for through members or representatives of a group or institution that appropriates their voices (van Dijk, 1999).
- *Naturalisation/neutralisation/depoliticisation*: to present a specific policy/political/ideological context as a given, as naturally emerging and, therefore, as neutral and potentially unchangeable.
- *Positivisation*: aspects that are portrayed as negative by school communities and some researchers are presented in positive terms.
- *Silencing/exclusion*: the elision of a specific actor or its voice in the construction of the text.

While a myriad of examples of these strategies emerges from the study, only a small set of illustrative cases is offered here. First, a contrast of voices is presented in the excerpts below to show how positivisation and affiliation generate an invisibilisation of the experience of school communities. One of them comes from an expert who was the National Coordinator of SIMCE between 2008 and 2012 and who continues to hold high management positions in connection to this policy in the Agency for Quality in Education:

> [1] Additionally, results from the tests have allowed for feedback in relation to pedagogical and management practices in schools, through the distribution of reports for teachers and school managers, the organisation of seminars to disseminate results and days of analysis in schools. (Author's translation of document from expert also in charge of the unit that develops SIMCE, 2011)

The second set of quotations emerges from the ethnographic strand of the study, showing the way in which SIMCE is experienced in schools:

[2] I would say it is chaotic at some points. The stress is too much. It goes beyond you being supervised or not. It is personal; it is related to the goals they set for you, along with the reality of children. (Author's translation of Rosa School teacher)

[3] I am worried that my colleagues, because I retire in a couple of years, my new colleagues, the young ones, won't have a job, because you know that a school with bad results can even be closed. (Author's translation of Nova School teacher)

[4] Standardised instruments do not consider diversity . . . they do not include the context, they do not consider these children who learn in a different way (Author's translation of head teacher, Sur School)

The contrast is apparent between a positive view of the potential value of the test for improving pedagogical reflection and practice, repeatedly articulated in the policy documents that were analysed, and a distressing narrative of the experiences of teachers in connection to the test—and who also highlight the way in which diversity and context are sidelined to the detriment of students' learning needs; this is very distant from the pedagogical use of results that is intended in policy documents. The findings from the ethnographic work repeatedly show that schools participate in SIMCE mainly due to its consequences and not because they see the test as pedagogically useful. This is consistent with findings from research in the last decade about the effects of SIMCE in schools (see, for example, Acuña et al., 2014; Centro de Investigación y Desarrollo de la Educación, 2012; de la Vega & Picazo, 2016; Falabella 2016a, 2016b, 2020; Flórez, 2013; Ortiz, 2012; Rojas & Leyton, 2015, among many others), so this is far from being an isolated experience. In that sense, the idea of the potential impact of high-stakes standardised tests used for market accountability purposes on teaching and learning is more an assumption of experts and policymakers than a reality in schools. This interpretation is supported by research that reveals the lack of use and understanding of results in school communities (Taut et al., 2009), along with perceptions of the limited potential of the test in terms of feedback for pedagogical practice (Manzi et al., 2014). Policy documents, however, place the responsibility for this lack of use in the need for capacity building in schools in terms of learning how to use results (see, for example, Equipo de Tarea, 2015), without any consideration that the information provided by these assessment policies might be pedagogically irrelevant. This analysis illustrates how the strategies of affiliation, positivisation, and silencing/exclusion operate in policy and media texts. In Excerpt 1, the voice of actors from schools is spoken by the expert author to assimilate it to his views. The negative experiences of schools, presented in Excerpts 2, 3, and 4, are erased from public discourse (silencing/exclusion) and are presented from a positive narrative about the impact of SIMCE in their pedagogical practice (positivisation). In another policy document, the distressing narrative of teachers is positivised by using the term "urgencia por mejorar [urgency for improvement]" (Equipo de Tarea, 2015, p. 6) to describe the effect that the test generates in schools as a means of presenting negative experiences in a positive fashion.

Another example that illustrates the deployment of these strategies derives from the analysis of the ways in which some actors are portrayed throughout the trajectory of the SIMCE

policy. In the case of students, for example, before the series of social movements for education that began in 2006, media and policy texts tended to address them in a passive way; they were represented as recipients of actions related to SIMCE and its results or as illustrative examples of success from disadvantaged schools. However, since their attempt to boycott the test in 2013, students became agents in the policyscape by opposing and acting against SIMCE, and their representation in discourse changed accordingly. As the following examples from an editorial in a national circulation newspaper illustrate, their position began to be neutralised, affiliated, and delegitimised by those actors who held a higher degree of power over public discourse.

