



# Enacting performance-based accountability in a Southern European school system: between administrative and market logics

Marcel Pagès<sup>1</sup> 

Received: 13 August 2020 / Accepted: 9 March 2021

© The Author(s), under exclusive licence to Springer Nature B.V. 2021

## Abstract

Performance-based accountability (PBA) policies are increasingly adopted in a wide range of education systems in order to reform school governance and to improve students' results and schools' performance. Countries around the world have been implementing national large-scale assessments to make school actors more accountable and responsible for students' results. This policy model has been generalized in countries with different administrative traditions, including those with a short tradition in New Public Management. This is the case in Spain, where PBA has been adopted unevenly in different regions, with Madrid being one of the earliest adopters. In recent decades, Madrid has developed a model that combines administrative test-based accountability with a system of broad parental school choice, which also facilitates the activation of market forms of accountability. However, the combination and interaction between market and administrative forms of accountability is understudied. This paper adopts a policy enactment perspective to analyze, through a case study approach, the interaction of administrative and market forms of accountability and its enactment at the school level. The case study is based on a set of 41 semi-structured interviews with teachers, principals, and school inspectors in a sample of eight schools in Madrid, combined with document analysis of school educational projects and improvement plans. The evidence suggests that administrative and market forms of accountability tend to generate dynamics of interdependence, resulting in increasing external pressures which schools tend to address with superficial responses, including teaching to the test, or second-order competition between schools.

**Keywords** Performance-based accountability · Policy enactment · Education policy · Decoupling · School choice

---

✉ Marcel Pagès  
Marcel.pages@uab.cat; Marcellpages.pm@gmail.com

<sup>1</sup> Departament de Sociologia, Universitat Autònoma de Barcelona, Edifici E1 - Campus Bellaterra, 08193 Bellaterra, Barcelona, Spain

## 1 Introduction

In increasingly competitive and interdependent world economies, school improvement programs, oriented to foster student learning and instrumental competences, are currently at the core of managerial reform agendas in education. Performance-based accountability (PBA) policies are one specific manifestation of this global reform trend, aiming to make school actors more accountable to different audiences on the basis of student performance in external national tests. Indeed, this reform approach tends to be adopted together with the prescription of learning standards, but it is also accompanied by the promotion of higher levels of decentralization and school autonomy, implying new governance arrangements for educational systems (Verger et al., 2019b). In this policy context, schools are expected to make use of their margins of autonomy to develop educational projects better aligned with student needs in order to enhance learning, reach the standards, and improve student performance in external and standardized tests. Schools are also expected to use data generated by the standardized national tests in order to identify weaknesses and implement corrective measures. The combination of these elements implies new modes of external and internal regulation of schools and, in certain circumstances, could involve increasing performative pressures among school actors (Ball, 2003; Falabella, 2014; Perryman et al., 2011) producing unintended effects such as teaching to the test, student triage, or cream-skimming (Mittleman & Jennings, 2018).

This policy approach shares some global common traits but is adapted and translated in different settings according to institutional variables and to political or economic junctures. The specific translation of PBA reforms in different countries interacts with the prevailing administrative traditions, the governance architecture of education systems, the association with other education policies, and other contextual conditions. Despite divergences, accountability reforms have been increasingly disseminated around the world, and the use of national standardized tests to assess student performance is being generalized, even in countries without managerial-oriented administrative traditions (Verger et al., 2019a). This is the case in Spain, where accountability policies have been adopted and implemented unevenly across regions, Madrid being one of the most active in deploying and using external and standardized tests to reform the governance of education (Verger et al., 2020a; Pagès & Prieto, 2020). In Madrid, schools are subject to an accountability regulatory framework that combines (i) bureaucratic control through inspection services; (ii) external and standardized evaluation of school performance, and (iii) parental school choice. The combination of these regulatory elements provides an excellent context to study and analyze the interaction and juxtaposition of different forms of accountability, mainly market-oriented and administrative-based. Both forms of accountability have often been analyzed separately, sometimes under the formal assumption that both work autonomously.

Nonetheless, the implementation of accountability schemes is not a linear process and is highly reliant on how school actors interpret and translate the different prerogatives of multiple accountability policy mandates, often implying a

superposition of multiple external demands which schools face simultaneously and need to de-codify and translate into different institutional and socio-economic contexts. Such contexts can foster or inhibit specific forms of policy enactment. For this reason, a deeper and more nuanced analysis of the interaction between administrative and market forms of accountability would improve our understanding of the policy enactments of PBA at the school level.

Aiming to address this gap, this study examines the enactment of accountability policy mechanisms in a quasi-market educational setting. Specifically, its main objectives are (i) to analyze the interaction of different forms of accountability to which schools are subject; (ii) to better understand how school actors make sense of these forms of accountability, and (iii) to identify the strategies and practices through which schools address external accountability pressures. In order to achieve these objectives, this paper presents a case study of the enactment of accountability reforms at the school level in Madrid, combining interviews with school actors ( $n=41$ ) in eight primary schools with document analysis of school educational projects and improvement plans.

The paper is structured as follows. First, the analytical approach of the study is presented, which is followed by a section outlining the main characteristics of the study's policy context, including recent accountability reforms. The next section presents the methodological strategy of the research, specifying aspects of data gathering, sampling, and analysis. Subsequently, the results of the empirical research are presented and organized according to the three main objectives of the study. The paper ends by presenting the main conclusions and implications in terms of policy and research.

## **2 Beyond implementation: the policy enactment of PBA**

PBA systems aim to enhance the responsibility of school actors regarding the quality and performance of student learning, measured on the basis of external testing. However, accountability schemes can follow different approaches and be formally configured through multiple policy designs. Furthermore, the implementation of accountability arrangements is not a linear process. It implies a complex and contradictory dynamic in which school actors have to translate and negotiate policy prerogatives — in this case those attached to different accountability policy tools — into specific school practices (Ball et al., 2012). In this paper, instead of policy implementation, the notion of policy enactment is used to analyze the interpretation and translation made by policy actors in order to “put policies into practice” in a “creative, sophisticated and complex but also constrained process” (Braun et al., 2011, p. 586).

### **2.1 Interpretation and sense making**

From the point of view of policy enactment theories, policies are not mechanically implemented but result from complex processes of negotiation, de-codification,

and adaptation of meanings. Indeed, the meanings, objectives, and ideas attached to policies are not always obvious but can be rather presented in ambivalent and ambiguous ways. This applies in particular to “meanings and practices surrounding accountability” which are not “absolute, but rather inherently problematic” (Gawlik, 2015, p. 396). Hence, within accountability systems, teachers and principals act as enactors who have to interpret and make sense of different policy messages according to collective and individual subjective accounts.

The way school actors make sense of a new reform is contingent to how these actors have received and understood previous policies in the past. Indeed, “what a policy comes to mean for implementing agents depends to a great extent on their repertoire of existing knowledge and experience” (Spillane et al., 2002, 393). School actors are more likely to incorporate familiar ideas and to ignore other ideas that are not in line with their current practices and beliefs. This is a cognitive process that often leads to misunderstandings within the implementation of new policy mandates (Spillane, 2009). Thus, in order to analyze the policy enactment of PBA, we have to consider school actors’ previous beliefs, knowledge, practices, and routines. Sense making refers here to an individual sphere; however, it should be also considered as a relational phenomenon build on social interaction and shared understandings (Spillane et al., 2002) that affect how “individuals and groups actively construct understandings and interpretations” about specific policy messages by “placing them into pre-existing cognitive frameworks” (Coburn, 2005, p. 478). Moreover, the process of interpretation does not take place in a vacuum but is embedded in organizational contexts, institutional environments, and professional cultures, among other dimensions (Spillane et al., 2002).

