


**ADVERTIMENT.** L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons:  <https://creativecommons.org/licenses/?lang=ca>

**ADVERTENCIA.** El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons:  <https://creativecommons.org/licenses/?lang=es>

**WARNING.** The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license:  <https://creativecommons.org/licenses/?lang=en>



PhD programme in Sociology  
Department of Sociology  
Faculty of Political Science and Sociology

**Performance-based accountability, educational innovation  
and teachers' work: global and national perspectives on an  
intricate relationship**

Doctoral Dissertation by  
Laura Mentini

Directed by  
Dr. Antoni Verger Planells  
Dr. Antonina Levatino

Barcelona, 24/09/2024

The project leading to these results has received funding from “la Caixa” (ID 100010434) under the agreement LCF/BQ/DR19/11740004.

## ACKNOWLEDGEMENTS

They always say acknowledgements of a dissertation are the hardest part to write.

The opportunity given me by my supervisor, Toni, to apply to La Caixa scholarship and continue my research in a PhD programme is invaluable. I cannot express my gratitude to Toni for forging my scientific and critical mind, always encouraging and seeing where my motivation and spark lies, for supporting and illuminating any opportunity that arose.

Anto, in the hardest moments, her sentence in: “you will get to know yourself, your limits and strengths during your PhD”, always pushed me to go further. Thank you for delving into creative discourses with me, challenging my thoughts and for being a great friend I have discovered.

Thank you, Andrea, Marcel, LLuis, Marjoleine, Natalie, Edgar, Giulia for being the best colleagues with which to share this process and for your great minds. I admire you all.

I formed my critical and scientific mind during the Master programme and deepened it during my PhD at UAB. Thank you to all the GEPS research members and the stimulating and familiar environment you created, to Xavi, Aina, Mauro, Alex, Andrea, Marti for being a source of inspiration, personally and professionally.

Thank you to the Italian scholars for opening my eyes about my own country. Thank you to the interviewees, the school teachers, educators, and students for making me see what was there to understand and the friendly chats that arose often. To the academics met around the world, the discussant and reviewers: you have made me understand what the academic world is practically about. Thank you to the members of the thesis examining board defense for having read my thesis and provided valuable feedback.

Thank you to LaCaixa colleagues Flaminia, Livia, Damiano, Sonia, for being a great Italian team to have around. To LaCaixa staff for the support, and to Tina for being a great coach that helped me organize thoughts, priorities and steps. Thank you to UAB administration, for being flexible, and accommodating when needed.

Thank you to my family for opening arms and being there in the hardest times, and pushing me to “close the circle”. To my lifelong friends, and all those which I encountered in these years, in Rome and Barcelona, for always believing in me, representing a source of distraction and enthusiasm during this long journey. To Luca for encouraging me to start this process and throughout, to Daniele for helping me recognize that research minds are everywhere and Dani for the fundamental support and joy in the last months and days. To myself for the determination and having learnt to be soother with myself during these years.

## **ABSTRACT**

Performance-based accountability (PBA) policies have become increasingly prevalent in the education sector, aiming to improve educational outcomes by holding schools, teachers, and administrators accountable for their performance. Concurrently, there has been a growing emphasis on fostering innovation within educational systems to meet the evolving needs of learners and societies in the 21st century. The intersection between PBA and innovation policies presents a contradictory and multifaceted landscape, shaping the policy ideas of educational stakeholders and influencing enactment dynamics within schools in very different ways.

This doctoral dissertation explores the intricate relationship between PBA and innovation policies in the education sector, and the role of teachers' wellbeing therein. Through an in-depth analysis of policy discourses, school enactment strategies, and teacher emotional responses, the dissertation aims to uncover the interactions and implications of these two mandates, examining both the academic literature at an international level as well as the policy frameworks and school-level attitudes and practices in the specific Italian context.

The dissertation employs a qualitative and multi-method approach combining different research instruments and analytical approaches, each tailored to a specific research question within three distinct sub-studies. This includes (a) a scoping review of the literature (b) semi-structured interviews with key educational stakeholders and decision-makers and document analysis (c) semi-structured interviews with school actors. Methodologically, the research is guided by a multi-sited case study framework to examine the deployment of policies across different policy levels and a realist evaluation approach that helps uncover the mechanisms driving policy implementation and its outcomes in different contexts.

The findings of the research highlight the professional and emotional implications of PBA on teachers, highlighting that PBA can reinforce socially desirable emotions and self-regulation and lead to both positive as well as negative feelings. Findings also underscore emotional consequences of PBA that discourage teachers from engaging in innovative pedagogical practices. The review emphasizes that factors such as school culture, student composition, principal leadership, and individual characteristics can influence how teachers emotionally experience PBA.

The findings related to the policy process underpinning the Italian National Evaluation System (SNV), highlight the assumptions, rationales, and tensions of PBA and autonomy to drive educational innovation in Italy. The research also highlights the fragmented and irregular policy adoption of PBA reforms in the country, the interplay of various interest groups, and the heterogeneous meanings and justifications given to PBA and innovation by key stakeholders in the country.

Empirical evidence from Italian schools finally illustrates the varied range of understandings, filtering and adaptations of PBA and innovation mandates by school actors. PBA policies, while intended to promote self-evaluation and improvement, can in some contexts also hinder innovation in schools due to the performative pressure they generate. The study finally identifies key factors such as actors' sense-making, schools' socio-economic status, organizational culture, and material conditions that mediate the enactment of PBA instruments, contributing to a deeper understanding of how institutional, organizational and individual factors interact to shape educational practices in the context of PBA reforms.

## LIST OF ABBREVIATIONS

COBAS	Confederazione dei Comitati di Base
CGIL	Italian General Confederation of Labour
CBE	Competency-based education
INDIRE	National Institute for Documentation, Innovation and Educational Research
INVALSI	National Institute for the Evaluation of the Education and Training System
NPM	New Public Management
PBA	Performance-Based Accountability
SNV	National Evaluation System
TBA	Test-Based Accountability
WP	White Paper

## TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION .....	11
PBA and innovation policies in education: main contributions and gaps in the literature.....	13
PBA linked to school autonomy and innovation as global education reforms .....	14
Understanding the “paradoxical” relationship between PBA and innovation .....	16
Gaps in understanding the adoption and enactment of PBA policies and innovation .....	17
Contextual background .....	19
Research questions.....	21
Methodology .....	22
Dissertation structure .....	26
References.....	26
Compendium of publications .....	34
CHAPTER 2: TEACHING AS AFFECTIVE LABOUR IN A DATAFIED WORLD: A SCOPING REVIEW OF THE RELATIONSHIP BETWEEN PERFORMANCE-BASED ACCOUNTABILITY AND TEACHERS’ EMOTIONS .....	35
Introduction .....	35
Data and methodology .....	36
Findings .....	39
Exploring teacher emotions: emotional rules and emotional labour in changing educational contexts .....	39
Teacher emotions under PBA regimes .....	41
The emotional effects of PBA .....	41
Mechanisms behind the emotional effects of PBA.....	42
Dealing with the emotional effects of PBA: teacher agency and coping strategies .....	44
Between PBA and teacher emotions: factors intensifying or weakening the emotional effects of PBA .....	45
Macro-level factors: systemic environment and cultural norms .....	45
Meso-level factors: the importance of school culture, student composition and internal relationships .....	46
Teacher-level factors: professional and personal characteristics .....	47
Discussion and conclusions .....	49
References.....	53
CHAPTER 3: A “THREE-LEGGED MODEL”: (DE)CONSTRUCTING SCHOOL AUTONOMY, ACCOUNTABILITY, AND INNOVATION IN THE ITALIAN NATIONAL EVALUATION SYSTEM.....	60
Introduction .....	60



A genealogy of school autonomy and accountability reforms in the Italian context .....	63
Theoretical framework .....	65
Politico-administrative regimes in the recontextualization of education reforms .....	65
The relevance of “ideas” and the discursive and intricate nature of policymaking.....	67
Methodology .....	68
Findings .....	72
The complex relationship between autonomy, accountability, and innovation: Underlying assumptions and intrinsic contradictions.....	72
Assumption 1: If schools have autonomy in curricular and organizational matters, they will innovate and introduce more change .....	72
Assumption 2: If schools are evaluated externally through standardized tests, they will be more effective in using school autonomy .....	73
Assumption 3 & 4: If schools use external test results for formative feedback, they will innovate and improve, and if schools innovate in their teaching strategies, they will obtain better results in external tests .....	75
Intervention: An equilibrated “three-legged system”?.....	77
Rationales behind the SNV: A solution in search of many problems.....	79
The context of intervention: A theory of no-change?.....	83
Structural features of the Italian educational context.....	84
School actors’ personal and professional characteristics .....	84
Cultural aspects and ingrained beliefs.....	86
The achievement of the SNV outcomes: Misunderstanding and misuse .....	87
Conclusions .....	90
References.....	93
 CHAPTER 4: NEGOTIATING BETWEEN THE ACCOUNTABILITY AND THE INNOVATION MANDATES: EVIDENCE FROM ITALIAN SCHOOLS.....	 99
Introduction .....	99
The Italian context and local meaning of educational reforms.....	102
Assessment and accountability policies .....	102
The Italian policy approach to innovation .....	104
School actors’ reception of the double-mandate .....	105
Theoretical framework .....	106
Making sense of the policy messages by school actors .....	106
Institutional filtering into local realities .....	107

The mediating role of school's contexts.....	107
Methods.....	109
Sampling procedure.....	110
Data collection and participant selection.....	113
Access to schools.....	113
Interview structure and data analysis .....	114
Results.....	115
(1) Privileged school settings .....	115
High performance results and embracing innovation without tensions.....	116
(2) Middle - income school settings .....	120
Homogenous contexts: Balancing Innovation and Performance amid internal conflict	121
Heterogeneous contexts: Divergent approaches and malleability of the double mandate	124
(3) Disadvantaged school settings.....	127
Experiencing contradictory pressures: Filtering and Decoupling policy.....	128
Discussion and conclusions .....	132
References.....	135
CHAPTER 5: SUMMARY OF THE MAIN RESEARCH FINDINGS .....	142
CHAPTER 6: CONCLUSIONS.....	150
Theoretical insights: integrating cognitive approaches and institutional theories in policy	
adoption and enactment.....	150
Methodological insights.....	151
PBA and innovation in a Southern-European setting: Learning about the Italian experience	152
Understanding the role of discursive ideas and sense-making.....	154
Understanding the role of institutional contingencies and local school contexts .....	155
Limitations and future lines of inquiry .....	156
Policy relevance.....	157
References.....	160

## **TABLE OF FIGURES**

Table 1: Overview of the three sub-studies .....	25
Table 2. Overview of documents.....	70
Table 3. Overview of key actors interviewed.....	71

Table 4. Characteristics of selected schools, by income and performance (Source: own elaboration).....	112
Figure 1. Flow chart of the scoping review (Source: adapted from Page et al. (2021) .....	38
Figure 2. Syntax.....	59

## CHAPTER 1: INTRODUCTION

In recent years, there has been a growing recognition of the interplay between performance-based accountability (hereafter PBA), school autonomy, and innovation in global education agendas. Policymakers are increasingly seeking to balance accountability requirements, typically manifested in standardized tests and PBA instruments, with efforts to promote innovation and flexibility within educational systems (Fahey & Koester, 2019; Verger & Parcerisa, 2017; Vincent-Lancrin, 2019).

The relationship between PBA, school autonomy, and innovation represents a critical yet understudied area within the field of educational policy and practice. While PBA policies are often framed as mechanisms to enhance accountability and transparency (OECD, 2013a), they can also exert significant influence on the degree of autonomy afforded to schools and educators (Farvis & Hay, 2020). Furthermore, previous research found that PBA shapes educator's identities, practices and experiences in different ways (e.g., Avalos et al., 2020; Ball, 2003; Holloway & Brass, 2017; Lewis & Hardy, 2018), and that the tension between accountability and innovation poses challenges for school teachers seeking to foster innovative and creativity in classrooms (Appel, 2020; Falabella, 2020; Knight, 2020), often generating a conflict between the political agenda and their appropriation in schools (Barbana et al., 2020; Candido, 2019).

Understanding how these reform agendas interact is essential for informing policy and practice in education. By examining the relationship between PBA, school autonomy, and innovation, and the role of contextual mediating factors, this dissertation aims to contribute to a deeper understanding of how PBA and innovation policies are conceptualized in policy processes, how they unfold in schools, and how they influence teachers' work and emotions. Moreover, research has shown that accountability is a global phenomenon, but how it is understood and enacted is locally contextualized (e.g., Gunnulfsen et al., 2018; Holloway et al., 2017; Quilabert, 2024).

Italy offers an interesting case to explore the aforementioned dynamics. Despite its historically centralized educational system (Mattei, 2012), the country has witnessed moves toward decentralization and greater school autonomy (Colombo & Desideri, 2018) with the aim to incentivize pedagogical innovation and teaching flexibility. Moreover, innovation is inherently part of the National Evaluation System (hereafter SNV) (Faggioli & Mori, 2018), a comprehensive accountability apparatus which integrates external accountability and internal school self-

evaluation mechanisms to foster reflexivity and change in pedagogical practices (Paletta et al., 2020), yet through an incoherent, fragmented and irregular PBA implementation (Capano & Lepori, 2019; Kickert, 2007).

To comprehensively examine the interplay between accountability and innovation, this doctoral dissertation aims to:

1. Explore how teachers experience PBA, and the specific role that emotions have on the enactment of PBA from an interdisciplinary and international perspective.
2. Analyze the policy framework underlying recent autonomy and PBA reforms in Italy, particularly analyzing the assumptions and tensions behind the National Evaluation System (SNV) in fostering innovation.
3. Investigate how school actors interpret and respond to the dual mandates of PBA and innovation in the Italian context, considering the schools' context as a mediating factor.

The study adopts a multi-method approach, including a scoping literature review, qualitative interviews with a broad range of key education informants, document analysis, and qualitative interviews to school actors. By exploring both literature contributions as well as empirical outcomes, and contextualizing global education reforms in Italy's historical trajectory and institutional landscape, this research seeks to provide insights into the relationship between PBA policies, pedagogic and educational practices, and institutional dimensions. Moreover, through a combination of theoretical perspectives, including new institutional theories and constructivist approaches, the dissertation aims to provide a comprehensive analysis of the complexities involved in the policy adoption and enactment of these policies.

In this introductory chapter, I will first situate the study in the existing literature, by giving a brief overview of the main contributions of research on PBA and innovation reforms. Additionally, I will outline the main gaps within this body of literature. Following this, I will describe how this dissertation aims to address and contribute to filling some of these gaps. Subsequently, the rationale behind choosing Italy as an empirical case study for this research is explained, followed by a presentation of the research questions and methodology employed. This introductory chapter concludes with an outline of the structure of this dissertation.

## **PBA and innovation policies in education: main contributions and gaps in the literature**

In recent decades, the governance of educational systems has witnessed a paradigm shift towards increased autonomy and accountability measures, with an emphasis on PBA policies. PBA in education generally refers to the monitoring of educational outcomes through external assessments (Fahey & Koester, 2019), the definition of learning standards and the establishment of school accountability mechanisms to monitor teachers' performance and promote competitive pressures among schools (Teltemann & Jude, 2018; Verger et al., 2019). Test-based accountability systems (TBA) are also part of PBA, focusing on holding educational institutions accountable for student outcomes, through external evaluation of students' competences in certain areas of knowledge by means of national large-scale assessments (NLSAs) (Verger et al., 2019)<sup>1</sup>.

At the same time, innovation emerges as a policy imperative in education (Fahey & Koster, 2019; European Commission, 2018), calling schools to innovate in organizational and pedagogical approaches, and respond to the challenges of the 21st century (Stoll & Kools, 2017). Learner-centered (LCE) and competency-based education (CBE) are considered part of these innovative approaches (Peterson et al., 2018), emphasizing the application of skills and knowledge to real-world challenges, and lifelong learning over traditional, teacher-centered methods (Sahlberg, 2006). While educational innovation is a multilevel phenomenon (EU Commission, 2018), a dynamic and complex process (Sahlberg, 2006) and contextually different according to countries (Licht et al., 2017), it typically involves the introduction of novel pedagogical or organizational practices aimed at enhancing student outcomes (European Commission, 2018). Moreover, innovation in education generally refers to two main types of innovations: *process innovations*, focused on changes in teaching and learning, including curricular and pedagogical changes, and *administrative innovations* which involve changes in the organization, management or governance of a school (Lubienski, 2009). Innovation in education is associated either with pedagogical (i.e. curriculum, assessment, pedagogical approach) or organizational (i.e. leadership, teacher collaboration, networks) innovation (Stoll & Kools, 2017). In terms of classroom innovations, according to the OECD,

---

<sup>1</sup> However, despite PBA being prominent in the global education agenda, we should recognize that “accountability” is a multidimensional concept, encompassing test-based, outcome-based logics, but also professional or internal forms of accountability (Elmore, 2005a; Firestone 2004) and is enacted differently according to the contexts (e.g., Barbana et al., 2020; Candido, 2019; Gunnulfson et al., 2018; Holloway et al., 2017; Quilabert., 2024).

three main components underpin innovation in teaching and learning (Looney, 2009). An “innovative” teaching is based on student-centered, constructivist, and student active engagement as opposed to a traditional teaching based on direct transmission. Project-based or problem-based learning (PBL) and the personalisation of student learning are part of such a pedagogical framework in teaching and learning (Pedrò, 2023). Secondly, innovative teaching focuses on the development of thinking and reasoning skills, and skills for life-long learning such as creativity, entrepreneurship, collaboration, problem solving, ICT literacy. This also implies increased attention to the development of competences, with emphasis on transversal, STEM or digital competences, as opposed to transmission and memorization of contents in curriculum, (Pedrò, 2023). Finally, educational innovation is closely linked to school autonomy, whereby schools tailor their approaches and personalize teaching to local contexts and student needs as opposed to top-down policies dictating educational content (Looney, 2009). Administrative types of innovations, which include both marketing and organizational innovations are however not directly linked to changes in the classroom, but involve changes in the behavior of a school and may impact the position of a school within a school market (Lubienski, 2009).