> Although there were very specific cases, the test boycott attempt is absurd, as it goes against themselves and their own schools. Although the instrument can be improved and complemented with other measurements, it is one of the few tools with which the education system counts to measure the level of learning of students and it is a device that allows to identify strengths and weaknesses, thus facilitating the correction of formative processes. Additionally, it is a mechanism that provides parents with useful information in the election of the school for their children; in the same vein, it is used as a factor in the provision of performance-related funding.

> Students must recognise that actions of this sort are counter-productive and reflect an inconsistency with their own demands, which point at the strengthening of school teaching. Additionally, criticisms towards SIMCE expressed by minority sectors of students have been favourably welcome by the authority. (Author's paraphrase of La Tercera, 2014)

A first mechanism of delegitimation in these excerpts is portraying the movement as a small group or a minority (very specific cases/expressed by minority sectors). A second delegitimation strategy apparent in the editorial referenced above is the use of negative terms to refer to the intentions of students (absurd/counter-productive), along with the naturalisation of the good effects they attribute to the test (facilitates correcting formative processes/provides parents with useful information/a factor in the provision of performance-related funding). The demands of students in relation to the negative effects of this policy and its connection to a market-oriented model of education, therefore, are silenced and criticism is reduced to aspects of the test that can be improved. Additionally, a strategy of affiliation is detected in the use the students' demand for 'the strengthening of school teaching' against themselves; if students want an education of good quality, then they should agree with SIMCE. A positivisation strategy is also found, for example, when the provision of performance-related funding is presented in a positive light—contrary to evidence from research and also from our own ethnographic data that refers to the perception of teachers and school leaders about the negative consequences of these mechanisms. These include effects such as the decontextualisation of pedagogical practice to the detriment of diversity in learning, teacher burnout, and stress for schools. Finally, depoliticisation is illustrated by the way in which policies that are part of a market-oriented neoliberal model of education, such

as competition through parental choice and performance-related funding, are presented as a given and as inherently positive.

It is important to add, with regards to students, that the way in which they were considered in the process of revision of SIMCE by a committee of experts in 2014 also reveals the use of delegitimation, affiliation and silencing/exclusion strategies. The executive report of the committee (Equipo de Tarea, 2015) portrays in its introduction the student movement as having a crucial role among the motivations for a critical review of SIMCE. In terms of their participation in this review process, however, not only their representatives are absent from the committee itself; student organisations are mentioned once in the report in a footnote that indicates that they were summoned to one of the meetings but they could not attend. This means that one of the main voices mentioned at the beginning of the document as motivating changes to SIMCE was eventually absent from this debate. Their role in the context of influence is, therefore, recognised only in terms of a phony inclusion that mentions them but then rules them out of the formulation of policy texts.

A third strand of findings that contributes to the focus of this article is related to the way in which actors see each other. The following contrast of quotations offers a clear example of this problem. The first illustrates the way in which teachers are seen by experts and policymakers in terms of their role in the construction of policy. The second excerpt is an example of how teachers portray those actors who are in charge of assessment policy development.

Interview with expert, policy-maker, and member of an education private foundation:

> Interviewer: What is the role that you think teachers should have in the construction of education policy?

> Interviewee: Well, the topic of teachers in Chile is quite complicated because teachers, especially primary teachers, don't know much, I mean, it is difficult that they can contribute in a very informed way to a well-designed policy. I would spend all the money in the world to teach teachers well but I know is very difficult.

Interview with former member of the Teachers Union studies area:

> I think there is a weight of the technical above life, then, with all the respect I have for [name of expert in charge of the initial design of SIMCE], she only had a technical perspective, real life was not there and in real life this technique didn't work and she realised this later.

The distance and lack of dialogue between these actors is apparent in these quotations. On the one hand, teachers are portrayed by experts and policymakers as lacking the necessary knowledge to be able to participate in policy formulation, thus excluding experiential professional knowledge as a legitimate and necessary source in the design of assessment policies. On the other hand, teachers see experts as people with valuable technical knowledge but also as distant from reality, which has as a consequence that technically sound solutions

do not work when they are taken into practice. This mutual misunderstanding could be part of the reasons why large-scale assessments and pedagogical enhancement seem to be disconnected. Both experiential and technical knowledge are potentially valuable for the design of successful assessment policies but they remain two separate realms under charge of different actors who do not interact with each other. This dissociation contributes to the development of assessment policies that are disconnected from practice, which nevertheless have to be implemented by teachers who are told what to do, even when these policies do not make sense to them. Nothing different from this can emerge when policy is conceptualised through the separation of the stages of design and implementation, as Bowe et al. (1992) criticised.