Overall, the success of a given policy reform is not only based on “efficient” designs but is highly reliant on how enactors interpret and make sense of policy messages in particular school contexts. Moreover, the subjective perception of different external policy pressures — either market-based or administrative oriented — is a constitutive part of this process and affects its subsequent responses and translations (Verger et al., 2020b).

## 2.2 External demands and accountability pressures

Under accountability regimes, schools face multiple and complex policy demands, which can be internally contradictory or can contradict other overlapping policy mandates. For instance, accountability relationships activate reputational concerns that act as a filtering mechanism to interpret and respond to multiple audiences (families, educational authorities, school owners, etc.), with diverging and even contradictory demands (Busuioc & Lodge, 2017). Administrative and market forms of accountability activate different forms of policy pressures towards schools, although “what makes schools reactive is not only the level of pressure that regulations exert but also the pressure that school actors perceive, live and experience” (Verger et al., 2020b, 223). From this perspective, the perceived pressure is a central element to understand how school actors make sense of accountability regulations and how they respond to different external demands.

The so-called decoupling argument, developed by neo-institutionalist theorists in the late 1970s, suggests that policy reforms tend to fail because a disconnection emerges between the external demands of the policy environment and actual organizational practices. This approach suggests that organizations tend to face and address external pressures of the institutional environment by strategically adopting symbolic changes in the organizational structure, without internalizing relevant changes into day-to-day practices. Hence, decoupling appears when organizations aim to solve conflicting and contradictory institutional demands by disconnecting formal organizational structures from actual practices (Meyer & Rowan, 1977).

The emergence of decoupling dynamics could be explained by different factors, including the various coercive pressures experienced by organizational actors, the trust relationships among policy and school actors, as well as the subjective beliefs regarding the efficacy of the measures which have to be implemented (Boxenbaum & Jonsson, 2017, p. 21).

As a result of decades of research and theoretical refinement, the main argument of the early contributions of neo-institutionalism regarding the notion of decoupling has been nuanced, suggesting increasing dynamics of policy alignment in contexts of accountability and more complex responses to external policy demands and accountability pressures.

### **2.3 School responses and policy translations**

Schools react to external policy pressures with complex and multiple organizational and instructional responses, which should be analyzed beyond a binary approach of alignment and decoupling, in order to focus on a broader range of responses “from passive conformity to active resistance, depending on the nature and context of the (external policy) pressures” (Oliver, 1991 p. 146). Focusing on schools, Coburn (2004) suggests that we should “rethink the decoupling argument” providing evidence on how the policy environment “penetrates schools in substantial ways” and how school actors respond to them. Such responses include forms of non-incorporation of the policy messages (with strategies of open rejection, decoupling, and parallel structures) and forms of incorporation (with assimilation and accommodation responses), which are developed by means of teachers’ agency in a context of “bonded autonomy” (Coburn, 2004).

Another aspect to take into consideration when “rethinking” the decoupling argument is that the policy context has changed significantly since the 1970s. The introduction of accountability mechanisms was intended, and partially achieved, to reinforce the alignment between instructional practices and policy expectations. In fact, more recent neo-institutionalist studies stressed that current accountability schemes “have led to a shift to more tightly coupled and narrowed controlled practices in organizations” (Meyer & Rowan, 2006, p. 2). Certainly, accountability mechanisms, especially those with high-stakes policy designs, appear to be able to partially penetrate classroom practices (Diamond,

2012), despite sometimes disrupting teacher autonomy (Hallett, 2010, p. 61) and resulting in policy outcomes that differ from initial expectations, including narrowing the curriculum or the reinforcement of test-driven educational approaches (Au, 2007; Lipman, 2004).

Despite the fact that that high-stakes accountability seems to operate with a greater capacity for policy alignment, a recent sequential and mixed-methods study (Verger et al., 2020b) has suggested that high-stakes systems do not always function “as a performative device in all circumstances” (p. 18). This research suggests that schools react to accountability pressures with divergent responses “which go beyond conventional classifications mainly focusing on alignment vs decoupling” (p. 17) and are significantly mediated by how school actors interpret the accountability mandates and, even more importantly, by how they experience the subjective pressures resulting from accountability policy instruments in different school contexts. Accordingly, schools react to accountability pressures with different responses, namely, accommodation, induced alignment, dilution, fabrication, and *de facto* opting out.

An increasing body of research has also analyzed school responses to low-stakes accountability systems, those which do not attach clear and explicit schemes of incentives and sanctions to the test results, and formally rely on the reflexivity and data-use capacity of school actors in addressing policy expectations (Landri, 2018; Maroy, 2015; Skedsmo, 2011). In a comparative study between two different schemes of accountability, Dumay et al. (2013) found that networked forms of accountability tend to promote better policy alignment than bureaucratic forms of accountability, whereas other studies suggest that low-stakes accountability systems in quasi-market settings induce schools to respond superficially to accountability demands, generating symbolic changes in schools and dynamics of policy decoupling (Barbana et al., 2020; Maroy, 2015). Indeed, low-stakes accountability systems generate multiple school and policy responses, which to a great extent rely on policy design variables. For instance, in a comparative study, Maroy and Pons (2019) found that the re-contextualization of low-stakes accountability in Quebec generated pedagogical practices that were better aligned to managerial expectations with similar effects as high-stakes schemes. However, in France, accountability tools were adopted with no relevant impact on teaching practice and instruction. At the school level, other decoupled responses to accountability regulations were reported in French-speaking Belgium, where the more reactive schools responded with forms of “horizontal decoupling,” meaning a disconnection between the policy implementation of testing with its ends and means, “implementing the mandatory part of the accountability instruments while avoiding a real penetration into the cognitive scripts of schools” (Barbana et al., 2020, p. 12).

Informed by these theoretical and empirical accounts, this paper aims to understand the enactment of PBA in Madrid by analyzing the way policy design mediates the sense-making process of accountability tools and how school actors respond, generating dynamics of alignment and decoupling of low-stakes accountability arrangements in a quasi-market education system.

### 3 Policy context

South European countries are particularly interesting contexts in which to study the adoption and enactment of PBA reforms. In countries such as Spain, Italy, and Portugal, recent accountability reforms have been introduced under the tenets of New Public Management (see Verger & Curran, 2014; Parcerisa, 2016; Landri, 2018; Carvalho et al., 2020). However, these are countries where NPM does not have a strong or long-lasting tradition in public administration, and the policy adoption of such reforms has tended to be erratic and contradictory. Spain is a late adopter of accountability reforms in education and, as in other sectors of public administration, has adopted these mechanisms following a quality assurance rationale. Within the Spanish education context, Madrid is an exception to this dynamic in the sense that accountability reforms not only have followed a quality assurance motivation but also aimed at the introduction of market logics and competition dynamics in the school system (Villamor & Prieto, 2016).

At the beginning of the 2000s, Madrid stated a governance reform of education which included the introduction of accountability policy tools, school autonomy measures, and the promotion of school choice. In parallel, the publicly subsidized private school sector was expanded<sup>1</sup> with the intention of promoting school diversification and increasing school choice opportunities (Carpintero & Siemiatycki, 2015). This education policy approach contributed to enhance the already high levels of school segregation in Madrid (Bonal & Zancajo, 2018) and further consolidated educational inequalities (Escardíbul & Villarroya, 2009).

In Madrid, the external test was introduced in 2005 to increase curriculum control and improve transparency in school results in order to enhance parental school choice. Nonetheless, the trajectory of accountability policy instruments has been far from linear and suffered at least two main changes: the dissemination of the test results and the evaluative model of the test.