### PBA linked to school autonomy and innovation as global education reforms

Global education reforms are intended as “traveling policies” (Ozga & Jones, 2006): this is when a particular educational reform moves across territories and appears simultaneously in a number of countries. PBA is increasingly adopted globally, implemented across various contexts, yet yielding disparate effects on education systems (Smith, 2014; Verger & Parcerisa, 2017). This reform model typically coupled with school autonomy, involves decentralizing organizational and pedagogical decisions to lower government levels and schools, and the simultaneous adoption of TBA, where schools are asked to give account of their results to the public administration, focusing on students’ learning outcomes in certain areas of knowledge, measured by external standardized tests. Through this set of policies, governments advocate for granting schools greater autonomy, based upon their willingness to undergo external assessments and embrace accountability measures (Fahey & Koester, 2019; Verger & Parcerisa, 2017).

In this context, many European and OECD education systems have embraced PBA policies as a means to enhance the quality, equity, efficiency and innovation of education systems (Lubienski, 2009; Sahlberg, 2006), by linking educational outcomes to various forms of assessment, including standardized testing and data-driven evaluations. At the same time, schools are provided with the

necessary autonomy to introduce these changes and promote innovation processes (European Commission, 2015; OECD, 2013). The underlying assumption is that "balanced school autonomy, with built-in accountability mechanisms, improves schools' capacity for innovation" (European Commission, 2018: 15). Accountability mechanisms are seen as necessary for balancing increased school responsibilities and ensuring students meet centrally defined standards (OECD, 2013a). In addition, granting schools greater autonomy is believed to enhance their ability to adapt their learning strategies to the local contexts and individual student needs, fostering innovation and improvement (Looney, 2009). By providing information to school actors on their educational outcomes, PBA policies are expected to promote effective teaching and learning models (Herman, 2004), at the same time incentivizing actors to challenge standard practices (Fahey & Koester, 2019). Through performative pressures, innovation in schools can be used to disseminate new pedagogical and/or curricular practices (Lubienski, 2009), or challenge inequalities and move teachers to expose students to high-quality instruction through content and differentiated pedagogy (Diamond, 2007).

At the same time, in recent decades, education innovation emerged as a policy imperative based on the need to reform education systems to better suit contemporary demands. Countries are asked to innovate in their organization and pedagogies (Stoll & Kools, 2017), and to respond through education to the 21<sup>st</sup> century education challenges, pushed by demands of the economy, globalization and rising interest in information, knowledge and communication (Salhberg, 2006; Tromp, 2007)<sup>2</sup>. Thus, a social consensus around the need to explore new school models becomes prominent (Pedrò, 2023), where the priority is placed on development of transversal and transferable competences, such as practical skills, citizenship education, conceptual understanding, problem-solving abilities, and socio-emotional skills (Tromp, 2017). Moreover, a global trend towards progressive pedagogies emerges, with a notable shift from traditional teacher-centric approaches to student-centered or learner-centered pedagogies (LCP), and the emphasis on active learning, student motivation and autonomy, metacognitive and affective qualities as opposed to memorization of facts and the acquisition of specific knowledge (SITES, 2009).

---

<sup>2</sup> External factors such as demographic and social changes, the demand for skills, technological changes and international competition on student learning contribute to the need to promote innovation in education (Pedrò, 2023).



## Understanding the “paradoxical” relationship between PBA and innovation

While the rationale behind PBA policies is to improve educational quality and student achievement, evidence suggests that PBA reforms yield contradictory effects on education systems and innovation. The simultaneous effort to decentralize power at local level on the one side, and the legitimation of central control and standardization on the other (Karlsen, 2000) have been characterized as a “paradox” (e.g., Falabella, 2014) that brings to light tensions between autonomy, accountability and innovation (Fahey & Koester, 2019; Knight, 2020; Looney, 2009).

With regards to teaching practices and instructional approaches, research has highlighted the role of PBA in fostering innovative approaches, by communicating good models of teaching and learning (Herman, 2004), and motivating school actors to challenge conventional practices (Fahey & Koester, 2019). On the other hand, however, the imperative to achieve favorable outcomes in standardized assessments may deter educators from taking pedagogical risks or engaging in creative practices (Falabella, 2020; Sahlberg, 2009; Knight, 2019). Accountability measures, while intended to foster innovation, can limit the adoption of alternative practices, collaboration, and creativity in teaching and learning (Lechasseur, 2015; Sahlberg, 2009), as teachers feel pressured to prioritize achieving good results in standardized assessments (Falabella, 2020; Knight, 2020). External pressure and growing expectations on teachers may lead to more comfortable attitudes rather than uncertainty and unrecognized effort. Moreover, testing and accountability shape schools and teacher’s priorities with regard to content and the attention paid to different subject areas (Spillane & Diamond, 2004). A narrow emphasis on standardized test scores and basic skills frequently leads to 'teaching to the test', curriculum content alignment, and fragmentation of knowledge (Au, 2007; Barret, 2009; Sahlberg, 2009). The pressure of external accountability may require teachers to prioritize test-focused material, test-taking skills (Valli & Buese, 2007), and a teacher-centered transmissive instruction (Barrett, 2009) instead of fostering holistic understanding and higher-order thinking (Au, 2007). Overall, a tight framing of teaching, focusing on student outcomes, with little room for discretion, innovation, and risk, hampers possibilities of system innovation (Sorensen & Robertson, 2017), and can obscure attention from the processes that are responsible for generating innovations in the first place (Fahey & Koster, 2019).

Previous studies have also explored the socio-emotional dimensions of PBA, shedding light on teachers' experiences, emotions, and well-being within accountability-driven environments. Research has shown that PBA reforms may have significant emotional effects among public sector

employees (Dubnick, 2005). Within the realm of education, research underscores the emotional toll of performance-based accountability on teachers, which may impact their ability to effectively navigate policy mandates and promote classroom innovation (Holloway & Brass, 2017; Braun & Maguire, 2020; Lambert & Gray, 2021), along with the implications on educators' identities, professional experiences, and the teaching profession (Ball, 2003; Day & Smethem, 2009; Day, 2018; Holloway & Brass, 2017; Lewis & Hardy, 2018).

In addition, scholarly inquiry into PBA policies in education has yielded significant insights into the role of school actors' opinions in shaping the enactment of PBA in education, including those teachers and school leaders (see Ball, et al. 2012; Braun, et al. 2010; Coburn, 2001; Diamond & Spillane 2004). While policymakers may advocate for accountability measures as a means of modernizing the system, improving educational equity, quality and efficiency (see Hursh, 2005; Verger et al., 2019), school leaders and teachers can have divergent understanding about the importance and implications of these policies for teaching and learning, shaping policy enactment and schools' responses in ways that differ significantly from the policy intentions, and may even contradict key assumptions of the theory of change underpinning such reforms (e.g. Barbana et al., 2020; Candido, 2019).

Finally, the outcomes of PBA on innovative practices in schools may vary depending on contextual factors such as the schools' socioeconomic status or performance levels. Authors underscore that the impact of PBA policies is not uniform across all school contexts (Diamond, 2007; Mittleman & Jennings, 2018). The socio-economic context of schools can significantly shape teachers' practices in response to PBA, by encouraging teachers in low-performing schools to rely on didactic pedagogy (e.g., lecture, seat work, memorization, recitation) (Diamond, 2007), categorize students based on test data (Hardy et al., 2019), or customize and adjust instruction and teaching methodologies according to student's academic performance (Mittleman & Jennings, 2018).

## Gaps in understanding the adoption and enactment of PBA policies and innovation

Despite the expanding body of research on PBA policies, significant gaps persist in our understanding of the adoption and enactment of PBA policies and innovation. Firstly, research indicates that performance monitoring and PBA have altered teachers' practices, identities, and

experiences in the education sector (e.g. Day, 2018; Holloway & Brass, 2017; Lewis & Hardy, 2018). Given the emotional nature of the teaching profession (Chávez, 2021; Perold et al., 2012), these changes influence teachers' relationships and their response to educational reforms. Hence, understanding how emotions affect and are affected by PBA is crucial for comprehending how teachers manage the pressures of these policies.

Secondly, diverse theories have attempted to explain the reasons why certain policies are adopted in different contexts in a process of globalization and increasing territorial interdependencies (e.g. Ochs & Phillips, 2002; Ozga & Jones, 2006; Steiner-Khamsi, 2014). In this context, uncovering the rationales behind the adoption of PBA and innovation in education, through an analysis of the politico-administrative regime (see e.g., Pollitt, 2007; Pollitt and Bouckaert, 2011; Verger et al., 2019), is deemed relevant. Also, understanding the role of ideas in defining the issues that define the policy agenda (Andreas et al., 2022) as well as the complexity of policymaking to explain institutional dynamism and public policy change (Cairney, 2012) represents a promising avenue for research. Furthermore, the implementation of policy is multifaceted and context-dependent (Holloway et al., 2017), thus the logics of accountability implementation can vary according to the historical backgrounds and structures of the education system. Still, while prior research has examined the enactment of PBA policies across diverse countries (e.g., Barbana et al., 2020; Candido, 2019; Falabella, 2020; Gunnulfson et al., 2018; Quilabert., 2024), there is still a scarcity of specific studies focusing on southern-European contexts. Within this gap, Italy emerges as an understudied yet pertinent case, offering insight into how, at both policy and school level, education stakeholders conceptualize and navigate the intertwined accountability and innovation imperatives, in a country with a specific managerialist and bureaucratic administrative tradition (Capano & Lepori, 2019; Hall et al., 2015; Verger et al., 2019).

The congruence between policy intentions and actual outcomes of PBA remains another subject of inquiry. In Italy, despite the centrality of autonomy, accountability and innovation, there is a research gap in assessing the rationales behind the adoption of the recent accountability policies, as well as the effectiveness of the accountability tools adopted (Fondazione Agnelli, 2014). This is linked to the relative youth of the Italian National Evaluation System (SNV), which has completed its first accountability cycle and the social reporting in 2019. As such, there has been limited opportunity to assess the alignment between the intended policy goals of the SNV and its actual outcomes in schools.

Lastly, despite extensive research on the deployment and outcomes of PBA, an examination of the complex interplay between accountability policies and innovation is lacking. Questions persist regarding the factors and mechanisms that explain the adoption and enactment of PBA policies in education and the extent to which they might influence or hinder innovative practices in schools. Indeed, the meaning and real goals of PBA on educational innovation are blurred and multiple, and the innovation culture, its forms and features remain unexplored in educational settings (Syariff et al., 2020; Hofman et al., 2012). In addition, actors' meaning-making of a new policy is useful, but is probably insufficient to explain the recontextualization and enactment of PBA policies, why they are adopted in a specific context, how they are transformed and adapted by school actors, and how they result in a range of responses. Thus, understanding the contextual factors that shape the adoption and effects of PBA policies is imperative. Exploring the context of policy enactment (Ball et al., 2012) helps comprehend how schools divergently navigate and address policy expectations, and the extent to which they facilitate or hinder the realization of expected policy objectives.

## **Contextual background**

The dissertation is situated within the broader context of global education reforms, with a particular focus on Italy, as a Southern-European case study.

The first reason why Italy forms a particularly relevant case for this object of study relates to its Napoleonic administrative system legacy (Hall et al., 2015; Verger et al., 2019). For many years, the Italian education system has been characterized by a centralized and bureaucratic structure (Grimaldi & Serpieri, 2012; Mattei, 2012). Italy has been perceived as one of the last European countries to embrace New Public Management (NPM) reforms (Hood & Peters, 2004), lagging behind in implementing educational reforms related to evaluation and innovation (Kickert, 2007; Barzanò & Grimaldi, 2012). Traditionally, its educational model is based on the predominance of transmissive teacher-centered practices (Bifulco et al., 2010; Ferrer-Esteban, 2016; OECD, 2013b), and is characterized as knowledge-based rather than skill-based (Piro, 2016). OECD evaluations revealed that Italian teachers, particularly at the lower secondary level, employ traditional teaching practices, such as frontal teaching, over innovative methods like simulations or cooperative learning (OECD, 2013b). The prevalence of traditional teaching practices, including directive didactics, hinders the adoption of project-based activities and other forms of active participation by students (De Sanctis, 2010).

The second reason is that, notwithstanding its centralized and traditional legacy, since the late 1990s, the Italian education system has undergone complex and contested processes of restructuring and reculturing (Grimaldi & Barzanò, 2014). This included the introduction of school autonomy and decentralization and the implementation of school, staff, and student evaluation to modernize governance and improve efficiency. Accountability and student evaluations were seen as solutions to the country's educational challenges (Barzanò and Grimaldi, 2012). Concerns about efficiency and quality have been highlighted by international organizations, leading to a desire and necessity to adhere to international norms and discourses on educational governance (Verger et al., 2019). Additionally, since the last decade, there has been a growing interest in innovation discourses and initiatives. The 107/2015 law ("The Good School Reform"), facilitated methodological innovation and teaching flexibility within the framework of school autonomy, encouraging personalized educational plans, interdisciplinary collaboration, and competence-based education. Educational innovation in Italy revolves around methodological changes, including cooperative learning, project-based workshops, and assessment innovations, aimed at shifting away from traditional frontal teaching models (Biondi et al., 2009).

The third reason is that the Italian accountability system is rather unique, combining external and internal school accountability mechanisms to foster improvement processes (Paletta et al., 2019) with an explicit focus on school improvement and innovation (DPR 80/2013). Accountability and innovation are deeply intertwined in the PBA system, with emphasis on changing professional practices and the ability of schools to adapt to their specific contexts (Paletta et al., 2019). It is a relatively "soft" accountability model, expected to lead to a process of continuous reflection and improvement in organizational and pedagogical aspects (Faggioli & Mori, 2018), with no material consequences attached to the performance results.

The fourth reason relates to the public perceptions and erratic policy implementation process of PBA policies in the country. In Italy, policies aimed at evaluating the educational system have only recently become a systemic routine, primarily due to a "policy impasse" stemming from opposition and contestations from unions and collegial bodies (Fondazione Agnelli, 2014). The Italian narrative is characterized by teacher resistance and counteraction, with protests and opposition to invasive forms of evaluation and accountability that challenge professional autonomy (Barzanò & Grimaldi, 2013). With the advent of neoliberal managerialist discourse and pressures, Italy witnessed a "war of discourses" (Serpieri, 2009) characterized by a dichotomous logic between

evaluation advocates and opponents (Barzanò & Grimaldi, 2013). In Italy, teacher and school evaluations have historically been viewed negatively (Martini & Papini, 2015), seen as intruding professional autonomy from external demands. The national tests have faced rejection and criticism in Italy, with protests and boycotts organized by civil society, teachers, and students (Barzanò & Grimaldi, 2013).

This last aspect is also related to the conflictual policy decision process which occurs in Italy in relation to comprehensive reforms of the school system or changes in the main policy instruments. Authors argue that decisions related to institutional autonomy of schools, teacher careers or student assessment are highly conflictual and produce tangible results only in the long term (Capano et al., 2022). Moreover, multiple and competing sources of legitimacy in Italy are important to the decision (Galanti & Lippi, 2023), such as the role of the interest groups appears prominent during the «Good School» (Buona Scuola) Reform promoted by Renzi Cabinet (Capano & Terenzi, 2019) or the cohesive and powerful role of “the insiders of the school systems” such as the teachers, the principals, the unions, and the bureaucracies that act as a dominant advocacy coalition that matter for policy success (Capano & Lippi, 2017; Malandrino, 2021).

Overall, the mix of bureaucratic culture and managerialism makes the Italian case interesting to explore. The performance-based policy implementation, compared to the anglo-american accountability setting, is erratic and its design incoherent (Capano & Lepori, 2019), which makes the case quite unique.