Concluding remarks: The need for a deep reconceptualisation of assessment policies and their relationship with practice

The findings of this study tend to confirm that current approaches to assessment policy design, where experts and authorities are in charge of thinking and developing policies that schools are expected to implement, play an important role in the disconnection between large-scale assessment systems and the improvement of teaching and learning. The separation between the stages of policy design and implementation, that Bowe et al. (1992) saw as characteristic of managerial perspectives and market-oriented education systems, relegates the formulation of policy to "idealizations of the 'real world'" (Bowe et al., 1992, p. 21) that are doomed to fail when their enactment is carried out in real and highly complex contexts like schools. In his research on the micro-politics of the school, Ball (2012) analysed how traditional organisational approaches, highly prevalent in current education policies, tend to impose an ideal model on how schools should work—a prescriptive attempt he understood as fundamentally ideological—instead of understanding school communities from an approach that accepts their complexity and their contextual diversity. This is the type of understanding that schools in our ethnographic evidence seem to be expecting from large-scale assessment policies.

The analysis of SIMCE through the policy cycle framework has not only allowed for a more complex, multi-voiced, and dynamic understanding of policy processes but has also highlighted the need to deeply rethink these processes in order to develop assessment policies that become beneficial for the education system.

The paper has illustrated how one of the aspects that requires a deep and urgent reconceptualisation is related to questions around who participates in policy and how this participation is enacted. Nancy Fraser's (2008) concept of *parity of participation* in her discussion around social justice might be of help in systematising the aspects that require transformation in the process of construction of large-scale assessment policies. According to Fraser (2008), a socially just society requires the establishment of social agreements where all interest groups participate as peers in the formulation of norms. This requires that

institutionalised obstacles for equality of participation are overcome, thus preventing the over-representation of some voices to the detriment of others whose legitimacy has not been justly recognised. The findings discussed here reveal how currently predominant processes of assessment policy formulation fail in the equal representation of voices. Both the definition of problems to be addressed and the means to do this are limited to the community of national and international experts, along with political authorities. In a very narrow concept of the expert, teachers' experiential and professional knowledge is excluded as a relevant source, with their participation not moving beyond consultation initiatives where their perspectives, as was illustrated in this article, are often appropriated to legitimise the assumptions of dominant groups of interest or are simply silenced (Flórez & Olave, 2020). From the perspective of policy makers, their knowledge is not seen as a legitimate source in the construction of assessment policies.

This lack of recognition in the processes of deliberation around assessment policies leads to the perception in school communities of these systems as imposed, decontextualised, and alien to the meanings they attribute to their work and, therefore, as meaningless to their practice (Flórez et al., 2018). These views, however, are scarcely represented in public discourse, where an over-representation of the voice of policy elites generates an illusion of consensus that legitimates the continuity of these measures despite criticisms. The paper has illustrated how this injustice of participation could be considered among the reasons why large-scale assessment systems are not meeting the assumptions of market-based accountability policies. Bowe et al. (1992) argued that the separation between the stages of design and implementation is predominant in managerial and market-oriented approaches. An increased parity of participation, therefore, requires moving away from these approaches and exploring new avenues for other conceptions of accountability that are able to construct genuine links between large-scale assessment systems and the improvement of learning.

The response from experts and authorities to this problem, however, is far from a reformulation of policy construction processes. Consistent with the design/implementation divide, they resort to the responsibilisation of teachers for their lack of capacity to give these results a relevant use. The lack of genuinely respectful dialogue and collaboration between these actors does not contribute to solve the problem. The question remains as to how we move from current modes of thinking assessment policies to a different approach, where parity of participation enables better and healthier connections between large-scale assessment systems and the enhancement of pedagogy.