Regarding the policy of test result transparency, the results of the external test were publicly disseminated in different formats between 2005 and 2015, altering the school governance dynamics and generating notable external pressures. Initially, the test results were made public in the form of school rankings, which triggered great opposition from teachers, principals, and public-school families (Verger et al., 2020a). However, since 2015, the results are no longer published. Interestingly, this change did not result from a governmental turn but from a variety of factors including a lack of policy consensus, administrative barriers, and issues of political leadership. Very briefly, it is important to highlight that the adoption of a high-stakes accountability policy model in Madrid was the result of a very particular policy leadership dynamics within the regional government. Once this leadership declined, a shift towards a more pragmatic policy approach was adopted by educational authorities in 2015, assuming that a lower-stakes accountability model would allow at the same time to reach higher levels of political consensus and a better alignment

<sup>1</sup> According to Carpintero and Siemiatycki (2015), between 2005 and 2012, new Public–Private Partnerships were agreed to build and operate schools to provide education for about 60,000 students and with a total investment of around 650 million euros (439).



with the national policy framework (Pagès & Prieto, 2020). In this context, the same administration that established the standardized test ended the policy of transparency without major political noise. Thus, Madrid represents a very particular policy context to study accountability in the sense that it has moved, in a relatively short period of time, from a high-stakes accountability system to a lower-stakes approach.

The second main change refers to the evaluative model of the test. The first models of the external test were designed with a content-oriented and memory-based approach. This continued until 2015 when the test changed and adopted a more complex design based on skills and instrumental competences.

Currently, the test results are disseminated to the families and the schools, which are obligated to develop a Performance Improvement Plan based on the analysis of the external test data. Schools are asked to analyze retrospectively the results of the external test, compare it with the internal evaluation, and identify aspects of improvement in instrumental skills. The improvement plans should incorporate general and specific objectives, didactic methodologies, specific interventions, and evaluation indicators. These plans are monitored and evaluated by the school inspection body and are expected to guide the process of internal evaluation for school improvement on the basis of teachers' and principals' reflexivity.

The policy context and the trajectory of accountability policies in Madrid are characterized by a sort of back and forth dynamics regarding the design and the policy expectations attached to PBA. In this context, this research suggests two preliminary hypotheses regarding the way that the policy framework has influenced the enactment of accountability arrangements in Madrid. First, the uneven policy trajectory of PBA and the back and forth dynamics regarding the use of accountability policy tools and the accompanying expectations would contribute to a superficial implementation of accountability policy instruments in schools, without altering the organizational routines of schools or the teaching practices. The second hypothesis suggests that the policy precedents related to the publication of school results generated a testing performative pressure that tended to persist among school actors. This occurred even in lower-stakes systems, especially when combined with market forms of accountability and mechanisms of parental school choice, resulting in school-level effects similar to those seen in high-stakes accountability models.

## 4 Methods

To better understand the sense-making process of PBA and its translations into specific school practices, this paper conducts qualitative empirical research on the basis of the case study approach. More specifically, this research develops an explanatory single case study with embedded units of analysis (Yin, 2003). The case study approach is appropriate when studying a contemporary social phenomenon within its real-life context, especially when the boundaries between the object of study and the context are not clear (Yin, 2003, p. 13). In this research, the enactment of PBA in Madrid is conceived of as the analyzed case, whereas the different schools included in the sample are treated as embedded units of analysis. This approach is adopted to analyze how a particular social phenomenon — the



enactment of PBA in Madrid — takes place in different circumstances and conditions, namely divergent institutional and contextual school settings. The results are not presented as differentiating by each unit of analysis, but as focusing on the conditions and the generative mechanisms of certain sense-making processes and school practices. The analysis, therefore, pays special attention to the potential mediating conditions of policy enactment, looking at school-level factors such as student performance, the role of school culture, and the subjective pressures experienced by teachers and principals.

The primary data for the study come from semi-structured in-depth interviews with teachers and principals. Assuming that the institutional and contextual factors influence different forms of policy enactment, the case study gathers data from eight schools with different social and organizational characteristics. The schools have been selected following theoretical sampling criteria according to different factors including (i) the type of school provider (public or private); (ii) the socio-economic status of the school (defined by the income level of the school neighborhood as a proxy), and (iii) the school culture (defined inductively as more traditional or more innovative according to accounts based on explorative and qualitative fieldwork in a broader sample of schools). Moreover, three extra interviews with school inspectors have been included to address three main issues. The first issue is to better understand bureaucratic forms of accountability, the second issue is to incorporate *meso* elements and interschool dynamics into the analysis, and the third issue is to add an external account of how schools implement accountability arrangements.

As can be observed in Table 1, a sample of key informants were interviewed in each school. These interviews have been complemented and triangulated with document analysis and preliminary non-systematic observations in each school. The document analysis and the primary observations were used as a strategy to define the sample of schools included in the research and as complementary sources of information, whereas the interviews were treated as the main fieldwork data for the analysis. This methodological decision has allowed more schools to be covered in different institutional and social contexts in order to focus on the conditions that influence policy enactment. The schools are treated as sampling units and the individual actors (teachers and principals) as the main units of analysis embedded in a particular school context. When the responses of teachers and principals within one school concur and school documents and preliminary observational notes are in agreement, then aggregated individual responses are analytically inferred to be school-level responses.

The main research instrument used is the semi-structured interview script, which included questions about personal background, school context and culture, subjective opinions regarding PBA, and school practices. Textual qualitative data were analyzed with inductive emerging codes and predefined deductive codes regarding the school context and culture, the interpretation of accountability policy instruments, and the translation into specific school practices. Specialized software for qualitative analysis of textual data was used to conduct the analysis of interviews.

**Table 1** Sampling and detail of interviews

School	Provider	SES level	S. Ethos	Teachers	Principals	Inspectors
A	Public	Low	Mixed	3	1	
B	Public	Med-low	Traditional	1	2	
C	Public	Med	Innovative	2	3	
D	Public	Med-high	Traditional	3	3	
E	Private	Med	Traditional	2	2	
F	Private	Med-high	Mixed	2	3	
G	Private	High	Traditional	4	3	
H	Private	High	Innovative	2	2	
Total				19	19	3

## 5 Results: understanding policy enactment of PBA in Madrid

In this section, the results of the research are presented and analyzed according to the different objectives that guide the study as outlined in Section 1. Accordingly, this section gives accounts of the different policy interpretations and rationalities of school actors regarding PBA, the interaction of different forms of accountability and its associated external pressures, as well as the resulting school practices and responses.

### 5.1 Policy interpretations and multiple rationalities

The results of the interviews show how school actors make sense of PBA in diverging and non-univocal forms. The interpretation of testing policy instruments varies between and within schools and is conditioned by both contextual elements and factors of a subjective nature. In this section, an analysis of the different interpretations of PBA is presented according to the different rationales and dimensions of the policy. By doing this, the findings highlight how the sense-making process of accountability policy instruments goes beyond con/pro positions. That is, the analysis reveals that neither schools nor individuals can be easily labeled regarding their opinion about PBA since their positions tend to be more nuanced and contingent on the policy dimensions in question.

The school actors interviewed tended to assume a vague positive opinion of the testing instrument, specifically regarding the evaluative approach of the test. Here, the precedents of the evaluative design of external tests are crucial to understand this specific form of interpretation. The school actors interviewed stress the virtues of the current model, which is perceived to be a “better idea” due to the “competence-based approach” (Teacher 7, School D), more appropriate and adapted to the teaching–learning process, relatively well “aligned to what students should know, understand and do” (Principal 4, School C), involving “creativity” and “establishing relationships between different problems” (Teacher 19, School H).

Another positive aspect that most of the principals highlight is the non-dissemination of the test results. Here again, policy precedents play a key role, and the end of the policy of school scores publication is seen as a positive change by all the school principals interviewed. Nevertheless, although school results are no longer published, such a precedent still operates generating a subjective sense of pressure as will be further analyzed.

The general nature and the purposes of PBA are interpreted in a more ambivalent way. Some teachers and principals suggest that “a control mechanism is important” (Principal 2, School B) and “some kind of external evaluation is positive” (Principal 10, School E). It is understood as an improvement mechanism to “make sure people do not relax” (Teacher 6, School C). However, this vague positive opinion is broadly surpassed by criticisms of PBA, which place the emphasis on different meaning frames: pedagogical, professional, social justice-oriented, and reputational.