## **Research questions**

This dissertation seeks to examine the adoption, enactment and mediating factors of PBA policies on innovation, with a particular focus on the Italian context. In doing so, I aim to assess the ways in which the relationship between PBA and innovation emerges in policy discourses, school enactment dynamics and teacher emotions. This dissertation addresses the following overarching research question:

*How are PBA, school autonomy, and innovation policies interlinked at multiple levels, and how do they affect teachers' work, wellbeing and practices?*

This overarching research question is examined through three sub-studies, each of which focuses on different scales in the analysis of PBA policies, and is guided by a specific research question:

- (1) *How are PBA policies experienced by teachers and, specifically, how do these policies affect them at the emotional level internationally?*
- (2) *What are the rationales behind different policy instruments' selection and assemblage in the context of the accountability agenda in Italy?*
- (3) *How do school actors in Italy experience and make sense of the accountability and innovation mandates, considering the mediating role of the local school context?*

While the first question focuses on teachers' experiences in a range of different countries, the second question explores the policy adoption of PBA in Italy, and the third question finally explores the policy enactment at the school level in the Italian context. Thus, whereas the first question focuses on exploring what is already known about how teachers experience PBA internationally, the following questions analyze these and other aspects in the specific case of Italy. Nonetheless, while each sub-study focuses on different levels and focus of analysis, it is important to emphasize the assumed interconnectedness of these different scales in the adoption and enactment of accountability and innovation policies. In other words, while the different sub-studies move from an international to a national to a local school geographical scope of analysis, and focus of different thematic locus of analysis (teachers, policy and school), the deployment of accountability policies does not occur in a top-down manner, but rather, policy design, policy enactment and contextual contingencies interact in complex and multiple ways.

## **Methodology**

To address these research questions, this dissertation relies on different research methods. A scoping review of the literature is used to provide an overview of the main concepts underpinning the research topic and identify studies carried out and the outcomes they have generated (Colquhoun et al., 2014). Qualitative methods, including semi-structured interviews and document analysis, are used to explore key informants' perspectives and within-schools enactment processes.

Moreover, the dissertation follows a multi-sited case study research (Bartlett & Vavrus, 2014; 2017), incorporating simultaneous attention to different policy levels alongside a horizontal analysis of how policies unfold across different scales and across different school contexts. Case studies (Yin, 2009) are particularly valuable for in-depth exploration of real-life phenomena, observing effects in real contexts, understanding the role of contexts, and focusing on actors' perspectives. A multi-sited case study poses further attention to how a policy operates at different scales, however, not equating these scales to bounded categories or geographical sites (such as the spatial dichotomy global/local), but as a “metaphoric reminder to conduct research across concatenation of sites” (Bartlett & Vavrus, 2014: 133). In this sense, the research incorporates different elements: a “horizontal” attention of the state, schools and teachers as units of analysis, and a comparison across different school contexts; a “vertical” attention to international, national and local school sites; and, in a minor way, the research also gives attention to a “longitudinal” axis, by conducting the research of education policies in a specific time, and by giving attention to the policy process and trajectory that occurred in the country until the most recent configuration of the PBA system.

The study also employs a realist evaluation approach, as introduced by Pawson & Tilley (2004), to provide a methodological lens for the research. This theory-driven approach to evaluating programs and interventions seeks to understand not only whether a program works but also how, why, and in what contexts it works. At its core, realist evaluation aims to uncover the underlying mechanisms that generate outcomes, recognizing that interventions work through various mechanisms in different contexts.

Alongside the common approaches adopted in the dissertation, in each substudy a specific analytical strategy is adopted, that responds to the specific research question. In the first substudy, I utilize a scoping review strategy to explore the relationship between emotions and PBA in the teaching profession. This strategy allows to map out key bodies of literature on a topic by assessing existing studies that address the review questions (Petticrew & Roberts, 2006) and provides an overview of the main concepts, types of studies and outcomes related to a research topic (Colquhoun et al., 2014). The research protocol involved a literature search using the SCOPUS database, which resulted in 127 articles, followed by a screening of each article to be included in the analysis and a detailed review of the articles following specific inclusion criteria, which resulted in a final sample of 63 articles. Using an inductive analytic strategy, the articles were finally sorted



into broad themes and subthemes to reflect each article's contributions regarding the link between PBA and teacher emotions.

In the second substudy, to explore the selection and assemblage of policy instruments, I analyze two main sources of information: policy and technical documents and semi-structured interviews. Eight key documents were analyzed, selected for their relevance to the research questions and twelve interviews with key educational stakeholders were conducted between June 2021 and January 2022. Participants were selected using purposive (9) and snowball sampling (3). Data analysis was performed using ATLAS.ti software, combining inductive and deductive approaches. More specifically, I used a qualitative, interpretive approach combined with thematic analysis to identify relationships between themes (Boyatzis, 1998) and the concept of the "theory of change" as an analytical tool, which helps make explicit the assumptions underlying a change project (Reinholz & Andrews, 2020).

In the third sub-study, to analyze how schools uniquely respond to the accountability and innovation policy mandates, I employ a case study approach (Yin, 2009) focusing on twelve lower-secondary and public schools in Rome. The schools were selected through a sampling process based on different steps and two main variables: performance in standardized tests and average income of the urban area of the school. I collected data through semi-structured interviews with principals, management teams, and teachers, selected through a purposive sampling strategy (Patton, 2015). Additionally, I used documentary analysis to triangulate responses, using publicly available school documents and fieldnotes. To analyze the data, I employed an ideal case analysis, intended as generalizations or mental representations of social phenomena that help make reality understandable, following a seven-step methodology (Stapley et al., 2022). The process involved grouping cases or participants into types based on common features.

An overview of each sub-study, and the data sources and analytical strategy adopted in each of them, is provided in Table 1.

Table 1: Overview of the three sub-studies

<i>Sub-study</i>	<i>Research question</i>	<i>Data sources</i>	<i>Data analysis</i>
1	Chapter 2: <i>How are PBA policies experienced by teachers and, specifically, how do these policies affect them at the emotional level internationally?</i>	63 included articles published between 2003 and 2021, based on SCOPUS literature search	Inductive analytic strategy
2	Chapter 3: <i>What are the rationales behind different policy instruments selection and assemblage in the context of the accountability agenda in Italy?</i>	8 key policy and technical documents, 12 semi-structured interviews with key educational stakeholders	Inductive and deductive approaches, thematic analysis through ATLAS.ti, “theory of change” as analytical tool
3	Chapter 4: <i>How do school actors in Italy experience and make sense of the accountability and innovation mandates, considering the mediating role of the local school context?</i>	7 semi-structured interviews with school leaders; 20 semi-structured interviews with teachers, documentary analysis of publicly available school documents, fieldnotes.	Ideal case analysis, content analysis with ATLAS.ti

## Dissertation structure

This dissertation is structured as follows: Chapter 2 delves into the first sub-study (publication A), followed by Chapter 3, which discusses the second sub-study (publication B). Chapter 4 then explores the third sub-study (publication C). Moving forward, Chapter 5 provides a synthesis of the primary findings of this dissertation, providing an answer to the research question. In the concluding Chapter 6, the methodological, empirical, and theoretical contributions are highlighted, along with a discussion on the limitations of the dissertation and avenues for future research. This chapter also concludes with reflections on the policy implications derived from the findings.

## References

- Appel, M. (2020). Performativity and the demise of the teaching profession: The need for rebalancing in Australia. *Asia-Pacific Journal of Teacher Education*, 48 (3), 301–315.
- Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher*, 36 (5), 258–267.
- Andreas B, Fernie S and Dainty A (2022) Understanding policy and change: Using a political economy analysis framework. *Construction Management and Economics* 40: 865–883.
- Avalos, M. A., Perez, X., & Thorington, V. (2020). Comparing secondary English teachers' ideal and actual writing practices for diverse learners: Constrained professionalism in figured worlds of high-stakes testing. *Reading & Writing Quarterly*, 36(3), 225–242.
- Ball, S. J. (2003). The teacher's soul and the terrors of performativity. *Journal of education policy*, 18(2), pp. 215-228.
- Ball, S., 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.
- Barrett, A. M. (2007). Beyond the polarization of pedagogy: Models of classroom practice in Tanzanian primary schools. *Comparative Education*, 43(2), 273–294.

- Bartlett, L., & Vavrus, F. (2014). Transversing the vertical case study: A methodological approach to studies of educational policy as practice. *Anthropology and Education Quarterly*, 45(2), 131–147.
- Bartlett, L., & Vavrus, F. (2017). *Rethinking Case Study Research. A Comparative Approach*, Routledge.
- Barzanò, G., & Grimaldi, E. (2012). “Policy” valutative e contesti di applicazione. *Caratteristiche Procedurali. nECPS - Educational, Cultural and Psychological Studies*, n06(6), 159–189.
- Barzanò G and Grimaldi E (2013) Discourses of merit. The hot potato of teacher evaluation in Italy. *Journal of Education Policy* 28(6): 767–791.
- Bifulco L, Monteleone R, Mozzana C, et al. (2010) The institutional mapping of Italy’s education system: Europeanization, centralization and regionalization. In: *Work.Able 3rd Workshop - FP7 EU research project “Making Capabilities Work,”* Lausanne, 17–19 November.
- Biondi, G., Mosa, E., & Panzavolta, S. (2009). *Autonomia e innovazione: Scenari possibili tra teoria e pratica, programma education*. FGA Working Paper, 16(2), 1–37.
- Boyatzis RE (1998) *Transforming Qualitative Information: Thematic Analysis and Code Development*. Thousand Oaks, CA: SAGE.
- Braun, A., Maguire, M., & Ball, S. J. (2010). Policy enactments in the UK secondary school: Examining policy, practice and school positioning. *Journal of Education Policy*, 25(4), 547–560.
- 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.
- Braun, A., & Maguire, M. (2020). Doing without believing—enacting policy in the English primary school. *Critical Studies in Education*, 61(4), pp. 433–447.
- Cairney P (2012) *Understanding Public Policy: Theories and Issues*. New York: Palgrave Macmillan.
- Candido, H. H. D. (2019). Datafication in schools: enactments of quality assurance and evaluation policies in Brazil. *International Studies in Sociology of Education*, 0214.
- Capano, G., & Lippi, A. (2017). How policy instruments are chosen: Patterns of decision makers’ choices. *Policy Sciences*, 50(2), 269–293.
- Capano G., Terenzi P (2019). Interest Groups and the «Good School» (Buona Scuola) Reform, on "Rivista Italiana di Politiche Pubbliche, Rivista quadrimestrale" 2/2019, pp. 247-276,
- Capano, G., Zito, A., Toth, F., & Rayner, J. (2022). *Trajectories of governance: How states have shaped critical policy sectors in the neoliberal age*. Palgrave Macmillan

- Chávez, R. C., Etcheberrigaray, G., Vargas, S., Assaél, J., Araya, R., & Redondo-Rojo, J. (2021). Actividades emocionales del trabajo docente: un estudio de shadowing en Chile. *Quaderns de Psicologia*, 23(1), pp. 1–26
- Coburn, C. E. (2001). Collective sensemaking about reading: How teachers mediate reading policy in their professional communities. *Educational Evaluation and Policy Analysis*, 23(2), 145–170. <https://doi.org/10.3102/01623737023002145>
- Colombo, M., & Desideri, A. (2018). Italy. The Italian education system and school autonomy. In S. Martins, L. Capucha, & S. O. J (Eds.), *School autonomy organization and performance in Europe. A comparative analysis for the period from 2000 to 2015* (pp. 101–113). CIES – ISCTE (Centre for Research and Studies in Sociology).
- Colquhoun, H. L., Levac, D., O'Brien, K. K., Straus, S., Tricco, A. C., Perrier, L., Kastner, M., & Moher, D. (2014). Scoping reviews: time for clarity in definition, methods, and reporting. *Journal of clinical epidemiology*, 67(12), pp. 1291-1294.
- Day, C., & Smethem, L. (2009). The effects of reform: Have teachers really lost their sense of professionalism?. *Journal of educational change*, 10(2), pp. 141-157.
- Day, C. (2018). Professional identity matters: Agency, emotions, and resilience. In P. A. Schutz, J. Hong, & D. C. Francis (Eds.), *Research on teacher identity* (pp. 61- 70). Cham: Springer.
- De Sanctis, G. (2010). *TALIS. I docenti italiani tra bisogni di crescita professionale e resistenze*. Fondazione Giovanni Agnelli Programma Education, FGA Working Paper, 24(2), 1–34.
- Diamond, J. B. (2007). Where the rubber meets the road: Rethinking the connection between high-stakes testing policy and classroom instruction. *Sociology of Education*, 80(4), pp. 285–313.
- DPR 80/2013, *DECRETO DEL PRESIDENTE DELLA REPUBBLICA 28 marzo 2013, n. 80 Regolamento sul sistema nazionale di valutazione in materia di istruzione e formazione*, Retrieved from: [https://www.istruzione.it/valutazione/allegati/DPR\\_%2028\\_03\\_13.pdf](https://www.istruzione.it/valutazione/allegati/DPR_%2028_03_13.pdf)
- Dubnick, M. (2005). Accountability and the promise of performance: In search of the mechanisms. *Public Performance & Management Review*, 28(3), pp. 376-417.
- European Commission, Directorate-General for Education Sport and Culture (2018) *Study on supporting school innovation across Europe: final report*. Report, European Commission: Publications Office.
- Falabella, A. (2020). The ethics of competition: Accountability policy enactment in Chilean schools' everyday life. *Journal of Education Policy*, 35(1), 23–45.

- Faggioli, M., & Mori, S. (2018). Valutare l'innovazione scolastica: Vincoli ed opportunità nel sistema nazionale di valutazione. In M. Freddano e S. Pastore (a cura di) Monografia, *Per una valutazione delle scuole, oltre l'adempimento* (pp. 88–99). Milano: Franco Angeli Edizioni
- Fahey, G., & Koster, F. (2019). *Means, ends and meaning in accountability for strategic education governance*. OECD education working papers, (204).
- Farvis, J., & Hay, S. (2020). Undermining teaching: How education consultants view the impact of high-stakes test preparation on teaching. *Policy Futures in Education*, 18(8), 1058–1074.
- Ferrer-Esteban G (2016) *Deconstructing the [Italian] Black-Box. Gli Approcci Didattici Degli Insegnanti Italiani e Le Performance Degli Studenti*. Torino: Fondazione Giovanni Agnelli.
- Fondazione Agnelli (2014) *La Valutazione Della Scuola. A Che Cosa Serve e Perché È Necessaria all'Italia*. Roma-Bari: Laterza.
- Grimaldi, E., & Serpieri, R. (2012). The transformation of the education state in Italy: A critical policy historiography from 1944 to 2011. *Italian Journal of Sociology of Education*, 4(1), 146–180.
- Grimaldi, E., & Barzanò, G. (2014). Making sense of the educational present: Problematising the “merit turn” in the Italian eduscape. *European Educational Research Journal*, 13(1), 26–46.
- Gunnulfson, A. E., & Roe, A. (2018). Investigating teachers’ and school principals’ enactments of national testing policies: A Norwegian study. *Journal of Educational Administration*, 56(3), 332–349.
- Hall D, Grimaldi E, Gunter HM, et al. (2015) Educational reform and modernisation in Europe: The role of national contexts in mediating the new public management. *European Educational Research Journal* 14(6): 487–507.
- Hardy, I. J., Reyes, V., & Hamid, M. O. (2019). Performative practices and ‘authentic accountabilities’: Targeting students, targeting learning? *The International Education Journal: Comparative Perspectives*, 18(1), 20–33.
- Herman, J. (2004). The effects of testing on instruction. In S. Fuhrman & R. Elmore (Eds.), *Redesigning accountability* (pp. 141–146). Teachers College Press.
- Hofman, R. H., de Boom, J., Meeuwisse, M., & Hofman, W. H. A. (2013). Educational Innovation, Quality, and Effects: An Exploration of Innovations and Their Effects in Secondary Education. *Educational Policy*, 27(6), 843–866.
- Holloway, J., & Brass, J. (2017). Making accountable teachers: the terrors and pleasures of accountability. *Journal of Education Policy*, 33(84), pp. 1-22.