An essential starting point for this transformation is to overcome the use of large-scale assessment systems for market accountability purposes and to prioritise, instead, formative purposes in a low-stakes context. This is the case, for example, in countries such as Portugal, Uruguay, and Denmark where large-scale assessment systems are explicitly low-stakes and are mainly focused on providing detailed feedback to schools (Luaces, 2010; Simoes & Pereira, 2017; Wandall, 2009). Parity of participation is also crucial and some good examples of co-constructed assessment policies where teachers' judgment is valued come from experiences such as the Externally Moderated School-Based Assessment in Queensland,

Australia (Maxwell & Cumming, 2011), the assessment system of the School Success Trajectories Project that was piloted at a municipal level in Valparaíso, Chile (Flórez & Olave, 2020), and the Ontario Secondary School Diploma (OSSD) that is based on consensus moderation processes carried out by teachers (Cuff, 2018). Along with more agency and participation from school communities, these systems also allow for an increased contextualisation and pedagogical pertinence of assessment activities because they are developed by schools based on shared criteria, a feature that enhances respect for diversity as well as the possibility of developing meaningful learning rather than focusing on training for a multiple-choice test. These are only a few examples of the direction that different countries have been following towards school-based assessment systems—Finland among them (Partnership for Educational Revitalization in the Americas, 2009)—in order to overcome the shortcomings of high-stakes testing regimes.

All these changes require, of course, political willingness. But an emerging horizon can be envisioned from the experiences mentioned above, characterised by equal and respectful participation of all groups of interest in assessment policy design and enactment, the development of low-stakes systems with formative purposes, and priority given to meaningful teaching and learning.

Acknowledgements

References

Acuña, F., Assaél, J., Contreras, P., & Peralta, B. (2014). La traducción de los discursos de la política educativa en la cotidianeidad de dos escuelas municipales chilenas: La metáfora médica como vía de análisis [The translation of educational policy discourses in the daily lives of two Chilean municipal schools: The medical metaphor as a means of analysis]. *Psicoperspectivas, 13*(1), 46–55. http://dx.doi.org/10.5027/psicoperspectivas-Vol13-Issue1-fulltext-363

Ball, S. J. (1993). What is policy? Texts, trajectories and toolboxes. *Discourse: Studies in the Cultural Politics of Education, 13*(2), 10–17. https://doi.org/10.1080/0159630930130203

Ball, S. J. (2012). *The micro-politics of the school. Towards a theory of school organization.* Routledge.

Ball, S. J., Maguire, M., & Braun, A. (2012). *How schools do policy*. Routledge.

Berryhill, J., Linney, J. A., & Fromewick, J. (2009). The effects of education accountability on teachers: Are policies too-stress provoking for their own good? *International Journal of Education Policy and Leadership*, *4*(5), 1–14. https://doi.org/10.22230/ijepl.2009v4n5a99

Bowe, R., Ball, S., & Gold, A. (1992). *Reforming education and changing schools: Case studies in policy sociology*. Routledge.

Centro de Investigación y Desarrollo de la Educació. (2012). *IX Encuesta a actores del sistema educativo 2012* [Ninth survey of stakeholders in the education system 2012]. https://educacion.uahurtado.cl/wpsite/wp-content/uploads/2017/03/Informe_IX_Encuesta_CIDE_2012.pdf

Comisión Sistema de Medición de la Calidad de la Educación. (2003). *Evaluación de Aprendizajes para una Educación de Calidad* [Assessment of learning for an education of quality]. Ministerio de Educación, República de Chile. https://bibliotecadigital.mineduc.cl/bitstream/handle/20.500.12365/17739/EV03-0009.pdf?sequence=1&isAllowed=y

Cuff, B. (2018). *International approaches to the moderation of non-examination assessments in secondary education*. Office of Qualifications and Examinations Regulation. http://dx.doi.org/10.13140/RG.2.2.17898.21442

Dalgalarrando, G. (2008, May 29). Colegios vulnerables narran nuevos métodos: Estrategias para triunfar en el SIMCE [Vulnerable schools narrate new methods: Strategies to triumph in SIMCE]. *El Mercurio,* Section C, p. 9.