**Pedagogical meaning frame** Most teachers interviewed criticize the external test because of concerns for student well-being. They point out that the test implies great levels of stress and pressure for students. This argument is usually present in schools that are more focused on socio-emotional approaches than student performance and are usually those schools attended by disadvantaged students. In middle-class and more innovative schools, a similar critical discourse with testing was found. In these cases, teachers highlight that the external standardized test implies significant tensions with normative educational approaches and schools' pedagogical models such as personalized learning or cooperative methodologies.

They [the educational authorities] ask us to adapt teaching to our students and their capacities, so why there is a single external exam for all? What kind of adaptation is that? (Teacher 6, School C)

Our students struggle with the external test because they are used to working differently, here we teach on self-reflection, they should do a process and work cooperatively. They are not used to working with this kind of tasks. (Principal 19, School H)

**Professional meaning frame** An argument in parallel with this pedagogical rationale is related to the teachers' sense of professionalism, which appears to be challenged when teachers "feel questioned if the results are not good enough" (Teacher 5, School C). Despite this, most teachers and principals agree that the results of the external test "do not reflect the work a teacher can do with students" (Principal 1, School A) and "do not reflect (education) quality" (Principal 19, School H). This feeling of judgment is shared among tested teachers. Even though the test is not formally used to evaluate teachers, teachers from different schools and contexts agree that the external test implies a certain level of distrust regarding teachers' work, professional duties, and their judgment capacity.

I think we are, as teachers, honest enough to be externally judged about our work (...) you do not have to be judged or exposed to people that do not know how we work. (Teacher 14, School G)

Deep inside, this test denotes distrust towards teaching professionals and implies a devaluation of their qualitative and professional function. (Inspector 2)

This argument is shared by teachers in schools with different pedagogical approaches and with very different student profiles but is mainly an argument articulated by senior professionals teaching tested subjects.

**Social justice meaning frame** Another critical interpretation is based on a social justice approach and suggests that an external and standardized test is not a fair evaluation instrument because it is not adapted to the contextual realities of students and schools. The main claim in this regard is that the test is a "de-contextualized" (Principal 5, School C) evaluation instrument because of its standardized nature. Moreover, some teachers perceive that the

external test contributes to classify schools in different positions according to student background and simply verifies the lack of familiar support and previous deficits in the acquisition of instrumental competences of disadvantaged students.

I think this test is just to classify schools and I think this is not fair considering our student population. For example, it is impossible to achieve the same results here than in my neighbourhood, or in Las Rozas, or in Majadahonda [affluent areas], because children are already much more stimulated in every way. (Teacher 2, School A)

The social justice rationale approach is present in schools from different contexts but is clearly a more intense and common discourse in public schools enrolling students from disadvantaged backgrounds regardless of the school's pedagogical orientation.

**Reputational meaning frame** Contingent to the social justice rationale, a final critical interpretation is based on a negative understanding of testing due to the reputational pressures that schools face because of the publication of the test results, which were seen as “terrible, offensive and shameless” (Principal 2, School B), “uncomfortable” (Principal 15, School G), “used perversely” (Teacher 8, School D), and contributing to make “first class and second class schools” (Principal 1, School A). These arguments are broadly shared by teachers and principals and very well represented by a common idea that suggests the test is internally relatively useful, but externally damaging, as expressed by one teacher:

The test can give you an approach on where you fail, which is not bad, so I would use it internally but not for external exposition. (Teacher 14, School G)

Not surprisingly, if the non-dissemination of school results is seen as positive, the transparency of school scores is interpreted critically for different reasons, including the increasing market pressures associated (related to student attraction capacity and loss of students if the school performance is worse than expected) or the pressures exerted by other actors including parents, owners, or the educational authorities.

The next section addresses the way these pressures are perceived and how school actors respond to them with a range of different school practices.

## 5.2 The double face of accountability pressures

Accountability pressures come from different sources and audiences, imply different expectations, and are perceived unevenly by school actors with different levels of intensity and importance. In the section below, I focus on how administrative and market forms of accountability interact and generate specific forms of external pressure in different school contexts, and how school actors face them.

### 5.2.1 Administrative accountability pressures

Administrative forms of accountability imply the interaction between schools and educational authorities and combine reputational effects with elements of administrative support and control. Principals report uneven levels of administrative pressure, and when these are present, they seem to be driven, especially in public schools, by the hierarchical component of the accountability relationship between inspectors and principals. Interestingly, despite there being no formal scheme of sanctions and incentives, this relationship is often understood as an external form of work control.

Well, yes, I feel pressure (...) if your results are very bad it seems like the inspection is *running your work*. I know that the inspection gives importance to these results. (...) So, if your *superiors* give importance to that, there is a certain pressure to ensure that everything is going well. (Principal 7, School D)

Principals in public schools feel more directly pressured by educational authorities than those in publicly subsidized private schools, especially in terms of bureaucratic requirements, but not regarding school performance. Formally, both types of schools were equally accountable to public administration, but the subjective perceptions of school actors regarding the role of educational authorities differ slightly. In both types of schools, the role of the educational authorities is generally seen as a form of bureaucratic control for the accomplishment of certain legal requirements, but the administrative pressure regarding student performance is not a generalized concern among teachers. In contrast, principals, especially in public schools, emphasize the role of school inspectors as monitoring agents of school performance and learning standards. At the same time, it is mainly in public schools attended by disadvantaged students where teachers and principals report a lack of administrative support and PBA is seen as a trigger for shaming and blaming dynamics.

Despite poor results, we have never received any support (...). They cannot *blame us* for the shortfalls this school could have, or any other in Madrid's deprived areas. Where are the resources? (Teacher 3, School A)

I think external testing is useless, they spend a lot of money analysing these results (...) but what do they do with underperforming schools? (...) They do not do anything, just *point at schools* and say "how badly the students in this school are doing", so it is useless. (Principal 2, School B)

Other teachers state that they do not feel directly pressured by inspectors, but by other complementary and indirect sources, which tend to be self-generated, yet are also driven by reputational concerns regarding educational authorities and families.

No, no, [the educational authorities] don't say anything but you pressure yourself because you want to leave your school in a good position. (Teacher 12, School F)

In private institutions, the school owner develops functions very similar to those exerted by educational authorities and are often seen as “another school inspection” which “asks for a lot of paperwork” and requires specific forms of school organization (Teacher 14, School G). In some cases, school owners act as administrative account-holders controlling the evaluation and monitoring of performance standards. Interestingly, school owners also develop an administrative role ensuring the accomplishment of guidelines “dictating certain norms” and being “something apart which regulate schools” (Teacher 19, School H). In fact, some publicly subsidized private schools suggest that they are held accountable by their owners more than by public authorities. In these cases, school owners tend to act as constrainers of pedagogical school autonomy by defining specific appealing educational approaches for families, often used as a strategy of differentiation in order to improve the school’s relative position in the marketplace. School owners are in this sense a very interesting accountability actor, who can skilfully synthesize market and administrative demands.

### 5.2.2 Market accountability pressures

Under quasi-market regulations, organizations tend to be increasingly accountable to their “clients” in order to deliver a “quality product” on the basis of choice and demand. Market forms of school accountability are based on the relationship between school and family preferences, often mediated by normative notions of quality education. The results of this research suggest that the pressures exerted by market forms of accountability appear to be more intense and powerful than expected in all types of schools studied. The logics of market accountability sometimes appear together with elements of PBA, but the results also indicate certain autonomy of the market mechanisms from PBA arrangements. In Madrid, accountability arrangements and market regulations have aimed to modulate the behavior of school actors, consolidating an increasing social interest in school performance, as well as a certain audit culture, which was recently channeled by informal means of communication, as suggested by a school inspector:

Even though rankings are not in place, schools still use the results of the external test. (...) If you talk with parents when they are going to choose schools, one thing that they ask principals is about school test performance. (Inspector 1)

In this context, school actors still perceive PBA as a policy instrument of a reputational and market-oriented nature, which can damage the image of schools in a free school choice environment with high levels of competition. Some school actors report feeling less pressured since the finalization of the policy of transparency but suggest that the main source of pressure was generated by school choice dynamics, mainly influenced by the levels of school demand.