- Holloway, J., Sørensen, T. B., & Verger, A. (2017). Global perspectives on high-stakes teacher accountability policies: An introduction. *Education Policy Analysis Archives*, 25(85).
- Hood C and Peters BG (2004) The Middle Aging of new public management: Into the age of Paradox?, *Journal of Public Administration Research and Theory* 14(3): 267–282
- Hursh D (2005) The growth of high-stakes testing in the USA: Accountability, markets and the decline in educational equality. *British Educational Research Journal* 31(5): 605–622.
- Karlsen GE (2000) Decentralized centralism: Framework for a better understanding of governance in the field of education. *Journal of Education Policy* 15(5): 525–538.
- Kim, T. & Yun, J. T. (2019). Logics of accountability: Cross-national patterns in schoollevel controls. *Education Policy Analysis Archives*, 27(119).
- Kickert W (2007) Public management reforms in countries with a Napoleonic state model. In: Pollitt C, Van Thiel S and Homburg V (eds) *New Public Management in Europe: Adaptation and Alternatives*. Basingstoke: Palgrave/Macmillan, pp.26–51
- Knight, R. (2020). The tensions of innovation: Experiences of teachers during a whole school pedagogical shift. *Research Papers in Education*, 35(2), 205–227.
- Licht, A.H, Tasiopoulou, E., Wastiau, P. (2017). (2017). Open book of educational innovation. *European Schoolnet, Brussels*. Retrieved from [www.europeanschoolnet.org-info@eun.org](http://www.europeanschoolnet.org-info@eun.org)
- Stoll, L., & Kools, M. (2017). The school as a learning organisation: a review revisiting and extending a timely concept. *Journal of Professional Capital and Community*, 2, 2–17.
- Lambert, K., & Gray, C. (2020). Performing ‘teacher’: exploring early career teachers’ becomings, work identities and the [mis-]use of the professional standards in competitive educational assemblages. *Pedagogy, Culture and Society*, 28(4), pp. 501–523.
- LeChasseur, K., Mayer, A., Welton, A., & Donaldson, M. (2016). Situating teacher inquiry: a micropolitical perspective. *School Effectiveness and School Improvement*, 27(2), pp. 255–274.
- Lewis, S., & Hardy, I. (2014). Funding, reputation and targets: the discursive logics of high-stakes testing. *Cambridge Journal of Education*, 45(2), pp. 245–264.
- Looney JW (2009) *Assessment and innovation in education*. OECD Education Working Papers, No 24.
- Lubienski C (2009) *Do Quasi-markets Foster Innovation in Education? A comparative perspective*. Report, OECD Education Working Papers, No. 25. Paris: OECD Publishing.

- Malandrino, A. (2021). Le politiche per la formazione iniziale degli insegnanti e l'accesso alla professione in Italia: Caratteristiche e criticità alla luce dell'institutional layering. *Rivista Italiana Di Politiche Pubbliche*, 16(2), 161–190.
- Martini, A., & Papini, M. (2015). *Che cosa ne pensano gli insegnanti delle prove INVALSI*. Working paper 24/2015. INVALSI.
- Mattei P (2012) Market accountability in schools: Policy reforms in England, Germany, France and Italy. *Oxford Review of Education* 38(3): 247–266.
- 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). Springer.
- Ochs, K. & Phillips, D. (2002) *Towards a Structural Typology of Cross-national Attraction in Education* (Lisbon, Educa).
- OECD (2013a) *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*. Report, Paris: OECD Publishing.
- OECD (2013b) *TALIS 2013 Results: An International Perspective on Teaching and Learning*. Report. Paris: OECD Publishing.
- Ozga, J., & Jones, R. (2006). Travelling and embedded policy: the case of knowledge transfer. *Journal of Education Policy*, 21(1), 1–17.
- Paletta, A., Basyte Ferrari, E., & Alimehmeti, G. (2020). How principals use a New accountability system to promote change in Teacher practices: Evidence from Italy. *Educational Administration Quarterly*, 56(1), 123–173.
- Patton, M. Q. (2015). *Qualitative research and evaluation methods: Integrating theory and practice* (4th ed.). Sage.
- Pawson, R., & Tilley, N. (2004). Realist evaluation. In *UK British Cabinet Office* (pp. 505–511). UK: British Cabinet Office.
- Pedro F. (2023), Chapter 10: Where is the school going? International trends in educational innovation, *Sociology, Social Policy and Education*, pp 147–164.
- Perold, M., Oswald, M., & Swart, E. (2012). Care, performance and performativity: Portraits of teachers' lived experiences. *Education as Change*, 16(1), pp. 113– 127
- Peterson, A., Hanna, D., Lafuente, M., & Law, N. (2018). Understanding innovative pedagogies: Key themes to analyse new approaches to teaching and learning. *OECD Education Working Papers* No. 172, (172), 135.



- Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide*. Oxford, UK: Blackwell Publishing.
- Piro F., (2016) Manuale di educazione al pensiero critico. Comprendere e argomentare, *Italian journal of management* Vol. 34, N. 100, 2016
- Pollitt C (2007) Convergence or divergence: What has been happening in Europe? In: Pollit C, Van Thiel S and Homburg V (eds) *New Public Management in Europe*. London: Palgrave Macmillan UK, pp.10–25.
- Pollitt C and Bouckaert G (2011) *Public Management Reform. A Comparative Analysis—New Public Management, Governance, and the Neo-Weberian State*, 3rd edn. Oxford: Oxford University Press.
- Reinholz DL and Andrews TC (2020) Change theory and theory of change: what's the difference anyway? *International Journal of Stem Education* 7(1): 2–12.
- Sahlberg, P. (2006). Education reform for raising economic competitiveness. *Journal of Educational Change*, 7(4), 259–287.
- Sahlberg, P. (2009). The role of education in promoting creativity: potential barriers and enabling factors School education and creativity. *Measuring Creativity*, (337), 337–344. Retrieved from <http://www.greenschool.org/wp-content/uploads/2012/03/Pasi-Sahlberg.pdf>
- Serpieri R (2009) A 'war' of discourses. The formation of educational headship in Italy. *Italian Journal of Sociology of Education* 4(1): 121–142.
- SITES. (2009). *Second information technology in education study. SITES 2006 technical report*. Retrieved from <https://www.iea.nl/sites-2006>
- Smith, W. C. (2014). The Global Transformation Toward Testing for Accountability. *Education Policy Analysis Archives*, 22.
- Sorensen, T. B., & Robertson, S. L. (2017). The OECD program TALIS and framing, measuring and selling quality Teacher™. In M. Akiba & G. K. LeTendre (Eds.), *International handbook of teacher quality and policy* (pp. 117–130). New York: Routledge
- Stapley, E., O'Keeffe, S., & Midgley, N. (2022). Developing typologies in qualitative research: The use of ideal-type analysis. *International Journal of Qualitative Methods*, 21, 1–9.
- Steiner-Khamsi, G. (2014). Cross-national policy borrowing: Understanding reception and translation. *Asia Pacific Journal of Education*, 34(2), 153–167.
- Syariff M Fuad, D. R., Musa, K., & Hashim, Z. (2020). Innovation culture in education: A systematic review of the literature. *Management in Education*, 36(1).

- Teltemann, J., & Jude, N. (2018, July). New accountability schemes? Assessing trends in educational assessment and accountability procedures in OECD countries. *Paper presented at the XIX International Sociological Association (ISA) World Conference*, Toronto, Canada.
- Tromp E. (2017) Global Policies, local meanings: The Re-contextualization of Competence-Based Education in Mexico, in: *Global Education Policy and International Development: New Agendas, Issues and Policies*
- Valli, L., & Buese, D. (2007). The changing roles of teachers in an era of high-stakes accountability. *American Educational Research Journal* (Vol. 44).
- Verger, A., & Parcerisa, L. (2017). *Accountability and education in the post-2015 scenario: International trends, enactment dynamics and socio-educational effects*. Paper commissioned for the 2017/8 Global Education Monitoring Report, Accountability in education: Meeting our commitments. <http://unesdoc.unesco.org/images/0025/002595/259559e.pdf>
- Verger, A., Parcerisa, L., & Fontdevila, C. (2019). 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.
- Vincent-Lancrin, S., Urgel, J., Kar, S., & Jacotin, G. (2019). *Measuring innovation in education 2019: What has changed in the classroom?*. OECD Publishing.
- Yin, R. (2009). *Case study research: Design and methods* (4th ed.). SAGE.

## Compendium of publications

This doctoral thesis is based on a compendium of publications. Below, an overview of the articles that are part of the compendium of publications is provided.

**Publication A:** *Campbuijsen M.; Levatino, A.; Mentini L.; Parcerisa, L. (2022). Teaching as affective labour in a datafied world: a scoping review of the relationship between performance-based accountability and teachers' emotions, Foro de Educación, 20(2), pp. 61-84. <http://dx.doi.org/10.14516/fde.1023>*

**Publication B:** *Mentini, L., & Levatino, A. (2024). A "three-legged model": (De)constructing school autonomy, accountability, and innovation in the Italian National Evaluation System. European Educational Research Journal, 23(3), 321-346. <https://doi.org/10.1177/14749041221148280>*

**Publication C:** *Mentini L. (2024): Negotiating between the accountability and the innovation mandates: evidence from Italian schools, Nordic Journal of Studies in Educational Policy, DOI: 10.1080/20020317.2024.2385120*

## **CHAPTER 2: TEACHING AS AFFECTIVE LABOUR IN A DATAFIED WORLD: A SCOPING REVIEW OF THE RELATIONSHIP BETWEEN PERFORMANCE-BASED ACCOUNTABILITY AND TEACHERS' EMOTIONS**

### **Introduction**

In recent decades, new public management (NPM) has influenced public sector reform worldwide (Common, 1998) with the aim to modernize and «de-bureaucratize» public administration (Bislev et al., 2002; Verger, 2012). Generally speaking, NPM reforms include a wide range of policies and measures, such as new, data-intensive accountability systems focused on performance results, decentralization of decision-making, competition between providers and the empowerment of users through «exit» and «choice» mechanisms (Clarke et al., 2000, p.6). In the education sector, NPM has resulted in an intensification of the «datafication» of schooling, which involves «the collection of data on all levels of educational systems (individual, classroom, school, region, state, international), potentially about all processes of teaching, learning and school management» (Jarke & Breiter, 2019, p.1). Datafication is linked to the dissemination of new policy technologies, metrics, standardized tests and performance based-accountability (hereafter PBA) instruments.

Research has shown that in many contexts, the transformation of the public sector following NPM reforms has significant emotional effects on public sector workers (Dubnick, 2005). Focusing specifically on the education sector, educational research has highlighted that performance monitoring and PBA have affected and reshaped educators' practices, identities and experiences, as well as the possibilities of the teaching profession (e.g., Ball, 2003; Day & Smethem, 2009; Day, 2018; Holloway & Brass, 2017; Lewis & Hardy, 2018). Scholarly attention upon these aspects is not surprising considering the emotional nature of teachers' work. As argued by Hargreaves (1998, p. 838), teaching is an «emotional practice» which «activates, colours and expresses» the feelings of teachers and those with whom they work. Teachers' emotions, in turn, shape their relationships with students, parents and colleagues, as well as with school structures, pedagogy and educational reform. In this regard, we consider that it is essential not only to explore the emotions caused by PBA, but also to understand what role emotions have in the enactment of PBA policies and how teachers use them to cope with PBA pressures.

With this aim of contributing to this understanding, in this paper we attempt to bring together the academic literature produced within different disciplinary fields. Our goal is to provide an overview of knowledge on the relationship between PBA and teachers' emotions, to identify potential limitations of research on the topic and elaborate future lines of inquiry. To do so, we conduct a scoping review of the literature on teacher emotions and PBA, which is based on a final sample of 62 documents published between the years 2000-2021 obtained from the SCOPUS database.

Our review allowed us to identify two main bodies of research. The first deals with an examination of teachers' emotions, without specifically focusing on PBA. In this literature, PBA is a crucial part of a changing teaching environment that accentuates and/or modifies feelings and emotions already inherent to the teaching profession. In the second strand of research, the focal point is the effect of PBA on teacher emotions. More specifically, here we identify research exploring what kind of emotions are provoked by PBA policies in education, what mechanisms could explain different emotional experiences, how teachers deal with them, as well as what factors intensify or weaken the emotional effects of PBA.

The structure of the article is as follows. First, we describe the research methodology. Thereafter, we present the main findings, which are organized into two subsections. In the first subsection, we present the findings concerning the first body of research, while in the second subsection, we provide an overview of the studies that specifically analyze the relationship between PBA policies and teachers' emotions. The last section concludes the paper by discussing the findings, highlighting the main gaps of existing scholarship and giving insights on possible future lines of research.

## **Data and methodology**

Considering the exploratory nature of this research endeavor, methodologically this paper relies on a scoping review strategy. This method enables researchers to map out the main bodies of literature existing on a particular topic by assessing existing studies that address the questions of the review (Petticrew & Roberts, 2006). In so doing, a scoping review can provide an overview of the main concepts underpinning a research topic and identify the sorts of studies carried out, as well as where they are published and what sorts of outcomes they have generated (Colquhoun et al., 2014).

To conduct our review on the relationship between emotions and PBA, we follow a research protocol consisting of a number of steps, as displayed in Figure 1. During the first step, we conducted a literature search using the SCOPUS search engine. The keywords we used were related to two main areas: PBA reforms and teacher affective dimensions (i.e., teacher emotions, teacher lived experiences, teacher subjectivities). The use of this syntax resulted in 127 articles. The whole syntax used can be found in the Annex.

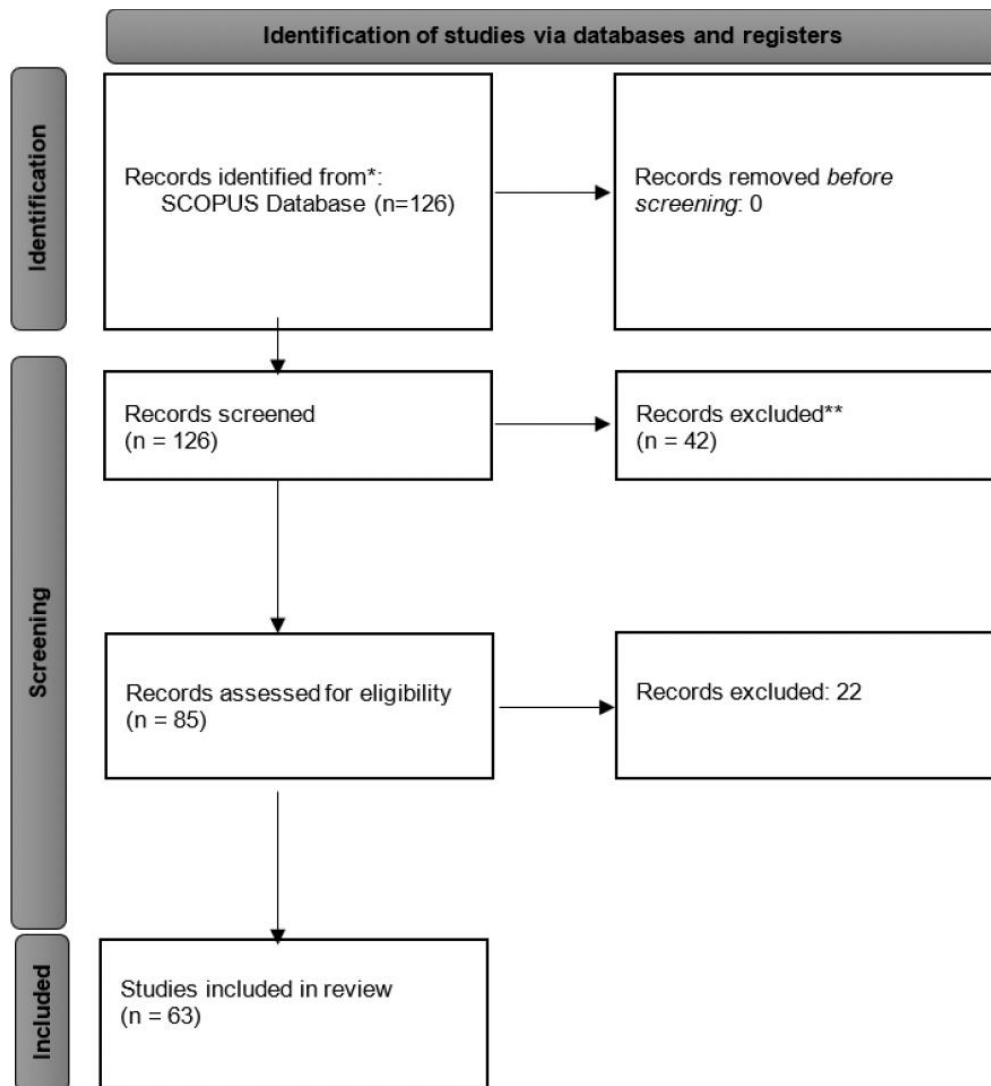
Subsequently, a matrix containing information on each retrieved article (i.e., title, author(s), year of publication, keywords, abstract) was constructed. The matrix allowed us to screen the abstracts in the following manner. Two reviewers separately read each abstract and expressed an opinion on whether the article was to be included, excluded or whether the inclusion was doubtful. In case of disagreement between the two reviewers, or where one expressed doubt, a third (and then, if needed, a fourth) reviewer was involved in this first screening. In this phase, the decision about the inclusion/exclusion was taken with a majority criterion. Criteria for inclusion were: 1) the article must deal with PBA and emotions; 2) the article must focus on the teaching profession and on teachers; 3) the article must be centered on compulsory education<sup>3</sup>. The first screening of the abstracts resulted in 85 articles.

The resulting sample was split into four and each quarter of articles was assigned to one author, who was in charge of downloading and carefully reading the assigned articles. In this phase, one article was excluded because it was unretrievable. Other articles (19 in total) were excluded because, despite the impression given by their abstract, they did not fulfill the inclusion criteria used during the first screening. In addition, we excluded three articles focused only on document analysis and without any empirical or reflective outlook on the relationship between PBA and teachers' emotions. Finally, we also excluded papers not written in English, nor in any language spoken by at least one of the authors. More specifically, this criterion resulted in the inclusion of one article in Spanish and one in Afrikaans, and in the exclusion of one article in Croatian. The second screening resulted in a final sample of 63 articles (see Figure 1).