Darling-Hammond, L., & Rustique-Forrester, E. (2005). The consequences of student testing for teaching and teacher quality. In J. Herman & E. Haertel (Eds.), *The uses and misuses of data in accountability testing: The 104th yearbook of the National Society for the Study of Education, Part* II (pp. 289–319). Blackwell.

de la Vega, L. F., & Picazo, M. I. (2016). La accountability de la calidad y equidad de una política educativa: El caso de la subvención escolar preferencial para la infancia vulnerable en Chile [Quality and equity accountability in educational policy: The case of preferential school subvention for vulnerable children in Chile]. *Revista Del CLAD Reforma y Democracia, 65*, 193–224. https://www.redalyc.org/journal/3575/357546620007/

Entman, R. (2007). Framing bias: Media in the distribution of power. *Journal of Communication, 57*(1), 163–173. https://doi.org/10.1111/j.1460-2466.2006.00336.x

Equipo de Tarea para la Revisión del Sistema Nacional de Evaluación de Aprendizajes (2015). *Informe ejecutivo* [Executive report]. http://archivos.agenciaeducacion.cl/documentos-web/Informe_Ejecutivo_Equipo_de_Tarea_Simce.pdf

Falabella, A. (2016a). ¿Qué aseguran las políticas de aseguramiento de la calidad? Un estudio de casos en distintos contextos escolares [What do quality assurance policies ensure? A case study in different school contexts]. *Revista Estudios Pedagógicos*, *41*, 107–126. https://doi.org/10.4067/S0718-07052016000100007

Falabella, A. (2016b). Do national test scores and quality labels trigger school self-assessment and accountability? A critical analysis in the Chilean context. *British Journal of Sociology of Education*, *37*(5), 743–760. https://doi.org/10.1080/01425692.2014.976698

Falabella, A. (2020). The ethics of competition : Accountability policy enactment in Chilean schools' everyday life. *Journal of Education Policy*, *35*(1), 1–23. https://doi.org/10.1080/02680939.2019.1635272

Falabella, A., & de la Vega, L. F. (2016). Políticas de responsabilización por desempeño escolar: Un debate a partir de la literatura internacional y el caso Chileno [School performance accountability policies: A debate around international literature and the Chilean case]. *Estudios Pedagógicos*, *2*(42), 395–413. http://dx.doi.org/10.4067/S0718-07052016000200023

Flórez, T. (2013). *Análisis crítico de la validez del SIMCE*. CNED.

Flórez, T. (2018). Repensar la evaluación a gran escala en función de una formación integral: análisis crítico de los sistemas vigentes y posibles caminos alternativos [Rethinking large-scale assessment for a well-rounded education: A critical analysis of current systems and potential alternative paths]. In A. Arratia & L. Osandón (Eds.), *Políticas para el desarrollo del currículum: Reflexiones y propuestas* (pp. 417–456). UCE-MINEDUC and UNESCO.

Flórez, T., & Olave, J. M. (2020). Hacia una construcción participativa de la política educacional: La propuesta alternativa del proyecto de Éxito Escolar en Valparaíso [Towards a participatory construction of education policy: The School Success project in Valparaiso as an alternative proposal]. *Revista on line de Política e Gestão Educacional*, *24*(1), 867–890. http://dx.doi.org/10.22633/rpge.v24iesp1.13790Flórez, T., Rozas, T., Gysling, J., & Olave, J. M. (2018). The consequences of metrics for social justice: Tensions, pending issues and questions. *Oxford Review of Education*, *44*(5), 651–667. https://doi.org/10.1080/03054985.2018.1500356

Fraser, N. (2008). *Escalas de justicia* [Scales of justice] (1st ed.). Herder.

Geertz, C. (2005). *La interpretación de las culturas* [The interpretation of cultures]. Gedisa.

Gutiérrez, P. (2003, May 2). La estrategia de los niños de Melipilla que superaron los límites del [SIMCE The strategy of children in Melipilla who overcame the limits of SIMCE]. *La Tercera,* p. 9.

Lingard, B., & Ozga, J. (2007). Introduction: Reading education policy and politics. In J. Ozga & B. Lingard (Eds.), *The RoutledgeFalmer reader in education policy and politics* (pp. 1–8). Routledge.

Luaces, O. (2014). *Evaluaciones en línea: Una innovación educativa en el marco del programa de una computadora por alumno* [Online assessments: An education innovation in the context of the one computer per-student programme]. Administración Nacional de Educación Pública. https://sea.anep.edu.uy/libro-sobre-la-evaluacion-en-linea

Manzi, J., Gutiérrez, G., Grau, V., & Volante, P. (2014). *Análisis sobre valoraciones, comprensión y uso del SIMCE por parte de directores escolares de establecimientos subvencionados* [Analysis of the assessment, understanding and use of SIMCE by headteachers in subsidised schools]. MINEDUC-FONIDE.