Yes, test results are very important, especially when they were published on the Internet. There was a lot of pressure related to parental school choice. However, our school has always had many students, so we were less worried.



I mean, we were not worried regarding loss of students but well, you always want to have a good image. (Teacher 19, School H)

The school ethos and the educational aspirations of families configure to a great extent the dynamics of market accountability. In high-performing schools that enroll students with advantaged backgrounds, performance is taken for granted and the academic culture is part of the school identity. In these contexts, achieving and maintaining good results in the external test seems to be a priority to sustain the school image and internal legitimacy. In addition, good results ensure the loyalty of families. Families emerge as the main pressuring actor regarding school performance, often more important than educational authorities.

This is not about the administration. Pressure comes from families, because of such a competitive environment, in which they want their children performing so much, out-standing from the beginning. There is more pressure from families and from a very broad school supply in this area. (Principal 17, School G)

However, the test not only works as an internal source of legitimacy but also as a strategy for external reputation building. Schools struggling to recruit students experience testing in a dual manner, both as a mechanism of reputation building and also as a potential threat for student recruitment. Indeed, principals in these schools suggest that families tend to “choose on the basis of the external test” (Principal 14, School F) using non-official data which circulates “by word of mouth” and which could “spread as an unfounded rumour” (Principal 12, School F). In these contexts, the external test appears to be a critical factor contributing to building or eroding the school reputation and its capacity for student recruitment, by means of a “*call effect*” mechanism related to the test performance (Teacher 11, School E).

The emergence of market forms of accountability is not exclusive to institutions with a high educational reputation, it also emerges in other types of school. In public schools with a more academic ethos, performative pressures remained, together with reputational concerns and dynamics of school competition. These are internalized as a form of “teachers’ responsibility in terms of image” (Teacher 4, School B). Despite the fact that test results are not published, the importance of performance among families and the role the non-formal means of information dissemination contribute to developing and maintaining market pressures in schools. As happens in private schools with higher reputation, principals in public schools enrolling middle-class students stress that parents actively collect information about school performance during enrolment periods and this has an effect on student recruitment.

These rankings, which arrive at the media, I don’t know how, make us have more enrolment. And I know that during the enrolment period parents do amazing and great studies on this. (Principal 8, School D)

Parents come to ask for information about the school and they directly ask for the external test, so yes, this is a pressure that makes us, externally, have a certain reputation or another. (Principal 3, School B)

Indeed, for certain schools, the test could be used for reputation building and marketing purposes as stated by a principal who suggests that “this test only serves

for some schools to reach a certain status” (Principal 5, School C). Interestingly, the reputational component of performance in a free school choice environment is indirectly generated by the increasing performance culture among parents which “exert pressure over principals like they never did before” (Inspector 1). Nevertheless, this trend is also context specific. Schools enrolling disadvantaged students do not experience such direct parental pressure. Such pressure is more likely to appear in schools enrolling middle-class and advantaged students. However, disadvantaged schools experience more indirect forms of market accountability pressures as the result of not being actively chosen. In these schools, the relationship between school performance, student population, and school choice is described as a “loop” that makes these schools look like a “ghetto” and become socially stigmatized. These schools can try to reverse such an image with efforts to attract families from other social backgrounds in order to “change the student population” and “improve the school reputation” (Principal 1, School A). Despite this, the stigmatization that these schools suffer often dilutes their student attraction capacity.

Overall, the results suggest that PBA does not work in isolation from market dynamics and seems to be relatively weak in generating administrative pressure on its own. However, when accountability and testing instruments interact with school choice and market mechanisms, accountability instruments are more able and likely to generate an increasing performance culture and pressure among school actors.

### **5.3 Policy translations and school responses**

The interpretation of the policy mandates of PBA and the different forms in which accountability and market pressures are experienced, derived in specific forms of policy translations which are expressed in particular school practices and responses. In the context of this research, four main school practices can be identified in response to PBA. Two of them are policy outcomes expected within the policy framework, namely the adoption of competence-based methodologies and data-use practices for improvement purposes. Two others can be labeled as undesirable or unexpected results according to the accountability policy framework, i.e., test preparation activities and reputation-building strategies.

#### **5.3.1 Competence-based methodologies**

The generalization of competence-based teaching is the most common practice reported in the interviews and can be partly understood as a response to PBA. Teachers and principals from all the schools report that they intensified the adoption of competence-based methods of evaluation and teaching “in order to align with the external test models” (Principal 1, School A). The adoption of a competence-based approach could be identified in the teaching as well as the evaluation methods. Teachers believe that this is powerfully driven by the external test, as a policy instrument which induces schools to introduce specific pedagogical changes without formally enforcing them:

R: Our exams are based on instrumental competences, so we prepare the exams just like in Madrid's external test

Q: And did you always do it this way?

R: No, since the external test was developed by instrumental competences we began to prepare our internal exams and tests just like the ones done by the Community of Madrid. (Teacher 9, School D)

These practices appear to be relatively internalized, although some barriers emerge in schools with more traditional pedagogical approaches. In these cases, the alignment to the external test is clearly an important factor explaining the adoption of competence-based teaching strategies instead of more traditional methods of teaching.

My method is more traditional; I am still working on morphological analysis, sentence structure, kinds of words... all of this is not asked in the external test, so then for the sake of my students I decided to work more on the reading and the writing. This is good for my students but it is also because of the external test. (Principal 14, School F)

However, the consolidation of competence-based approaches can also be explained by other factors. Competence-based teaching has been formally introduced in Spain as a result of international influences (Tiana, 2011) and has become an emerging *doxa* and a discursive consensus among teachers, although not necessarily translated into new teaching–learning practices (Bolívar, 2008). Therefore, the consolidation of competence-based teaching could be explained because it is considered to be a normative model of teaching and learning, which may be reinforced by the implementation of a competence-based external test.

### 5.3.2 Data use for improvement purposes

The theory of action of accountability regulations expects school actors to use the data generated from the external test in order to improve instruction and implement corrective measures. Interestingly, the results of this research suggest that decoupling dynamics prevails when it comes to data use and analysis derived from the external test as a means to introduce improvement measures. In fact, the analysis of data, the design of improvement plans, and their implementation is not always developed as defined in the current PBA framework. The engagement of schools in the process of data use and analysis differs among schools but is often understood merely as a legal and bureaucratic requirement. In fact, when asked for the process of definition and development of performance improvement plans, teachers and principals in schools with more innovative pedagogies tend to outline its mandatory character and dilute the importance of test results in their reflections about instructional strategies.

We do it because they [the educational authorities] compel us to analyse and to expose it to the school board but we explain just this, that these tests are biased. (Principal 4, School C)

In these cases, performance data analysis for school improvement purposes is often implemented in a superficial manner, frequently understood just as a formal requisite and without being internalized into daily school practices, as suggested by one school inspector:

Teachers are not at all used to performance indicators and all of this; you go there and well... What is done is merely playing to the gallery... I do not see this taking root, but it is true that this language is increasingly becoming more familiar. (Inspector 3)

In fact, some teachers and principals and even school inspectors openly state that data-use requirements are perceived as “senseless” and do not help to generate profound processes of organizational reflexivity. Instead, they are seen as just “another section to fill in the annual report” (Inspector 2). Some schools argue that data practices and improvement plans are useless if they do not come with administrative support and additional resources to implement corrective measures and improvement plans.