---

<sup>3</sup> We decided to focus exclusively on compulsory education because PBA has been particularly present at these educational levels, essentially under the form of national standardized assessments and teacher evaluation mechanisms.

Figure 1. Flow chart of the scoping review (Source: adapted from Page et al. (2021))



The 63 included articles were published between 2003 and 2021. The majority used a qualitative methodological approach, four were based on a mixed-methods design and eight were theoretical reflections and/or literature reviews. Geographically, the articles selected encompassed a range of continents and different countries; a third focused on English-speaking countries (US, UK and Australia), while the rest were based in European countries (Spain, Norway, Cyprus, Ireland, Belgium), Asia (India, Singapore, China), Africa (South Africa) and the Middle East (Bahrain, United Arab Emirates).

We categorized and analyzed the selected articles according to their theoretical framework, methodology, research question, geographical focus, discipline and main findings. Finally, to carry out the synthesis, we followed an inductive analytic strategy that was used to sort the articles

following emerging analytic categories. Specifically, we grouped and analysed the articles thematically according to common broad themes (e.g., PBA as context, type of emotion, mechanisms, dealing with emotions, conditions) and subthemes (e.g., tension, reputation, agency, coping strategy, etc.), which reflect the main contribution/s of the article in relation to the link PBA-teacher-emotions.

## Findings

Our review identified two main strands of research. The first deals with the exploration of teachers' emotions in general as well as with the identification of factors that shape and influence them. In this literature, the focus is not specifically on how PBA policies directly impact emotions. Rather, PBA policies are seen as part of a changing educational context that, with its new features, is reinforcing or modifying already existing mechanisms affecting teachers' emotionality. The second body of research specifically deals with the effect of PBA on teachers' emotions, trying to identify what types of emotions are generated by the existence of PBA in education, the mechanisms behind the emergence of these emotions and how teachers deal with them. This second strand of literature also highlights the conditions that can intensify or weaken teachers' experiences of the identified emotions.

### Exploring teacher emotions: emotional rules and emotional labour in changing educational contexts

According to the first body of literature, one of the most important factors shaping teachers' emotions are the emotional rules attached to the teaching profession. These refer to expectations and norms about how teachers have to manage and make use of their emotions during their day-to-day work (Bodenheimer & Shuster, 2020) and about which emotions are permitted and/or considered appropriate. This means that teachers are expected to show—and ideally feel—a range of desirable emotions (such as empathy, calmness, kindness, compassion), thereby conforming to professional norms (Bodenheimer & Shuster, 2020). In contrast, teachers should also control and repress other emotions considered professionally unacceptable, such as anger, anxiety and vulnerability (Dunn et al., 2020). Emotional rules are seldom explicitly stated, but rather operate as tacit prescriptions in the form of ethical codes, pedagogical practices and techniques. Teachers are socialized into these rules and work roles, learning what they should and should not feel, and which feelings they should and should not show (Dunn et al., 2020).



Although emotional rules are often internalized and embodied by teachers (Dunn et al., 2020), their fulfillment requires constant and substantial emotional labour. As Bodenheimer and Shuster (2020) highlight, emotional labour is a concept introduced by Hochschild (1983) and defined by Morris and Feldman (1996, p. 987) as the effort, planning, and control needed to express organizationally desired emotions during interpersonal interactions. The authors also remark how emotional labour, according to Hochschild (1983, p. 33), often involves surface acting, i.e., «portraying emotions to a public without actually feeling them». Emotional labour thus constitutes a rational act of self-regulation of emotions, which requires extreme intellectual work. In this sense, it challenges the common belief according to which emotions are antithetical to rationality (Dunn et al., 2020). The reviewed articles point out how the daily enactment of emotions not aligning with inner feelings can lead to the experience of guilt, shame, frustration, vulnerability, anger (Dunn et al., 2020) and even to burnout (Bodenheimer & Shuster, 2020). These negative feelings, considered outlaw emotions, are often silenced (Dunn et al., 2020) or expressed exclusively in the private sphere (Steinberg, 2008).

There is broad agreement in the scholarly literature that the teaching profession is highly emotional but also highly regulated (Bodenheimer & Shuster, 2020; Chávez et al., 2021). The self-regulation of emotions, similar to what happens in other service-oriented/caring occupations, is considered inherent to the teaching profession. Nonetheless, various studies have pointed out how education policy reforms and changes in the structure of education might alter professional emotional rules and, consequently, expectations of emotional labour performed by teachers (Bodenheimer & Shuster, 2020). In this regard, PBA policies may reinforce emotional rules of impersonal distance, neutrality and disengaged objectivity (Steinberg, 2008). This increased focus on objectivity and neutrality make several scholars claim that the reforms implemented in the framework of NPM, especially PBA policies, neglect the intrinsic emotional nature of the teaching profession (Chávez et al., 2021; Perold et al., 2012).

Beyond this impact of accountability measures on reinforcing undesirable emotional rules, Steinberg (2008) argues that the increasing use of external standardized tests is also seen to accentuate already existing emotions. As her review of research on assessment and emotions shows, research has indeed claimed that teachers experience a range of emotions (e.g., irritation, stress, shame, disgust, guilt, pride, joy, sympathy) during the assessment process (Steinberg, 2008). External assessments are seen to intensify these emotions, although a number of factors intervene

in this process. These include teachers' perceptions about who is responsible for low test results, teachers' beliefs about the validity of the test and the consequent (lack of) necessity to raise scores, teachers' perceptions and concerns regarding the emotional impact of testing and low test results on pupils, as well as whether results are made public (Steinberg, 2008).

## Teacher emotions under PBA regimes

### *The emotional effects of PBA*

Moving on to research that focuses specifically on the effect of PBA on teachers' emotions, this body of literature highlights a range of different emotional effects. On the one hand, research documents that teachers can experience positive emotions such as joy, pride and job satisfaction in the wake of PBA reforms (Holloway & Brass, 2017; Falabella, 2020). On the other hand, a wide range of reviewed studies report that teachers feel unpleasant and negative emotions deriving from PBA reforms, including unhappiness, irritability (van Wyk & Le Grange, 2016), stress, preoccupation, disillusionment (Murphy et al., 2020), frustration (Brown et al., 2018; Al-Whadi, 2020), fear (Aguilar et al., 2019), dissatisfaction, anxiety and pain (Rojas & Leyton, 2014), as well as feeling drained, confused, demoralized and criticized (Steinberg, 2013).

More specifically, different investigations refer to the «emotional costs» of PBA and how these are also associated with adverse effects on teachers' subjectivity and well-being (Braun & Maguire, 2020, p.11; Lambert & Gray, 2021). For example, qualitative research in Chile shows how PBA contributes to creating an environment of teacher discomfort, which is expressed through anger and fatigue as moods (Aguilar, 2019). Adverse effects on teachers' mental health and well-being have been also documented in the case of early-career teachers in Australia (Lambert & Gray, 2021), as well as in the UK context where teachers were found to experience severe stress and depression deriving from accountability surveillance measures (Hebson et al., 2007). Some papers also highlight how PBA can contribute to damaging trust between staff (van Wyk & Le Grange, 2016) or generating feelings of resentment amongst teachers (Niesz, 2010).

In other cases, it appears that as a consequence of PBA reforms, teachers can experience both positive and negative emotions at the same time (Al-Wadi, 2020). Indeed, the way PBA influences teacher emotions is complex and the reviewed literature recognizes both the «pleasures» and the «terrors» related to PBA reforms (cf. Holloway & Brass, 2017, p.380). That is, performativity

cultures associated with PBA exert continual pressure on teachers but, paradoxically, can also form a crucial source of professional job satisfaction (Troman et al., 2007), thus leading to the experience of mixed emotions, including joy as well as sadness, frustration, anger and fear (Zhang & Tsang, 2021). In some cases, an increase in stress and preoccupation with results seems to lead to feelings of disillusionment with the chosen career. The contradictory emotional effects on teachers, documented by various studies, suggest that PBA differently shapes teachers' well-being, self-efficacy, competence and subjectivities (Murphy et al., 2020).

### *Mechanisms behind the emotional effects of PBA*

In the reviewed literature, a number of explanations are put forward for the emotions that teachers experience as a result of PBA reforms. These explanations relate to (a) the embodiment of performance indicators as a valid proxy of professional success; (b) the datafication of education; (c) the material consequences of PBA reforms; (d) the competition between already internalized emotional rules, established beliefs about good teaching and the new prescriptions/expectations promoted by PBA.

To start with the first explanation, several studies highlight how teachers' emotional experiences can be explained as a result of teachers normalizing hyper-performativity and competition, and embodying new understandings of «the good teacher» as promoted by PBA reforms (e.g., Webb, 2005; Falabella, 2020; Lambert & Gray, 2021). When results are increasingly seen as a valid proxy of professional success, high performance can result in positive emotions and feelings of pleasure. On the other hand, failure to live up to the internalized ideal of «the good teacher» can result in a range of negative emotions, including disappointment, self-doubt and shame. For example, by conducting research in the Chilean context, Falabella (2020) shows how PBA affects teachers' self-image and identity as professionals, as well as their personal ego. PBA is thus conceived as a powerful policy technology that can shape who a teacher is and who a teacher aspires to be. Consequently, PBA can trigger feelings of fear in teachers, but at the same time «produce commitment and dedication with seductive offers for self-enhancement» (Falabella, 2020, p. 31). While this explanation highlights the seductive power of performance metrics, it simultaneously emphasizes how PBA can contribute to an environment of professional insecurity and can produce a set of psychic and social costs (Braun & Maguire, 2020; Falabella, 2020).

A second explanation for the emotional effects of PBA on teachers relates to the public exposure of the teaching profession, following the public dissemination of data, and the associated risk of

being judged with a potential effect on social recognition and teachers' professional reputation. As pointed out by Pierlejewski (2020, p. 472), this can lead to the fear «of not producing the right kind of data», thereby discouraging teachers from engaging with creative or innovative pedagogical practices and strongly limiting their agency. As such, the significant amount of data nowadays collected about teachers and their pupils can have the power to modify and regulate the subjectivities of its objects, thereby serving as a regulatory device. As the reviewed studies show, the fact that performance data are often made visible, not only among school actors but also among external audiences such as school administrative bodies, inspection regimes, parents and media, can produce feelings such as anxiety and stress related to the fear of public humiliation and/or loss of social recognition (Kelchtermans et al., 2009; Niesz, 2010; Lewis & Hardy, 2015; Brown et al., 2018).

A third explanation for the emotional effects of PBA is put forward by studies that focus on the material consequences attached to PBA policies. In particular, in contexts where significant material stakes are attached to performance, such as hiring and firing decisions or financial incentives or sanctions, studies document a range of emotional effects on teachers. For example, research in Chile shows how teachers can experience fear of the material consequences attached to performance indicators, which results in feelings of anxiety and stress (Aguilar, 2019).

Finally, a fourth mechanism behind the emotional effects of PBA relates to the emotional tension that emerges from the competition between the new emotional rules, prescriptions and expectations generated by PBA reforms and already internalized emotional rules. PBA policies, and their focus on performing, might indeed obscure and undermine the traditional focus of teaching on caring and nurturing. In this sense, existing research reports how teachers struggle to hold both roles in balance (Perold et al., 2012). Consequently, teachers are found to experience internal conflicts between deeply rooted emotional rules concerning the caring and relational nature of the teaching profession, and external beliefs about good teaching, normally involving a new jargon made up of concepts such as efficacy and effectiveness (Kelchtermans, 2016). As highlighted by Pereira (2018), neoliberal accountability practices cause teachers to feel a constant sense of being pulled in different directions, which results in emotional tension that is manifested in the double stress of having to reach out to students emotionally and having to meet the expectations of the school and school evaluators. According to Kelchtermans (2016, p. 3), teachers generally «feel not only “accountable” for their pupils' measurable outcomes, but also “responsible” for their pupils as human beings». This means that it is arduous, or even impossible,

to maintain emotionally neutral and objective teacher-pupil relationships, as PBA policies require. This misalignment between teachers' models of good teaching and those promoted by PBA policy discourses (e.g., Tsang, 2019; Chandran, 2021) is therefore put forward as a key explanatory factor for teachers' negative emotional experiences following PBA reforms. Moreover, the intensification of tasks and responsibilities in order to meet the accountability expectations can contribute to the need to adapt practices in ways that do not align with teachers' perceptions of good teaching, giving rise to negative emotions such as dissatisfaction, frustration, powerlessness and meaninglessness (Tsang & Kwong, 2017; Inostroza, 2020; Chandran, 2021).

### *Dealing with the emotional effects of PBA: teacher agency and coping strategies*

Beyond identifying the emotional effects of PBA on teachers and highlighting how emotional effects can be explained, various studies have also attempted to gain a deeper understanding of how teachers deal with the emotional effects of PBA. More specifically, a number of reviewed articles highlight how negative emotions emerging from PBA may have teachers engaging less with innovative or alternative pedagogical practices (Steinberg, 2008; Niesz, 2010; Terhart, 2013). For example, Lechasseur (2015) shows how pressure coming from accountability represents one of the major obstacles for teachers' engagement with alternative practices. One explanation for this finding is related to teachers' desire to avoid painful emotions following low results (Steinberg, 2008). Innovative practices, which require time and attention, would expose teachers to a higher risk and feelings of uncertainty, as they might even think they cannot 'waste time' on something that does not directly contribute to improving standardized test results (Falabella, 2020). Another explanation for this finding is connected to the adoption of an «instrumental and technical orientation to change» (Niesz, 2010, p. 389), which contributes to responding effectively to accountability demands (i.e., building up the image of a good school), even when it means to contradict educational and pedagogical values related to innovation (Niesz, 2010). Finally, this might represent a form of «pragmatic compliance» or resignation (Keddie, 2018; cited by Braun & Maguire, 2020, p. 8). In the English school system, for example, Braun and Maguire (2020, p. 8) observed how the participants in their study «were arguably also compliant» but simultaneously felt «a sense of deep unease».

A range of studies also highlight how teachers can develop different coping strategies to deal with the tension of conflicting priorities and demands (Kelchtermans, 2016). Some of these studies show how teachers can mobilize emotions as a way to deal with the pressures and tensions coming from PBA reforms. For example, in the Chinese context, Huang and Vong (2015) show how

positive emotions linked to Confucian philosophy are mobilized by teachers as a resistance strategy against PBA pressures. In the Irish context, Murphy et al. (2020) detect the potential to construct alternative conceptions to those embedded within PBA systems, by being reflexive and aware of the tensions experienced and by taking «an active role in their self-definition» (Murphy et al., 2020, p. 598). In Chile, Rojas and Leyton (2014, p.209) identify forms of resistance towards PBA policies that are enacted through «self-care tactics», which allow teachers to reconstruct their image as professionals. Nevertheless, representations of teachers' exhaustion and submission are also widespread among the participants in this study. In this regard, Waite (2011, p. 66) observes how teachers often «remain caught between perceived risks of resisting a system judged by narrow assessment criteria and an apparent warrant to embrace self-determination». Eventually, this tension might even lead to teachers feeling «emotionally out of control» (Steinberg, 2013, p. 438).

### *Between PBA and teacher emotions: factors intensifying or weakening the emotional effects of PBA*

Beyond offering explanations for the emotional effects experienced by teachers as a result of PBA reforms, the reviewed literature allowed us to identify a number of factors that seem to influence the relationship between PBA and teacher emotions. These factors provide a sense of the conditions under which the emotional effects of PBA on teachers might either be intensified or weakened. The factors we identified belong to macro-level system structures, meso-level school factors and individual-level factors (i.e., teacher biographies and professional characteristics). It is relevant to mention that some of these factors can correlate with each other.

#### *Macro-level factors: systemic environment and cultural norms*

Existing research highlights that systemic and cultural aspects can influence how teachers perceive and enact PBA, and their feelings towards these reforms. Teachers' position in the governance structure seems to play a role in shaping their emotional experiences of PBA (Chandran, 2021; Niesz, 2010). More concretely, teachers' subordination to the local authority (Chandran, 2021), as well as hierarchical relationships inside schools (Niesz, 2010), are found to generate more negative feelings towards PBA reform.

Cultural and societal values have been also found to play an important role in influencing teachers' emotional experiences of accountability reforms (Chandran, 2021; Huang & Vong, 2015; Mooney et al., 2020; Steinberg, 2013; Zhang & Tsang, 2021). A review of Ireland's official policy documents

highlights, for instance, how a traditionally Catholic culture—which emphasizes morality and consensus to norms and standards—shapes teachers’ subjectivities in the direction of technocracy, thereby reducing their creativity, pleasure and autonomous judgement (Mooney et al., 2020). Other examples are located in the Chinese context, where the traditional Confucian philosophy—associated with moral knowledge, harmony and appropriate behaviours—seems to influence the way teachers manage their emotions under PBA by generating spaces of empowerment (Huang & Vong, 2015) or leading to more rational knowledge and comprehensive understandings (Zhang & Tsang, 2021).