Maxwell, G., & Cumming, J. (2011). Managing without public examinations: Successful and sustained curriculum and assessment reform in Queensland. In L. Yates, C. Collins, & K. O'Connor (Eds.), *Australia's curriculum dilemmas: State cultures and the big issues* (pp. 186–206). Melbourne University.

O'Neill, O. (2013). Intelligent accountability in education. *Oxford Review of Education*, *39*(1), 4–16. https://doi.org/10.1080/03054985.2013.764761

Ortiz, I. (2012). En torno a la validez del Sistema de Medición de la Calidad de la Educación en Chile [Around the validity of the System for the Measurement of Quality in Education in Chile]. *Estudios Pedagógicos (Valdivia)*, *38*(2), 355–373. https://doi.org/10.4067/S0718-07052012000200022

Partnership for Educational Revitalization in the Americas. (2009). Tendencias en la evaluación del logro escolar: La experiencia de cinco países con alto rendimiento educativo [Trends in school achievement assessment: The experience of five countries with high educational performance]. *Serie Mejores Prácticas*, *11*(31), 1–4. https://www.thedialogue.org/wp-content/uploads/2016/05/31-Tendencias-en-la-evaluaci%C3%B3n-del-logro-escolar-la-experiencia-de-cinco-pa%C3%ADses-con-alto-rendimiento-educativo-Junio-2009.pdf

Pavez, K. (2008, May 27). El caso de la escuela más pobre con mejor puntaje [The case of the poorest school with the best score]. *La Tercera,* p. 15.

Rockwell, E. (2009). *La experiencia etnográfica* [The ethnographic experience]. Paidós.

Rojas, M. T., & Leyton, D. (2015). La nueva subjetividad docente: Construcción de subjetividades docentes en los inicios de la implementación de la Subvención Escolar Preferencial en Chile [New teacher subjectivity: The construction of teachers' subjectivities in the beginnings of the implementation of the Preferential School Subvention in Chile]. *Estudios Pedagógicos (Valdivia)*, *40*(Especial), 205–221. https://doi.org/10.4067/s0718-07052014000200012

Sahlberg, P. (2010). Rethinking accountability in a knowledge society. *Journal of Educational Change, 11*, 45–61. https://doi.org/10.1007/s10833-008-9098-2

Shepard, L. (1992). *Will national tests improve student learning?* CSE Technical report 342, CRESST, University of Colorado.

Simoes, P., & Pereira, S. (2017, October 2–6). *Low stakes national external assessment in Portugal: A singular case among European school systems* [Paper presentation]. Annual Conference of the International Association of Educational Assessment (IAEA), Batumi, Georgia.

Taut, S., Cortés, F., Sebastian, C., & Preiss, D. (2009). Evaluating school and parent reports of the national student achievement testing system (SIMCE) in Chile: Access, comprehension, and use. *Evaluation and Program Planning, 32*, 129–137. doi:10.1016/j.evalprogplan.2008.10.004

La Tercera. (2014, November 25). *Cuestionamientos al Simce* [Criticism towards Simce]. https://www.latercera.com/noticia/cuestionamientos-al-simce/

UNESCO. (2017). *Rendir cuentas en el ámbito de la educación: Cumplir nuestros compromisos. Informe de seguimiento de la educación en el mundo* [Accountability in education: To fulfil our commitments. Monitoring education around the world report]. http://unesdoc.unesco.org/images/0026/002610/261016S.pdf

van Dijk, T. A. (1999). *Ideología: Un enfoque multidisciplinario* [Ideology: A multidisciplinary approach]. Gedisa.

Verger, A., Fontdevila, C., & Zancajo, A. (2016). *The privatization of education: A political economy of global education reform*. Teachers College Press.

Wandall, J. (2009). *National tests in Denmark: CAT as a pedagogic tool*. Danish National School Agency.

Zúniga, V., & Aravena, N. (2003, May 5). Excepciones en el SIMCE: Nuevas estrategias derriban el mito del factor socioeconómico. Privados elevan el rendimiento de escuelas pobres [Exceptions in SIMCE: New strategies demolish the myth of the socioeconomic factor. Private [consultants] raise performance of poor schools]. *El Mercurio,* Section C, p. 3.