[The improvement plans] Do not serve to receive any support figure, so at the end of the day it is more *to play to the gallery* than to improve the results of the school. (Principal 8, School D)

We have to do the improvement plan because of bad results but what we want is more support, the improvement plan will still be failing if we don't have more support and resources. (Teacher 3, School A)

However, the mandatory character of improvement plans implied its generalized adoption in schools, and the formal use of external test data as complementary information to identify aspects of school improvement is generalized, despite in a superficial manner and playing a secondary role.

It is *more data* to use and take into consideration but it *cannot drive your improvement process*, it helps but is just some more information that you have. (Principal 19, School H)

It makes me reflect on where we fail and work a little bit more on it, just like that. (Teacher 14, School G)

Interestingly, the results suggest that despite the fragmented and superficial use of external test data, most schools are increasingly using their internal evaluation results to implement and develop improvement plans. Their plans are more oriented to learning and teaching process approaches, independent of external performance indicators, especially in schools with more innovative approaches. Overall, the use of the external test data to introduce improvement measures in schools is in general superficially adopted, not intensively implemented and only partially internalized. Moreover, external test data are often seen as opposed to internal performance data, which are assumed to be more nuanced and appropriate to identify aspects of improvement, especially in schools with vulnerable students.

### 5.3.3 Test preparation activities

Teaching to the test is one of the most salient undesired results identified in the existing research on school accountability. In this study, the practices of test preparation

are reported in all the interviews, independent of the school context, although these practices vary in terms of intensity, scope, and level of systematization. In general terms, all the school actors admit to preparing and training students for the external test, but they also stress that test preparation is not one of the core activities of their teaching task, neither is it part of the school mission and vision. Counterintuitively, the strategies of test preparation are reported as a certain normalized school practice beyond elements of social desirability. Nevertheless, the normalization of test preparation is often nuanced by considerations regarding the non-intensive character of such practices.

When analyzing how and why these practices emerge, three main school factors are identified, namely pedagogical models, performance culture, and student profile. The pedagogical approach of schools can inhibit or increase the intensity and systematization of test preparation. Schools with more innovative pedagogical approaches experience a certain tension between the logics of PBA, the strategies of test preparation, and the educational approach of schools. In some cases, school actors reported that the innovative pedagogical methodologies “take time from teaching the subject” and “make impossible” to prepare for the test (Teacher 19, School H). Others suggest a more intrinsic contradiction between the innovative pedagogical models and the logics of testing:

In general we don't give a lot of importance [to test preparation] because we are trying to advance other things, we are trying to have a more active methodology and this contradicts testing... but in the end it is a blend, we do not achieved teachers to untie from testing because they feel responsible for their students. (Principal 5, School C)

Schools with a more traditional and academic ethos tend to resort to test preparation activities with greater intensity and systematization. In these cases, test preparation is being more “integrated to the whole course and in the subject curricula” (Principal 10, School E), being “quite determining” and used in class “to teach content” and to know what content teachers should emphasize, or in the words of a principal “to know what needs to go over” (Principal 15, School G). Teaching to the test appears to be more frequent and intense in schools that are more performance-oriented, guided by concerns of fulfillment among families, teachers, and principals, and especially intensive in those schools where excellence and academic results are central elements of the school identity. In these schools, the test results could be seen as an external and internal threat for the school image and hence test preparation appears to be a proper strategy.

Another factor that conditions the adoption of test preparation strategies is the socio-economic background of students. Teachers and principals attending students in disadvantaged contexts suggest that test preparation and teaching to the test is a way to “familiarize students” with the model of the test in order to “avoid student stress and frustration” (Principal 1, School A; Teacher 3, School A). In these cases, the practices of test preparation are clearly linked with the pedagogical meaning frame previously identified in schools with more socio-emotional approaches. Here, test preparation appears to be less intense despite these not being residual practices.

Overall, the intersection of different school factors enables us to explain certain levels of intensity in practices of test preparation. For instance, the combination of pedagogic innovative approaches, well-being students, and a soft academic school culture serve to blur practices of test preparation. In contrast, schools with a high performative culture and more traditional pedagogies are more likely to develop more intensive practices of test preparation regardless of student population profiles.

### 5.3.4 Promotional and marketing strategies

A final approach to PBA policies is related to the promotional activities and the strategies of reputation building exerted by schools in a context of free school choice. Some of these practices are directly affected by the PBA regime, whereas others are more directly related to market dynamics and free-school choice arrangements.

The use of test results for promotional purposes is not a generalized response yet it is present both in public and private schools. These practices have been reported more intensively in schools with high levels of perceived market competition and a major academic school culture. In these schools, good performance in the external test and other performance data (mainly, the rates of promotion or the results in the university access exams, for those schools having secondary education levels) are used as a promotional and marketing strategy. This is especially true for school open days to attract students and improve the school position in a competitive market environment.

If you have to compete for the same students you have to offer a value added, and one of these elements is all the numbers and data for families. (Principal 10, School E)

In marketing terms it is always convenient to obtain good results because there is a lot of competition between schools. (Teacher 14, School G)

Yes, yes. We inform parents about school results, logically. It is our dynamic, I mean, if we have obtained good results, why not to say it? This attracts student enrolment. (Teacher 9, School D)

Well-performing schools have a greater capacity to engage in competitive logics using performance data for school promotional activities. They can develop marketing strategies to better respond to first- and second-order competition dynamics, as reported in other research on education markets (Gewirtz et al., 1995; Maroy & Van Zanten, 2009). Specifically, second-order competition takes place when schools compete not only to attract and maintain enrolment but also to recruit a certain profile of students which in turn could help enhance the schools' position in the school market by building their reputation:

[in case of bad results] our reputation in the neighbourhood can get worse and this could affect us having fewer or *a worse kind of student*. (Principal 3, School B)

In this regard, the performance culture of families and their increasing role in the process of school choice seems to be a condition of possibility for schools to develop strategies of second-order competition.

Parents are very conservative when it comes to choosing the school. So when they visit schools, schools show the external tests results, the rate of promotion, the English level.... They are promoting themselves, that is to say, they *are being accountable, but they are selling the school*. And this is how, little by little, *they reach a certain profile of students*. (Inspector 1)

In contrast, schools that are not as academic or performance-oriented, as well as those with more innovative pedagogical approaches, tend to avoid the use of external test results as a strategy for school promotion. These schools tend to stress the “individualized” and “familiar” character of the school (Principal 18, School H) or other pedagogical aspects such as “active methodologies” and “personalized learning strategies” as their main selling points (Principal 4, School C). This does not mean that promotional activities are absent from these schools, but they do not rely on test performance or other related forms of academic achievement data.

The adoption of certain forms of school methodological innovations is, on many occasions, a way to improve the reputation of schools in the marketplace. This strategy is reported to be more frequent in schools with a higher subjective sense of competitive pressure and is in fact activated in some schools as a response to market mechanisms. Such innovative measures that result from market pressures are often implemented superficially because they are not adopted for intrinsic pedagogical motivations, but as a strategy to enhance the school reputation, attract families, and improve the school market position. They do not imply substantial changes regarding pedagogical approaches, teaching methods, or instructional strategies.

Well, we carried along by what was there. *It is all about trends*. Now it is bilingualism, before it was quality [management] and currently it is about [methodological] innovation. This is so, now if everybody innovates you cannot be left behind innovating. (Principal 15, School G)

In Madrid, a particular but generalized expression of this dynamic is the adhesion of schools to the official Spanish–English Bilingual Programme, which, for all school actors interviewed, was seen as an imperative imposed by the market dynamics of parental school choice and school competition for student enrolment. In fact, all the schools in this research participating in the bilingual program acknowledged that participation in this project was mainly driven by competitive market pressures and as a differentiation strategy.