*Meso-level factors: the importance of school culture, student composition and internal relationships*

Another group of articles highlight school-level aspects that play a strong role in influencing the relationship between PBA and teachers’ emotional experiences, namely, the school culture, the socioeconomic background of the student population and the quality of relationships among school actors.

First, it seems that different feelings emerge within different school cultures (Ibrahim, 2020; Kostogriz & Doecke, 2013; Kelchtermans et al., 2009). More specifically, a collaborative school culture is associated with teachers experiencing fewer negative emotions in accountability environments. Such a culture is characterized by interdependent teachers «sharing a common vision and responsibility to enhance student learning» (Ibrahim, 2020, p.5) or a shared «working consensus» related to shared implicit norms, values and practices amongst school staff (Kelchtermans et al., 2009, p.225).

Beyond school culture, teachers’ emotional experiences of PBA policies seem to intensify in schools that cater to disadvantaged student populations (e.g., Falabella, 2020; Steinberg, 2013; Steinberg, 2008; van Wyk & Le Grange, 2016). Several articles show that in such schools, teachers may experience an «overwhelming sense of failure» (Steinberg, 2013, p. 438), as well as higher levels of dissatisfaction, anxiety and insecurity in their teaching, by not being able to generate adequate student achievement. The latter might be due to the perceived inappropriateness of curriculum pace and pedagogy (Steinberg, 2013), or derive from comparisons with schools catering to middle-class students (Kelchtermans et al., 2009). Teachers working in lower socioeconomic environments are also found to do more emotional labour in order to maintain a balance between trying to meet the accountability demands and being patient and creative with students’ slow progress (Steinberg, 2013).

Finally, the quality of relationships among school actors seems to influence teachers' perceptions and emotional experiences of PBA policies. These can refer to teacher-principal relationships, relationships amongst teachers and teacher-student relationships. Some studies highlight how the principal can mediate teachers' emotional experiences of PBA policies (Bodenheimer & Shuster, 2020; Kelchtermans et al., 2009). Principals are indeed considered key in reducing negative emotions generated by PBA and in minimizing teachers' burnout (Bodenheimer & Shuster, 2020) by providing understanding, support and clear directions (Kelchtermans et al., 2009). Although relational aspects of teachers' work can be affected under accountability environments (Mausethagen, 2013), Mahony et al. (2014) show how a good working relationship amongst teachers can make teachers experience assessment policies more positively compared to those who carried out the task in isolation (Mahony et al., 2004). The former group of teachers also seemed to experience higher teacher-pedagogical well-being and self-care (Murphy et al., 2020), higher motivation and efficacy (Mausethagen, 2013) and managed uncertainty in a better way (Gu, 2014). The same can be said about teacher-student relationships: although student-teacher relations may be negatively affected by the PBA regime (Mausethagen, 2013), a good relationship with students, high student engagement and teachers' recognition of the importance of «building a rapport with the students» are found to positively influence teachers' emotional experience and well-being by providing a sense of self-care in face of performative demands (Murphy et al., 2020, p.597).

#### *Teacher-level factors: professional and personal characteristics*

Teacher-related factors belong both to their working conditions and to their individual sense-making and personal characteristics. For example, years of experience in education may influence teachers' attitudes towards the policy and, specifically, the degree of acceptance of the reform. Early-career teachers seem to more often accept the employment conditions and the reform pressures as given (Lambert & Gray, 2021), or engage with them pragmatically and efficiently, instead of opposing them (Terhart, 2013). This may be explained by «the need to “fit in” the normalizing cultural apparatus as a driving force for new teachers to secure ongoing employment» (Lambert & Gray, 2021, p. 7). Nonetheless, the acceptance of such conditions reflected a state of anxiety and has a detrimental impact on teachers' mental health (Lambert & Gray, 2021).

The subject taught is also found to play a role in how teachers are emotionally affected by PBA. Interestingly, teachers who do not teach a subject that is tested in the school are also found to experience stress and frustration (Brown et al., 2018; Lambert & Gray, 2021; Shaw, 2016). This



seems to happen because accountability measures produce a number of changes in schools that indirectly affect teachers whose subjects are not evaluated in the test. Examples include changes in workload, staff relations and working conditions, or teachers having to support other subjects' test scores (maths and reading), all of which add to their day-to-day stressful occupation (Shaw, 2016).

Other aspects related to teachers' professional status that influence the relationship between PBA and teacher emotions are whether teachers have any leadership responsibilities (Niesz, 2010), their contract status (Tsang, 2019; Lambert & Gray, 2021) and whether or not they participate in training activities on the national test (Al-Wadi, 2020). More specifically, holding a leadership position has been found to increase teacher isolation and resentment, consequently negatively affecting the change and innovation that teachers may bring to the school (Niesz, 2010). The sense of powerlessness generated by the accountability environment is intensified for teachers without permanent contracts (Tsang, 2019). More specifically, research highlights that the «casualization», «flexibilization» or even «flexploitation» of the teaching workforce (Charteris et al., 2017), i.e., the tendency towards precarious working conditions, employment uncertainty, poor access to professional development opportunities, lower salaries and disconnection from teaching unions, has been found to increase the likelihood of experiencing harmful emotions such as powerlessness, anxiety and marginalization (Charteris et al., 2017).

Finally, teachers' participation in orientation sessions about the design of the test and its items is found to generate positive reactions, resulting in teachers judging the test as clear and useful. This is in contrast to teachers who did not participate in these training sessions and who felt the application of the tests was «stressful», «demanding», «confusing» and «difficult» (Al-Wadi, 2020, p. 211).

Individual sense-making and the personal characteristics of teachers also appear as crucial factors influencing how emotions in relation to PBA are experienced. Following the idea of emotions being rooted in cognition and evaluation (Nias, 1996, cited in Keltcherman et al., 2009; Nausbaum, 2001, cited in Steinberg, 2013), the emotional experience of teachers seems mediated by their personal beliefs, thoughts and values. A number of articles confirmed the existence of such interpretative lenses by which, although working in similar contexts, teachers' affectivities are personal and unique, and shaped by their own beliefs, perceptions and judgements (Perold et al., 2012; Shaw, 2016; Smyth, 2003; Steinberg, 2008). For instance, although located in the same

disadvantaged context, some teachers manage to gain «agency» and self-confidence which make them experience and respond to the performance regime in a less powerless and passive way (Perold et al., 2012). In another way, teachers' emotions are interwoven with their beliefs and educational goals. When educational goals are aligned to accountability demands, teachers experience a lower degree of stress (Shaw, 2016). Similarly, Steinberg (2008) found that teachers' beliefs about the validity of a test influences their emotional experience towards reform. In this line, teachers' pedagogical identity and task perceptions are found to play a role in how they experience and respond to changes in school reform and how they interpret the expectations deriving from such changes (Kelchtermans et al., 2009; Smyth, 2003). Whether teachers believe their role is to benefit students, care for their emotional well-being, or focus on students with special needs differently affects how they respond to calls for change, how they feel about such changes and their adopted practices (Kelchtermans et al., 2009). In other words, «educational discourses and practices are deeply interwoven with teachers' own narrative conceptions of their identities» (Smyth, 2003, p. 272).

Finally, personal characteristics such as age and gender are also found to play a role in how teachers experience PBA policies. Older teachers are found to experience performance demands and related work in such contexts more negatively. This may be explained by a loss of enthusiasm and motivation for education (van Wyk & Le Grange, 2016), or a higher insecurity in dealing with rapid school changes and changes in their teaching purpose in the later stages of their careers (Hebson et al., 2007). In relation to gender, the emotional rules of teaching are considered both gendered and historically contingent (Zembylas, 2002, cited in Dunn et al., 2020). Consequently, women teachers were found to respond to policy in a «gendered» manner, underrating their achievements and feeling «less» than men (Mahony et al., 2004, p.446) or conceptually pictured as «mothers and selfless saviours» encouraged to regulate their minds and emotions and sacrifice themselves to achieve higher educational standards (Dunn et al., 2020, para.8).

## **Discussion and conclusions**

By conducting a scoping review of the scientific literature, we have shed light on the relation between PBA policies and teachers' emotions and have identified a number of trends in current research on this topic. First, it appears that in different contexts worldwide, PBA policies are a crucial part of a changing educational environment. More specifically, various studies highlight how PBA policies can distance, neutrality and disengaged objectivity (Steinberg, 2008). These new

rules can have an important effect on accentuating feelings and emotions that are already inherent to the teaching profession.

Other studies specifically focused on the effects of PBA policies on teachers' emotions. This body of research shows how teachers can experience a range of different emotions following PBA reforms, including feelings of anxiety, frustration, shame and anger. A number of studies point out that the prescriptive cultures of neoliberal accountability and performativity have contributed to a widespread lack of deep trust in teachers' professional standing and judgement, and to increased levels of professional vulnerability and stress (e.g., Kelchtermans et al., 2009). Although in the reviewed literature PBA seems to predominantly generate unpleasant and negative emotions experienced by teachers, our review also identified different investigations that add a note of caution to this assumption, suggesting that the relationship between PBA and emotions can work in multiple ways and directions. Indeed, a few studies also document positive emotions, such as joy and pride. This highlights how teachers can also experience the pleasures of performativity (cf. Holloway & Brass, 2018); as argued by Falabella (2020), this might relate to the seductive power of performativity.

According to the reviewed research, several mechanisms might explain the documented emotional effects. Some of these mechanisms might be the embodiment of performance indicators and performance data as proxies of professional success and/or as powerful determinants of social recognition and professional reputation. Other explanations are connected to the consequences of PBA and to the severe emotional tensions between internalized professional values, rules and preferences and the new demands and prescriptions of PBA reforms.

The existing literature also highlights how the ways in which teachers deal with emotions is complex (Falabella, 2020); emotions can sometimes be used as strategies to deal with the tensions and pressures generated by PBA policies and to cope with the emotional effects of PBA. In this sense, emotions seem to be not only states that are passively felt, but also as states that can be consciously «activated».

Our review also shows how, despite the promise of PBA reforms of fostering educational quality and pedagogical change (cf. Diamond, 2007), the emotional effects of such reforms can, in some cases, lead to a reduction in the use of innovative or alternative pedagogical approaches by teachers. This seems to be because the fear of obtaining bad results in standardized tests or of not producing

the «right data» (Pierlejewski, 2020, p.472) discourages teachers from taking the risk and the time to innovate (Falabella, 2020).

The different cases covered by the reviewed literature demonstrate how similar emotional effects of PBA are experienced by teachers located in contexts that differ significantly from one another. This suggests how some emotional experiences might constitute an integral part of accountability reforms. Yet, the reviewed literature also highlights how specific features of PBA policies, characteristics and configurations of the education system, as well as specific cultural and institutional traditions, might have a crucial influence on the way in which (and intensity with which) emotions coming from PBA are felt. Moreover, different studies show how numerous factors at the meso-level (school) and micro-level (individual) can also intensify or weaken the emotional effects experienced by teachers.

The existing research has therefore contributed in important ways to our understanding of the relationship between teachers' emotions and PBA policies. Nonetheless, our scoping review also identified a number of limitations by which the existing literature is characterized. Future research could address these limitations in order to advance our understanding of teachers' emotional experiences of PBA reforms. For example, the literature commonly lacked detail about research participant recruitment and/or selection. As a result, it was difficult to evaluate the extent to which the reported emotional experiences belonged to a specific group of teachers (e.g., those more positive or more critical towards PBA reforms).

In addition, we noted that the topic of teachers' emotional experiences following PBA reforms has been predominantly addressed by studies using a qualitative research design. Considering the nature of the topic, this is perhaps not surprising. Nonetheless, while qualitative studies can help deepen understanding of the complex relationship between PBA and teachers' emotions, they do not always show how widespread particular emotional experiences are. Studies relying on robust mixed-methods designs would be especially valuable in this regard. In a similar vein, quantitative studies could try to examine how different factors interact in generating particular emotional experiences. Here, it is worth noting that field experiments also form a promising methodological approach to researching the emotional impact of PBA policies and some of their specific features. They may also help to examine the interactions between different factors and conditions in intensifying or weakening this effect.

It became clear that existing research often treats teachers as a homogenous group of professionals. An assumption underpinning some of the reviewed studies seems to be that teachers are socialized in the same way and that emotional rules of the teaching profession are the same for all teachers. As socialization processes differ according to contextual features, individual background and characteristics, future research might want to devote more attention to how teachers' social origins and class backgrounds influence the way they experience emotions under PBA regimes. Emerging research exploring the influence of cultural values on teaching emotional rules and emotional experiences related to PBA (Chandran, 2021; Huang & Vong, 2015; Mooney et al., 2020; Steinberg, 2013; Zhang & Tsang, 2021) constitutes another promising avenue that should be continued by further research in different contexts.

Finally, studies on this topic produced within different scholarly traditions are currently being performed in disciplinary silos with no mutual recognition. By conducting a review of existing research and by providing an overview of what is known about this research topic, we hope to have contributed to the integration of these bodies of literature and allow future studies to build on previous research. We emphasize the need for future research that takes an interdisciplinary approach in order to build on insights derived from different disciplines.

In addition to the promising lines of future inquiry that have already been mentioned, we would add that some of the findings of our scoping review point to additional directions for future research. For example, a key finding of our review is that emotions are not only an effect or consequence of PBA reforms, but can also be a resource when dealing with some of the pressures and tensions that emerge from PBA. Future research should pay more attention to this transformative potential of emotions. In a similar vein, the findings highlight a discrepancy between studies that place emphasis on teachers being disciplined by performativity, leaving little room for teacher agency and studies that do allow for the latter. Here, a question remains as to what extent teachers can play an active role in mastering the emotional effects, e.g., by activating emotional responses. Finally, more critical research is needed on the often-assumed link between PBA and innovative pedagogy and creativity, paying particular attention to how emotions play a crucial role in mediating this relation.

Our review contains some limitations. Our search relied on one database and on English keywords. Although we consider SCOPUS as the most reliable and comprehensive database, future studies

might want to extend the search to other databases and to articles that do not contain an abstract in English.

To conclude, our scoping review has highlighted the need for continuing research on the complex relationship between PBA reforms and teacher emotions, and the benefits of bringing together different areas of scholarship on this topic, not only to provide a more comprehensive understanding, but also to make fruitful interdisciplinary collaborative efforts possible.

## References

- Aguilar, C.C., Trujillo, A., & López, V. (2019). Identidad docente y políticas de accountability: el caso de Chile. *Estudios Pedagógicos*, 45(2), pp. 121–139.
- Al-Wadi, H. M. (2020). Bahrain's secondary EFL teachers' beliefs of English language national examination: 'How it made teaching different?' *International Journal of Instruction*, 13(1), pp. 197–214.
- Ball, S. J. (2003). The teacher's soul and the terrors of performativity. *Journal of education policy*, 18(2), pp. 215–228.
- Bislev, S., Salskov-Iversen, D., & Hansen, H. K. (2002). The global diffusion of managerialism: transnational discourse communities at work. *Global Society*, 16(2), pp. 199–212.
- Bodenheimer, G., & Shuster, S. M. (2020). Emotional labour, teaching and burnout: Investigating complex relationships. *Educational Research*, 62(1), pp. 63–76.
- Braun, A., & Maguire, M. (2020). Doing without believing—enacting policy in the English primary school. *Critical Studies in Education*, 61(4), pp. 433–447.
- Brown G.T.L., Gebril A., Michaelides M.P., Remesal A. (2018) Assessment as an Emotional Practice: Emotional Challenges Faced by L2 Teachers Within Assessment. In Martínez Agudo, J. (ed.), *Emotions in Second Language Teaching. Emotions in Second Language Teaching*. (pp. 205- 222). Cham: Springer.
- Chandran, M. (2021). Teacher accountability and education restructuring: an exploration of teachers' work identities in an urban school for poors in India. *International Studies in Sociology of Education*, Published Online before Print.