Altogether, the alignment between school practices and policy expectations are uneven in terms of internalization or superficial implementation according to different school features and contextual factors, including elements of school culture and context. Moreover, unexpected outcomes and instrumental school practices also emerge unevenly and are in some cases highly determined by the market dynamics of school choice and competition.



## 6 Conclusions

This research contributes to the academic debate on the enactment of accountability policy instruments and, more broadly, to the re-contextualization of these instruments in different educational settings. In general terms, the results suggest that the policy framework combining administrative and market-based accountability in Madrid has generated an ambiguous mandate of educational improvement which schools face in different intensities and translate into multiple practices. These range from superficial and instrumental responses to more significant and internalized school practices. More specifically, three main conclusions can be outlined in relation to the empirical results.

First, the results of this research contribute to a reflection of the complex relationship between administrative and market forms of accountability. The combination of administrative and market forms of accountability tend to generate dynamics of interdependence. Market forms of accountability are often tied to administrative tools of external evaluation, for example, when schools use performance indicators to improve their market position or when families base their choices on school performance data. In parallel, schools are also aware of how their capacity to attract and retain students with a certain profile could help them enhance their performance and improve their reputation in the opinion of educational authorities and parents. However, the results also show that market forms of accountability are relatively autonomous and independent of administrative mechanisms. This is the case, for instance, in schools that adopt innovative methodologies to satisfy family preferences without necessarily taking into account learning standards, performance results, or instructional improvement.

Overall, the relationship between market and administrative forms of accountability is ambivalent, and both mechanisms can generate superficial responses among school actors. The evidence provided in this regard also invites us to discuss the results of other studies which suggest that market logics tend to prevail over administrative mechanisms (e.g., Barbana et al., 2020; Maroy et al., 2020). The case study of Madrid indicates that administrative and market forms of accountability operate together to increase external pressures. Such external pressures are experienced unevenly in different schools and generate multiple school practices and responses. Moreover, because this process takes place in a highly competitive and segregated social context, it can, under certain conditions, reinforce practices of second-order competition and dynamics of social closure in education. Future research should explore how, as a response to administrative and market demands, schools activate different logics of action that may generate undesirable policy outcomes in terms of equity.

Secondly, this research has shown that school actors develop complex interpretations of the PBA mandates and can articulate critical, neutral, and engaging understanding with different dimensions of the same policy frameworks, sometimes even in contradictory forms. Hence, schools should be analyzed as complex organizations where their members develop their own understanding of the policy environment according to certain collective and organizational conditions,

but with great levels of discretion depending on subjective perceptions and personal dispositions. Policy precedents are especially relevant when it comes to making sense of policy, a fact that invites us to consider the historical component of sense-making and policy enactment and the accompanying methodological and research implications (as noted elsewhere, e.g., Coburn, 2004). The results presented here suggest that the fact that test results were previously published to promote school choice conditioned the subjective interpretation of PBA among school actors. This was due to earlier forms of performative pressure which persisted even when the test results were no longer published. Indeed, as outlined by cognitive theories, policy enactment is a process of de-codification and re-signification of policy messages into previous cognitive frameworks. However, policy instruments also contribute to building and developing architectures of perception and interpretation. Therefore, accountability policies should not be analyzed as mere regulatory arrangements but as powerful policy instruments able to build and develop imaginaries and modulate school actors' perceptions and behaviors.

Hence, this study invites us to consider the importance of subjective factors in explaining the effects of policy programs on human action. More specifically, the results of this research challenge the assumption of high-stakes systems as unique accountability designs able to generate performative pressures. In the case of Madrid, external pressures prevailed under a lower-stakes design because the perceptions and interpretations of school actors remained attached to the reputational concerns of testing in a quasi-market setting. Therefore, when analyzing schools' external pressures under accountability regimes, we should not only consider the formal policy designs (high-stakes vs low-stakes) but also the subjective perceptions of policy actors regarding such pressures, since, as noted by symbolic interactionism, "if men define situations as real, they are real in their consequences" (see Merton, 1995).

Third, the results confirm that schools adopt different strategies and practices in response to PBA, according to a complex interaction between different factors, namely, the schools' pedagogical approach, their performance culture, the student population, and the subjective perception of external pressures. Indeed, the appropriations of PBA are highly determined by school culture and other contextual elements. Nonetheless, PBA has contributed to the generalization of some school practices in different settings. For instance, competence-based teaching is significantly internalized in the schools studied. This is largely due to the enforcing capacity of the external test but also because teachers assume this approach is desirable and appropriate. However, instrumental practices, such as test preparation activities, also emerge and are generalized despite uneven levels of systematization and intensity. In contrast, promotional activities are more unevenly adopted and are highly reliant on a school's reputation and perceived competitive pressure. Data use is another of the expected responses to PBA, but it appears to be implemented fragmentally and superficially. In many cases, this simply meets legal requirements without being internalized into everyday school life. This is partly because most school actors are hesitant regarding the efficacy and fairness of accountability policy instruments, which is something that reinforces responses that decouple from the accountability mandate.

The results of this research suggest that the decoupling argument partly applies in the case of Madrid, especially with respect to practices of data use. Schools tend to adopt superficial changes without altering the formal structures and instructional dynamics when school actors do not believe in the efficacy and fairness of certain policy instruments. Moreover, the back and forth dynamics regarding the accountability regime, i.e., its low policy sedimentation and coherence, may have contributed to such superficial responses. In this sense, the absence of a stable and a formal scheme of incentives and sanctions, which is characteristic of the erratic trajectory of NPM measures in Spanish public administration, also favors school actors' misbelief in the accountability system and contributes to the prevalence of decoupling dynamics and superficial responses.

This research aims to improve our understanding of accountability policy enactments in different school settings and has important policy and research implications beyond the case study discussed. In terms of policy implications, this research provides new evidence on the prevailing gap between policy expectations and actual policy enactments. Teachers and principals are key policy actors in the enactment of education policies at the school level. Hence, the failure or success of education reforms relies heavily on the way school actors understand the principles, goals and mechanisms of the policies, and tools and instruments of the reforms. Therefore, taking into account the voices and perceptions of teachers and principals in education reform is a critical point, or at least a necessary condition, to favor its success. Reform processes should also take into account the fact that school actors are not exempt from misunderstanding policy messages or reinterpreting policy mandates in instrumental terms. This is one reason why policy makers should consider non-expected results and prevent potential undesirable policy outcomes, especially when these outcomes can have equity effects.

Regarding the research implications of this case study, at least two main issues should be stressed. First, this research reflects on the importance of policy precedents, which should be considered to a greater extent when analyzing contemporary forms of policy enactment. This implies that the chronological and historical dimension of policy process should be taken into account in the methodological designs of implementation and enactment research. The importance of policy precedents is closely related to the subjective dimension of the policy enactment because the pre-existing experiences of policy actors regarding specific policy instruments determine its interpretation. Therefore, another research implication of this work is related to the need for greater consideration of the subjective dimension of policy enactment. Human action can be understood not only as the response to objective conditions of reality but also as the response to the meaning given to such conditions (Merton, 1948). This research has also shown that the adoption and enactment stages of the policy process are intimately related and cannot be artificially disentangled. Therefore, a final research implication is the indication that a proper policy analysis requires a more holistic approach to iteratively address policy adoption and policy enactment as two sides of the same coin.

## References

- Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational researcher*, 36(5), 258–267.
- Ball, S. J. (2003). The teacher's soul and the terrors of performativity. *Journal of education policy*, 18(2), 215–228.
- Ball, S. J., Maguire, M., & Braun, A. (2012). *How schools do policy: Policy enactments in secondary schools*. Routledge.
- Barbana, S., Dumay, X., & Dupriez, V. (2020). Local implementation of accountability instruments in the French-speaking community of Belgium. *European Educational Research Journal*, 19(2), 94–108. <https://doi.org/10.1177/1474904119850964>.
- Bonal, X., & Zancajo, A. (2018). School segregation in the Spanish quasi-market education system: Local dynamics and policy absences. In X. Bonal and C. Bellei (Eds.), *Understanding school segregation: Patterns, causes and consequences of spatial inequalities in education* (pp. 201–221). Bloomsbury. <https://doi.org/10.5040/9781350033542.ch-010>.
- Bolívar, A. (2008). El discurso de las competencias en España: Educación básica y educación superior. *Revista de docencia universitaria*, 6(2). <https://doi.org/10.4995/redu.2008.6268>.
- Boxenbaum, E., & Jonsson, S. (2017). Isomorphism, diffusion and decoupling: Concept evolution and theoretical challenges. *The Sage handbook of organizational institutionalism*, 2, 79–104.
- Braun, A., Ball, S. J., Maguire, M., & Hoskins, K. (2011). Taking context seriously: Towards explaining policy enactments in the secondary school. *Discourse: Studies in the Cultural Politics of Education*, 32(4), 585–596.
- Busuioac, M., & Lodge, M. (2017). Reputation and accountability relationships: Managing accountability expectations through reputation. *Public Administration Review*, 77(1), 91–100.
- Carvalho, L. M., Costa, E., & Sant'Ovaia, C. (2020). Depicting the faces of results-oriented regulatory processes in Portugal: National testing in policy texts. *European Educational Research Journal*, 19(2), 125–141.
- Carpintero, S., & Siemiatycki, M. (2015). PPP projects in local infrastructure: Evidence from schools in the Madrid region, Spain. *Public Money & Management*, 35(6), 439–446.
- Coburn, C. E. (2004). Beyond decoupling: Rethinking the relationship between the institutional environment and the classroom. *Sociology of Education*, 77(3), 211–244.
- Coburn, C. E. (2005). Shaping teacher sensemaking: School leaders and the enactment of reading policy. *Educational Policy*, 19(3), 476–509.
- Diamond, J. B. (2012). Accountability policy, school organization, and classroom practice: Partial recoupling and educational opportunity. *Education and Urban Society*, 44(2), 151–182.
- Dumay, X., Cattonar, B., Maroy, C., & Mangez, C. (2013). The local institutionalization of accountability in education: Network and bureaucratic modes of implementation. *RISE*, 2(2), 99–141.
- Escardíbul, J. O., & Villarroja, A. (2009). The inequalities in school choice in Spain in accordance to PISA data. *Journal of Education Policy*, 24(6), 673–696.
- Falabella, A. (2014). The Performing School: Effects of Market & Accountability Policies. *Education Policy Analysis Archives*, 22(51). <https://doi.org/10.14507/epaa.v22n70.2014>.
- Gawlik, M. A. (2015). Shared sense-making: How charter school leaders ascribe meaning to accountability. *Journal of Educational Administration*, 53, 393–415.
- Gewirtz, S., Ball, S. J., & Bowe, R. (1995). *Markets, choice, and equity in education*. Open University Press.
- Hallett, T. (2010). The myth incarnate: Recoupling processes, turmoil, and inhabited institutions in an urban elementary school. *American Sociological Review*, 75(1), 52–74.
- Landri, P. (2018). *Digital governance of education: Technology, standards and Europeanization of education*. Bloomsbury Publishing.
- Lipman, P. (2004). *High stakes education: Inequality, globalization, and urban school reform*. Psychology Press.
- Maroy, C. (2015). “Comparing Accountability Policy Tools and Rationales : Various Ways, Various Effects?” In *Governing Educational Spaces : Knowledge, Teaching and Learning in Transition*, H.-G. Kotthoff, and Klerides E., dir., 35–58. Boston, MA: Sense Publishers.
- Maroy, C., & Van Zanten, A. (2009). Regulation and competition among schools in six European localities. *Sociologie du travail*, 51, e67–e79.
- Maroy, C., & Pons, X. (2019). *Accountability policies in education*. Springer.

- Maroy, C., Dupriez, V., & Pons, X. (2020). Accountability and datafication in francophone contexts: A reinforcing cognitive state still challenged by local actors. In S. Grek, C. Maroy, & A. Verger (Eds.), *World Yearbook in Education 2021: Accountability and datafication in education*. Routledge.
- Merton, R. K. (1948). The self-fulfilling prophecy. *The Antioch Review*, 8(2), 193.
- Merton, R. K. (1995). The Thomas theorem and the Matthews effect. *Social Forces*, 74, 379.
- Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, 83(2), 340–363.
- Meyer, H. & Rowan, B. (2006). Institutional analysis and the study of education. In H. Meyer & B. Rowan (Eds.), *The new institutionalism in education* (pp. 1–14). Albany, NY: SUNY Press.
- Mittleman, J., & Jennings, J.L. (2018). Accountability, achievement, and inequality in American Public Schools: A review of the literature. In: B. Schneider (Ed.), *Handbook of the sociology of education in the 21st century* (pp. 475–492). Cham: Springer.
- Oliver, C. (1991). Strategic responses to institutional processes. *Academy of Management Review*, 16(1), 145–179.
- Pagès, M., & Prieto, M. (2020). The instrumentation of global education reforms: an analysis of school autonomy with accountability policies in Spanish education. *Educational Review*, 72(6), 671–690. <https://doi.org/10.1080/00131911.2020.1803795>.
- Parcerisa, L. (2016). Modernización conservadora y privatización de la educación el caso de la LOMCE y la Nueva Gestión Pública. *Revista Educación, Política y Sociedad*, 1(2), 11–42.
- Perryman, J., Ball, S., Maguire, M., & Braun, A. (2011). Life in the pressure cooker—School league tables and English and mathematics teachers' responses to accountability in a results-driven era. *British Journal of Educational Studies*, 59(2), 179–195.
- Skedsmo, G. (2011). Formulation and realisation of evaluation policy: Inconsistencies and problematic issues. *Educational Assessment, Evaluation and Accountability*, 23(1), 5–20.
- Spillane, J. P. (2009). *Standards deviation: How schools misunderstand education policy*. Harvard University Press.
- Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research*, 72(3), 387–431.
- Tiana, A. (2011). Análisis de las competencias básicas como núcleo curricular en la educación obligatoria española. *Bordón: Revista de pedagogía*, 63(1), 63–75.
- Verger, A., & Curran, M. (2014). New public management as a global education policy: Its adoption and re-contextualization in a Southern European setting. *Critical studies in Education*, 55(3), 253–271.
- Verger, A., Fontdevila, C., & Parcerisa, L. (2019a). Reforming governance through policy instruments: How and to what extent standards, tests and accountability in education spread worldwide. *Dis-course: Studies in the Cultural Politics of Education*, 40(2), 248–270.
- Verger, A., Parcerisa, L., & Fontdevila, C. (2019b). The growth and spread of large-scale assessments and test-based accountabilities: A political sociology of global education reforms. *Educational Review*, 71(1), 5–30.
- Verger, A., Prieto, M., Pagès, M., & Villamor, P. (2020a). Common standards, different stakes: A comparative and multi-scalar analysis of accountability reforms in the Spanish education context. *European Educational Research Journal*, 19(2), 142–164.
- Verger, A., Ferrer, G., & Parcerisa, L. (2020b). In and out of the 'pressure cooker': Schools' varying responses to accountability and datafication. In S. Grek, C. Maroy, & A. Verger (Eds.), *World Yearbook in Education 2021: Accountability and datafication in education*. Routledge.
- Villamor, P., & Prieto, M. (2016). Reformas hacia la privatización de la educación en la Comunidad de Madrid. *Revista de Sociología de la Educación-RASE*, 9(2), 265–276.
- Yin, R. K. (2003). Case study research: design and methods (ed.). Applied social research methods series, 5.