- Charteris, J., Jenkins, K., Bannister-Tyrrell, M., & Jones, M. (2017). Structural marginalisation, othering and casual relief teacher subjectivities. *Critical Studies in Education*, 58(1), pp. 104–119.
- Chávez, R. C., Etcheberrigaray, G., Vargas, S., Assaél, J., Araya, R., & Redondo-Rojo, J. (2021). Actividades emocionales del trabajo docente: un estudio de shadowing en Chile. *Quaderns de Psicologia*, 23(1), pp. 1–26.
- Clarke, J., Gewirtz, S., & McLaughlin, E. (Eds.). (2000). *New managerialism, new welfare?* London: SAGE Publications Ltd.
- Colquhoun, H. L., Levac, D., O'Brien, K. K., Straus, S., Tricco, A. C., Perrier, L., Kastner, M., & Moher, D. (2014). Scoping reviews: time for clarity in definition, methods, and reporting. *Journal of clinical epidemiology*, 67(12), pp. 1291-1294.
- Common, R.K. (1998). Convergence and transfer: a review of the globalization of new public management. *The International Journal of Public Sector Management*, 11(6), pp. 440-450.
- Day, C., & Smethem, L. (2009). The effects of reform: Have teachers really lost their sense of professionalism?. *Journal of educational change*, 10(2), pp. 141-157.
- Day, C. (2018). Professional identity matters: Agency, emotions, and resilience. In P. A. Schutz, J. Hong, & D. C. Francis (Eds.), *Research on teacher identity* (pp. 61- 70). Cham: Springer.
- Diamond, J. B. (2007). Where the rubber meets the road: Rethinking the connection between high-stakes testing policy and classroom instruction. *Sociology of Education*, 80(4), pp. 285–313.
- Dubnick, M. (2005). Accountability and the promise of performance: In search of the mechanisms. *Public Performance & Management Review*, 28(3), pp. 376-417.
- Dunleavy, P., & Hood, C. (1994). From old public administration to new public management. *Public money & management*, 14(3), pp. 9-16.
- Dunn, A. H., Moore, A. E., & Neville, M. L. (2020). “There isn’t an easy way for me to talk about this:” a historical and contemporary examination of emotional rules for teachers. *Teachers College Record*, 122(9), pp. 1-36.
- Falabella, A. (2020). The ethics of competition: accountability policy enactment in Chilean schools’ everyday life. *Journal of Education Policy*, 35(1), pp. 23–45.
- Gu, Q. (2014). The role of relational resilience in teachers’ career-long commitment and effectiveness. *Teachers and Teaching: Theory and Practice*, 20(5), pp. 502– 529.

- Hargreaves, A. (1998). The emotional practice of teaching. *Teaching and Teacher Education*, 14(8), pp. 835–854.
- Hebson, G., Earnshaw, J., & Marchington, L. (2007). Too emotional to be capable? the changing nature of emotion work in definitions of “capable teaching.” *Journal of Education Policy*, 22(6), pp. 675–694.
- Hochschild, A.R. (1983). *The Managed Heart: Commercialization of Human Feelings*. Berkeley, CA: University of California Press.
- Holloway, J. (2019). Risky teachers: mitigating risk through high-stakes teacher evaluation in the USA. *Discourse*, 40(3), pp. 399–411.
- Holloway, J., & Brass, J. (2017). Making accountable teachers: the terrors and pleasures of accountability. *Journal of Education Policy*, 33(84), pp. 1-22.
- Huang, H., & Vong, S. K. (2015). The rhetoric of conscience and teachers’ resistance in an era of accountability: a Macao study. *Asian Education and Development Studies*, 4(3), pp. 357–370.
- Ibrahim, A. (2020). What hurts or helps teacher collaboration? Evidence from UAE schools. *Prospects*, 1-18.
- Inostroza, F. (2020). La Identidad de las Educadoras Diferenciales en tiempos de Políticas de Accountability. *Education Policy Analysis Archives*, 28(29), pp. 1–28.
- Jarke, J., & Breiter, A. (2019). Editorial: The datafication of education. *Learning, Media and Technology*, 44(1), pp. 1-6.
- Kelchtermans G., Ballet K., & Piot L. (2009) Surviving Diversity in Times of Performativity: Understanding Teachers’ Emotional Experience of Change. In Schutz P., & Zembylas M. (eds), *Advances in Teacher Emotion Research: The Impact on Teachers’ Lives* (pp. 215-232). Dordrecht: Springer.
- Kelchtermans, G. (2016). The Emotional Dimension in Teachers’ Work Lives: A Narrative-Biographical Perspective. In Zembylas M., & Schutz P. (Eds.), *Methodological Advances in Research on Emotion and Education* (pp. 31–42). Switzerland: Springer International.
- Lambert, K., & Gray, C. (2020). Performing ‘teacher’: exploring early career teachers’ becomings, work identities and the [mis-]use of the professional standards in competitive educational assemblages. *Pedagogy, Culture and Society*, 28(4), pp. 501–523.
- Lambert, K., & Gray, C. (2021). Hyper-performativity and early career teachers: interrogating teacher subjectivities in neoliberal educational assemblages. *Discourse: Studies in the Cultural Politics of Education*, 0(0), pp. 1–15.



- LeChasseur, K., Mayer, A., Welton, A., & Donaldson, M. (2016). Situating teacher inquiry: a micropolitical perspective. *School Effectiveness and School Improvement*, 27(2), pp. 255–274.
- Lewis, S., & Hardy, I. (2014). Funding, reputation and targets: the discursive logics of high-stakes testing. *Cambridge Journal of Education*, 45(2), pp. 245–264.
- Mahony, P., Menter, I., & Hextall, I. (2004). The emotional impact of performance-related pay on teachers in England. *British Educational Research Journal*, 30(3), pp. 435–456.
- Mausethagen, S. (2013). A research review of the impact of accountability policies on teachers' workplace relations. *Educational Research Review*, 9(1), pp. 16–33.
- Mooney Simmie, G., & Moles, J. (2020). Teachers' Changing Subjectivities: Putting the Soul to Work for the Principle of the Market or for Facilitating Risk? *Studies in Philosophy and Education*, 39(4), pp. 383–398.
- Morris, J., & Feldman, D. (1996). The Dimensions, Antecedents, and Consequences of Emotional Labour. *The Academy of Management Review*, 21(4), pp. 986- 1010.
- Murphy, T. R. N., Masterson, M., Mannix-McNamara, P., Tally, P., & McLaughlin, E. (2020). The being of a teacher: teacher pedagogical well-being and teacher self-care. *Teachers and Teaching: Theory and Practice*, 26(7–8), pp. 588–601.
- Nias, J. (1996). Thinking about feeling: the emotions in teaching. *Cambridge Journal of Education*, 26(3), pp. 293–306.
- Niesz, T. (2010). “That school had become all about show”: Image making and the ironies of constructing a good urban school. *Urban Education*, 45(3), pp. 371– 393.
- Nussbaum, M. C. (2001). *Upheavals of thought: The intelligence of emotions*. Cambridge: Cambridge University Press.
- Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372(71), p. 1.
- Pereira, A. J. (2018). Caring to Teach: Exploring the Affective Economies of English Teachers in Singapore. *Chinese Journal of Applied Linguistics*, 41(4), pp. 488– 505.
- Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide*. Oxford, UK: Blackwell Publishing.
- Perold, M., Oswald, M., & Swart, E. (2012). Care, performance and performativity: Portraits of teachers' lived experiences. *Education as Change*, 16(1), pp. 113– 127.

- Pierlejewski, M. (2020). The data-doppelganger and the cyborg-self: theorising the datafication of education. *Pedagogy, Culture and Society*, 28(3), pp. 463–475.
- Rawolle, S. (2013). Emotions in education policy: A social contract analysis of asymmetrical dyads and emotion. In Newberry, M., Gallant, A. & Riley, P. (Ed) *Emotion and school: understanding how the hidden curriculum influences relationships, leadership, teaching, and learning (Advances in Research on Teaching, Vol. 18)* (pp. 49-69), Bingley, England: Emerald Group Publishing.
- Rojas, M., & Leyton, D. (2014). The new teacher subjectivity Construction of teachers' subjectivities at the beginning of the preferential school voucher's implementation in Chile. *Estudios Pedagógicos*, 40, pp. 205–221.
- Schutz, P. A., & Zembylas, M. (Eds) (2009). *Advances in Teacher Emotion Research: The Impact on Teachers' Lives*, Dordrecht: Springer.
- Shaw, R. D. (2016). Music teacher stress in the era of accountability. *Arts Education Policy Review*, 117(2), pp. 104–116.
- Sherfinski, M. (2020). “Confidence” problems and literacy coaching: How a suburban kindergarten divided “good” and “bad” teachers in the accountability shove-down. *Journal of Early Childhood Literacy*, pp. 1-29.
- Simpson, A., Sang, G., Wood, J., Wang, Y., & Ye, B. (2018). A dialogue about teacher agency: Australian and Chinese perspectives. *Teaching and Teacher Education*, 75, pp. 316–326.
- Smyth, J. (2003). A high school teacher's experience of local school management: A case of the “system behaving badly towards teachers.” *Australian Journal of Education*, 47(3), pp. 265–282.
- Steinberg, C. (2008). Assessment as an “emotional practice. *English Teaching: Practice and Critique*, 7(3), pp. 42-64.
- Steinberg, C. (2013). Teachers dealing with failure in a low socio-economic status school in South Africa. In Craig, C.J., Meijer, P.C. & Broeckmans, J. (Ed.) *From Teacher Thinking to Teachers and Teaching: The Evolution of a Research Community (Advances in Research on Teaching)* (pp. 427-442) Bingley, England: Emerald Group Publishing.
- Terhart, E. (2013). Teacher resistance against school reform: Reflecting an inconvenient truth. *School Leadership and Management*, 33(5), pp. 486–500.
- Troman, G., Jeffrey, B., & Raggl, A. (2007). Creativity and performativity policies in primary school cultures. *Journal of Education Policy*, 22(5), pp. 549–572.

- Tsang, K. K. (2019). The Interactional–Institutional Construction of Teachers’ Emotions in Hong Kong: The Inhabited Institutionalism Perspective. *Frontiers in Psychology, 10*, pp. 1–15.
- Tsang, K. K., & Kwong, T. L. (2017). Teachers’ emotions in the context of education reform: labor process theory and social constructionism. *British Journal of Sociology of Education, 38*(6), pp. 841–855.
- Van Wyk, M., & Le Grange, L. (2016). Die geleefde ervarings van primêre skoolonderwysers binne ’n kultuur van performatiwiteit. *Tydskrif Vir Geesteswetenskappe, 56*(4), pp. 1149–1164.
- Verger, A. (2012). Globalización, reformas educativas y la nueva gestión del personal docente, *Revista Docencia, 46*, pp. 4-13.
- Vigoda, E. (2003). New public management. *Encyclopedia of public administration and public policy, 2*, pp. 812-816.
- Waite, S. (2011). Teaching and learning outside the classroom: Personal values, alternative pedagogies and standards. *Education 3-13, 39*(1), pp. 65–82.
- Webb, P. (2005). The anatomy of accountability. *Journal of Education Policy, 20*(2), pp. 189–208.
- Zembylas, M. (2002). “Structures of feeling” in curriculum and teaching: Theorizing the emotional rules. *Educational theory, 52*(2), pp. 187-208.
- Zembylas, M., & Schutz, P. A. (2016). *Methodological advances in research on emotion and education*. Cham: Springer.
- Zhang, Y., & Tsang, K. K. (2021). Performance-Based Accountability and Teacher Emotions: Role of Zhongyong Thinking. *Frontiers in Psychology, 12*, pp. 1-10.

## **Annex**

*Figure 2. Syntax*

(TITLE-ABS-KEY (“accountability” or “performativity” or “standardi\* test\*” OR “teacher evaluation”)) AND (“teacher\* emotion” OR “teacher\* affect” OR “teacher subjectivity” OR “teacher lived experience”)

# **CHAPTER 3: A “THREE-LEGGED MODEL”: (DE)CONSTRUCTING SCHOOL AUTONOMY, ACCOUNTABILITY, AND INNOVATION IN THE ITALIAN NATIONAL EVALUATION SYSTEM**

## **Introduction**

In recent decades, in a New Public Management (NPM) logic, autonomy, and accountability policies have increasingly informed educational reforms, aimed at improving the quality, equity, efficiency and innovation capacity of education systems (cf. Lubienski, 2009; Sahlberg, 2006). Within this reform model, the decentralization of organizational and pedagogical decisions to lower government levels and schools is usually combined with test-based accountability (TBA), that focuses on the acquisition of competences by students in certain areas of knowledge, measured by external standardized tests.

The emergence and widespread adoption of these policies in education has generated an intense debate around the relationship between autonomy, accountability and innovation<sup>4</sup>, which remains controversial. Behind school autonomy with accountability (SAWA) policies in education is the assumption that “balanced school autonomy, with built-in accountability mechanisms, improves schools’ capacity for innovation” (European Commission, Directorate-General for Education Sport and Culture, 2018: 15). In particular, giving more autonomy to schools is seen as a way of fostering their capacity to adapt to their learning contexts and to students’ individual needs, as well as to generate more freedom to improve and innovate (Looney, 2009; Lubienski, 2009).

---

<sup>4</sup> In this paper, when talking about “innovation” we generally refer to all three levels of innovation, following the definition given by the European Commission, Directorate-General for Education Sport and Culture (2018): (1) pedagogical innovation, that is, teaching practices which are usually new in a given context; (2) organizational innovation, that is, changes in the structure and organizational behavior of schools); (3) system level innovation, that is, the capacity of the system to consolidate change involving several stages and stakeholders. In this paper, when we wish to refer to a specific level of innovation, we will always specify this.

Accountability is viewed as an instrument aimed at balancing the greater responsibilities given to schools, as well as ensuring that students meet centrally defined standards (OECD, 2013a). Governments are willing to give more autonomy to schools to the extent to which schools are willing to be monitored through external assessments and accountability measures (Fahey and Koester, 2019; Verger and Parcerisa, 2017). These parallel pushes toward the decentralization of content and power at local level, and the legitimation of central control and standardization (Karlsen, 2000) have been considered a “paradox” by certain authors (e.g., Falabella, 2014) and tensions between autonomy, accountability and innovation have been highlighted (Fahey and Koester, 2019; Knight, 2020; Looney, 2009). In particular, it has been underlined how the pressure to obtain good results in external, standardized assessments might result in teachers taking fewer risks and having less time to engage in innovative or creative practices (Falabella, 2020; Knight, 2020). Further tension may also be present as regards the way in which standards-driven accountability may counteract the power of schools to organize themselves autonomously and to shape learning (Knight, 2020).

Accountability, autonomy, and innovation are central and are considered equally important constituent aspects in the current Italian National Evaluation System (hereafter SNV). Contemporary reforms in the Italian context have indeed focused on extending school autonomy (Law 57/1997; Law 207/2015), while also adopting a national evaluation system that conjugates both external accountability and internal self-evaluation mechanisms, with an explicit focus on school improvement and innovation (Presidential Decree 80/2013).

Italy represents an interesting case in which to analyze to what extent accountability policies are supposed to foster innovation or they might hinder it. This is because of its centralized system legacy (Mattei, 2012), that coexists with more recent pushes toward decentralization and devolution (Colombo and Desideri, 2018). Furthermore, interestingly, Italy is one of the countries with the most changes taking place in classrooms (OECD, 2013b), notwithstanding its educational model based on the predominance of transmissive and deductive, teacher-centered practices (Bifulco et al., 2010; Ferrer-Esteban, 2016; OECD, 2013b). Existing research into the Italian case has mainly focused on the trajectories, shaping forces and adoption of NPM reforms in education and has covered a timespan up until 2015 (Grimaldi and Serpieri, 2013; Hall et al., 2015; Serpieri, 2009). Studies have also focused on the adoption of headteachers’ and teachers’ evaluation programs, piloted between 2010 and 2012 (Barzanò and Grimaldi, 2013; Grimaldi and Barzanò, 2014; Grimaldi and Serpieri, 2014; Serpieri et al., 2015), the trajectory and governance effects

produced by national standardized tests (Landri, 2014) and the assumptions behind the 2015 reform on school governance (Barone and Argentin, 2016). Despite the centrality of autonomy, accountability and innovation in the Italian SNV, to our knowledge, no study to date has specifically delved into the relationship between these three elements, focusing on the current school autonomy with accountability reform configuration.

This paper aims to (1) identify the theory of change (or program ontology) behind the Italian SNV in its current configuration, with a special focus on how autonomy, accountability and innovation have been conceptualized and linked together within a broader reform package; (2) explore whether tensions exist that might hamper the achievement of the “substantive promises” (Malen et al., 2002: 114) of these policies. To do so, we build upon the analysis of key documents and interviews with key informants involved in the design and implementation of the SNV. The analysis relies on the components of a “theory of change” (i.e. assumptions, intervention(s), rationales/outcomes, context and measurement of outcomes) as conceptualized by Reinholz and Andrews (2020). In doing so, the paper contributes to analyzing the coherence of the theory of change behind the SNV, as well as exploring the main obstacles which, according to our informants, might hinder the achievement of the programmatic goals and initial intentions. This aspect is particularly relevant, considering that a review of the evaluation tools adopted is lacking (Fondazione Agnelli, 2014). Moreover, the Italian SNV is a relatively young school evaluation system, therefore, no policy feedback or impact evaluation has yet been conducted. Indeed, the first cycle of social accountability was conducted between 2016 and 2019, and the first social reporting only ended in 2019.

The paper is structured as follows. We first provide information relating to the reform context in which the policies have taken place. We then present the theoretical underpinnings on which our analysis is based. After outlining the data and methodology used, the findings are presented and discussed in different subsections following the components of the reform ontology. The last section concludes the paper and points to future directions of research.

## **A genealogy of school autonomy and accountability reforms in the Italian context**

For many years, the Italian education system has been characterized by its centralized and bureaucratic nature (Grimaldi and Serpieri, 2012; Mattei, 2012). Italy has been considered as one of the last European countries to introduce NPM reforms (Hood and Peters, 2004) and a latecomer in implementing educational reforms in terms of evaluation (Kickert, 2007). The idea of reforming the public education sector through accountability and autonomy reforms has represented a common thread over the last 30 years (Peruzzo et al., 2022) and has remained consistent across left- and right-center governments (Mattei, 2012). Nonetheless, policies aimed at evaluating the educational system in Italy have become a systemic routine only very recently (Grimaldi and Serpieri, 2014), due to a “policy impasse” generated by opposition and contestations from unions and collegial bodies (Fondazione Agnelli, 2014; Grimaldi and Serpieri, 2013).

In the late 1990s, the discourse on autonomy in Italy emerged under different center-left governments, influenced by a “third-way” discourse, in the context of financial and monetary crises and austerity measures (Peruzzo et al., 2022), pressures at international level and the crisis of the welfarist model (Grimaldi and Serpieri, 2014). The Italian school model thus started to undergo “complex and contested processes of restructuring and reculturing” (Grimaldi and Barzanò, 2014: 26) based on (1) the introduction of school autonomy and decentralization, (2) the formation of a new headteacher role, and (3) the introduction of school, staff and student evaluation. In 1997, a school-based management reform (Law 59/1997) was implemented, within the framework of “soft decentralization,” aimed at increasing the efficiency of the system by granting schools a greater degree of autonomy in organizational, pedagogical and administrative matters. Decree 275/99 was framed within a wider transformation of the public administration (Grimaldi and Serpieri, 2013) and awarded more autonomy and decision-making power to regions and schools, based on a NPM logic (Serpieri, 2009). According to the reform, schools could develop networks with other schools and/or public or private actors, and new responsibilities were granted to headteachers (Grimaldi and Serpieri, 2014).

In the early 2000s, under a center-right government (2001–2006) led by Berlusconi, the National Institute for the Evaluation of the Education and Training System (INVALSI), an in-house agency of the Ministry of Education, was reorganized (Legislative Decree 286/2004), attributing to the Institute the function of systematically assessing student knowledge and competences through standardized national assessments (Law 107/2007). During the course of the next center-right



government, led by Berlusconi (2008-2011) and characterized by themes such as the inefficacy of public education and its costs (Barzanò and Grimaldi, 2012) and the need to contain expenditure due to the economic crisis (Peruzzo et al., 2022), a national evaluation system in Italy began to be piloted (Fondazione Agnelli, 2014). Several schemes were tested, such as, for example, a headteacher voluntary evaluation scheme, a teacher-reward scheme based on reputational mechanisms (Valorizza) and a school reward scheme based on the measurement of the school's added value (VSQ), which involved financial prizes and salary rewards for the 30% top performing schools.

Notwithstanding the lack of success of all the aforementioned pilot programs (Grimaldi and Serpieri, 2014) and the strong resistance they met from teachers' unions and leftist parties, due to their merit-based and financial dimension (Fondazione Agnelli, 2014; Grimaldi and Serpieri, 2014), in 2011, INVALSI tests became mandatory for all schools on a census basis<sup>5</sup> but with no consequences attached. In this sense, TBA in Italy has been considered as "mild" (Pensiero et al., 2019: 84), meaning that, should schools fail to reach the established goals, no consequences are foreseen in relation to school resources or school actors' salaries. In 2012, under a technical government led by the economist Mario Monti, who was asked to ensure fiscal stability and promote human capital in light of an economic rationality (Landri, 2014; Peruzzo et al., 2022), a new 3-year pilot experiment (Vales), explicitly connecting evaluation and improvement, was implemented to evaluate headteachers and school effectiveness (Serpieri et al., 2015). At variance with prior projects, this experiment was characterized by the absence of any prize or merit-based or ranking consequences and paved the way for the current SNV (Sistema Nazionale di Valutazione) which was established in 2013/2014 (DPR 80/2013) and implemented in 2015/2016 under a center-left government. The system is based on a combination of external assessments (through INVALSI standardized tests and sample-based school inspections) and internal school evaluation (by means of a school self-evaluation report). The system also foresees the involvement of another in-house agency, the National Institute for Documentation, Innovation and Educational Research (INDIRE), as well as a range of other private external actors and consultants to support school improvement and innovation processes (Serpieri et al., 2015).

---

<sup>5</sup> From 2011 to 2018, INVALSI standardized tests also formed part of the final exam in lower-secondary education, influencing both the access grade to the exam as well as the final grade.

In 2015, under a center-left government (governo Renzi), school autonomy was further reinforced in terms of human and financial resources, strengthening the responsibilities of the manager-headteacher and highlighting the central role of flexibility in teaching autonomy and innovative teaching methodologies (Law 107/2015). A ministerial web portal (Portale Unico Dati della Scuola) containing figures measured by the SNV, such as schools' self-assessment in INVALSI test results, was also created. These reforms are considered to have introduced competitive dynamics between schools and a "meritocratic logic," which is viewed in sharp contrast to the principles of the center-left (Barone and Argentin, 2016: 138). This has fueled hostilities between the government and labor unions (Peruzzo et al., 2022) as the major strikes, boycotts and sabotage of the INVALSI tests, organized by teacher unions during 2014/2015, demonstrate (Poliandri, 2018).

### **Theoretical framework**

This study adopts an analytical perspective that combines politico-administrative and ideational factors in the study of institutional dynamism and public policy change (Cairney, 2012). This perspective, which is informed by new institutionalism currents, is also inspired by policy sociology approaches that have highlighted the intricacy and complexity of policymaking.

#### **Politico-administrative regimes in the recontextualization of education reforms**

Broadly speaking, institutionalist theory is concerned with attainment and change in the status quo, and how and why specific configurations emerge and become appropriate over time. The idea is that organizations do not exist in a vacuum but interact with the socioeconomic and political context (DiMaggio and Powell, 1991; Meyer, 2008). Accordingly, the way in which the broader cultural, socio-economic, political environment exerts an influence on organizations must be considered (Christensen and Molin, 1995). Sociological institutionalism, one of the most well-established approaches within new institutionalism, considers the state as being composed of multiple and broadranging institutions and agencies (Meyer and Rowan, 1977), referring to institutions as both formal structures of government and political systems, as well as informal rules/norms which guide behavior (Cairney, 2012). The policy process is structured by political institutions, state structures, state-interest groups and policy networks. In this study, we consider

that a multiplicity of state institutions, that is, evaluation, inspection or improvement agencies, research agencies, national leadership or teacher institutions, ministries of education and their staff, but also non-state actors such as teachers' unions shape the policy process in multiple ways. At the same time, we also acknowledge that wider social norms and priorities, promoted by international organizations, are key in understanding education policy-making at different levels (Grek and Ozga, 2010; Martens et al., 2010).

Another highly relevant perspective in relation to new institutionalism is historical institutionalism. From this perspective, institutional change is conceived as "path dependent" (Thelen and Steinmo, 1992: 2), meaning that "the range of options available to policymakers at any given time is a function of institutional capabilities that were put in place at some earlier period" (Krasner, 1988: 67) and that once one of these options is chosen over another, this constraints future possibilities (Krasner, 1988). The idea of "path dependency" rejects the view that the effects of the same forces will generate the same results everywhere, as they will be mediated by the contextual features of a specific situation inherited from the past (Hall and Taylor, 1996). Institutional legacies and politico-administrative regimes are thus considered as mediating the adoption of education reforms. Following this approach, comparative studies on the adoption of public sector reforms (e.g. Pollitt, 2007; Pollitt and Bouckaert, 2011; Verger et al., 2019) have identified three main categories of countries with different administrative regimes. These categories clearly differ in the ways global reform models have been adopted and justified. The first group includes countries with a more liberal organization of the state. In these countries, global education reforms have been adopted within a market-oriented rationale, involving the active participation of the private sectors and elements of competition between providers (Pollitt, 2007). At times, in these countries, the discourses have been instrumentally combined with equity arguments, focused on the importance of reducing achievement gaps (Hursh, 2005; Verger et al., 2019). Another group comprehends countries with a neo-Weberian state tradition, characterized by high levels of decentralization and a strong welfare state (Pollitt, 2007). In these countries, global education reforms have transformed teachers' work but do not seem to have challenged the very idea of public service professionalism. There, TBA has been adopted with the explicit goal of assuring quality (Verger et al., 2019) and as a way of promoting transparency to facilitate citizens' engagement and deliberation (Camphuijsen and Levatino, 2022). A final group consists of the Napoleonic states, which are characterized by centralized and hierarchical administrations, civil servants enjoying high levels of professional autonomy and powerful unions (Hall et al., 2015). In these countries, global education reforms have been adopted with the declared aim of

modernizing public service and administration and making educational systems more flexible (Verger et al., 2019). However, the implementation has often been fragmented and irregular, meeting obstacles and resistance (Kickert, 2007).

Italy, similar to other Southern European countries, has a Napoleonic administrative tradition (Hall et al., 2015; Verger et al., 2019). The country is characterized by a hierarchically structured public administration, a statist legacy and compulsory education, mainly provided by the state but with some options for private (especially religious) schooling. The school system is made of three key stages and is characterized by training tracks and a highly selective system (Bifulco, 2010; Grimaldi and Serpieri, 2012). The introduction of NPM reforms in this country, in particular, has been characterized by a “war of discourses” between performance-based managerial accountability, neoliberalism and the strength of welfarist legacies (Serpieri, 2009: 123).

### The relevance of “ideas” and the discursive and intricate nature of policymaking

The relevance of “ideas” in explaining change emerges more clearly in so-called discursive/constructivist institutionalism (Hay, 2006; Schmidt, 2008). Ideas are indeed central in terms of defining the issues and problems that will define the policy agenda and are influential when they interact with the political actors, who decide to frame them and use them to convince others (Cairney, 2012). However, it is important to stress that ideas do not operate in a vacuum and are context dependent. Political or economic contexts, together with institutional forces, create the conditions for the behavior of actors, as well as for the development, diffusion and translation of their ideas (Stone, 2012).

According to Cairney (2012), policymakers operate in a context of information complexity, in which the analysis of the main problems they face is never comprehensive and where time is limited. Moreover, they have to deal with competing demands and contradictory preferences that are difficult to articulate and order. These challenges are amplified by the differences in knowledge of the actors involved in the policy-making process, as well as by the different meanings they attach to the “language of policy” (Andreas et al., 2022: 3). Policies, from this perspective, can be considered as discursive strategies, drawn from particular and historically

contingent structures of knowledge, and produced by the actor's language, values, beliefs and practices (Ball, 1994). It follows that, far from being the results of a linear procedure, involving the identification of a problem and the search for the most adequate solution, policies are the result of an ambiguous, complex process (Cairney, 2012). In light of such theorizations, it becomes essential to take into account not only the context of text production, that is, the texts representing the policies (Ball, 1993), but also the multiple and sometimes even conflicting meanings provided by individual actors who have been involved in the entire process.

The non-linearity of the policy-making process has also been highlighted by Kingdon's multiple streams framework (Kingdon, 1984), according to which the policy process can be conceptually broken down into different, independent streams: problems, policies and politics. Problems are core components in any policy process, however, not all problems receive attention. In fact, as the political environment is marked by ambiguity and complexity, the attention received by certain problems in relation to others is essentially based on the actors' ability to frame the issue through a persuasive story, often by assigning blame to certain social groups (Zahariadis, 2016: 90). Thus, problems need to be socially constructed to penetrate policy agendas. Policies are also ideas, but in the form of solutions proposed by participants, as strategies to address a problem. Nonetheless, these ideas/solutions are often used to address different aims from those explicitly stated and can even precede the emergence of the problems (Kingdon, 1984). For this reason, it has been argued that contrary to the expected policy sequence, on many occasions, a "solution can be in search of a problem" (Zahariadis, 2003: 59). Politics has to do with how receptive the public is to certain policy ideas/solutions at particular times. We, therefore, take into account that "solutions" can occasionally be considered as the main drivers of policy change, independently of the problems. Furthermore, as the persuasiveness of an idea/solution can be more important than the solution/idea itself, we consider it important to evaluate how the adoption of a solution/idea is justified and communicated in the policy-making process.

## **Methodology**

The analysis is based on a qualitative, interpretive approach combined with a thematic analysis, which is useful when identifying the relationships between themes (Boyatzis, 1998). It draws on two sources of information. Firstly, we analyzed eight key policy and technical documents, which were purposely selected based on their relevance in relation to our research questions. More specifically, we included: the 2004 legislative decree (Legislative Decree 286/2004), which

establishes the reorganization of the INVALSI institute and its evaluation role; the 2007 White Paper (2007 WP), which “contains the knowledge base of standardization processes and the strategic vision of how evaluation is assessed through standards, data and performance” (Landri, 2014: 30); two ministerial directives (Ministerial Directive 88/2011; Ministerial Directive 11/2014), which define the strategic priorities of the SNV and the objectives of the INVALSI tests, respectively; one presidential decree (Presidential Decree 80/2013), which forms the normative basis of the SNV; the 2015 law (Law 107/2015) outlining school autonomy reform and two key technical documents issued by INDIRE (Technical Document, INDIRE) on school innovation and by the INVALSI (Technical Document, INVALSI) relating to INVALSI standardized tests. Complete references of the documents analyzed can be found in Table 2.

Table 2. Overview of documents.

Year	Type of document	Name of document	Issued by	Cited within text as:
2004	Legislative Decree	DECRETO LEGISLATIVO 19 novembre 2004, n.286 Istituzione del Servizio nazionale di valutazione del sistema educativo di istruzione e di formazione, nonché riordino dell'omonimo istituto, a norma degli articoli 1 e 3 della legge 28 marzo 2003, n. 53	Presidency of the Republic	Legislative Decree 286/2004
2007	White Paper	Quaderno Bianco della Scuola	Ministry of Economy and Finances, and Ministry of Public Education	2007 WP
2011	Ministerial Directive	Direttiva del 3/10/2011 n. 88, Obiettivi delle rilevazioni nazionali INVALSI sugli apprendimenti degli studenti - a.s. 2011/2012	Ministry of Public Education	Ministerial Directive 88/2011
2013	Presidential Decree	DECRETO DEL PRESIDENTE DELLA REPUBBLICA 28 marzo 2013, n. 80 Regolamento sul sistema nazionale di valutazione in materia di istruzione e formazione	Presidency of the Republic	Presidential Decree 80/2013
2014	Ministerial Directive	Direttiva del 18 Settembre 2014, Priorità strategiche del Sistema Nazionale di Valutazione per gli anni scolastici 2014/15, 2015/16 e 2016/17	Ministry of Public Education	Ministerial Directive 11/2014
2015	Law	Legge 13 luglio 2015, n. 107. Riforma della scuola "La Buona scuola" - « Riforma del sistema nazionale di istruzione e formazione e delega per il riordino delle disposizioni legislative vigenti.»	Parliament	Law 107/2015
2018	Technical Document	The INVALSI tests according to INVALSI	INVALSI	Technical Document, INVALSI
2019	Technical Document	Avanguardie Educative, l'innovazione possibile	INDIRE	Technical Document, INDIRE

Secondly, we drew on 12 semi-structured interviews with key educational stakeholders, conducted between June 2021 and January 2022, mostly using an online format. To select interviewees, we relied on both purposive (9) and snowball sampling (3). Table 3 provides an overview of the interview participants. When selecting interviewees, a multiple and broad range of institutions and stakeholders were considered. More specifically, we interviewed members of the Ministry of

Education, who were directly involved in the ideation and policy design of the SNV, representatives of the two in-house agencies involved in the design and implementation of education reforms in Italy (i.e., INVALSI and INDIRE), as well as academic scholars with expertise in educational accountability and innovation. Due to their role in influencing NPM discourse regarding education in Italy (Grimaldi and Serpieri, 2013) and their cooperation with INVALSI, INDIRE and the Ministry (Landri, 2014), we also interviewed exponents of private foundations (e.g., Compagnia di San Paolo and Agnelli Foundation<sup>6</sup>) carrying out projects and advocacy in the educational field in Italy. Finally, because of the historically powerful role of teacher and labor unions in educational policymaking (Barzanò and Grimaldi, 2013; Kickert, 2007), we also included representatives of the national school leaders' association (ANP) and trade unions (CGIL and Cobas).

*Table 3. Overview of key actors interviewed.*

Actors	No. of interviewees
Members of the Ministry of Education ( <i>Direzione Generale degli Ordinamenti e del Sistema Nazionale di Valutazione</i> )	2
Representatives of INDIRE Institute	2
Representatives of INVALSI Institute	1
Academic scholars	2
Private foundation exponents	2
School leaders' association	1
Union leaders	2

The interviews lasted approx. 50 minutes and were conducted in Italian, audio-recorded and transcribed verbatim. Key intercepts were translated into English and used as supporting evidence in the finding sections.

Data analysis was conducted using ATLAS.ti software, based on a combination of inductive and deductive approaches. The first group of codes was developed on the basis of the research questions and theoretical framework. This was further enriched and complemented by a set of

---

<sup>6</sup> According to Sorensen et al. (2021: 113), private foundations, such as the Compagnia di San Paolo School Foundation, the Tre Elle Foundation and the Agnelli Foundation, have been “among the most active players in promoting policies and measures that favoured the reculturing of the Italian education system according to a ‘private sector’ ethos.”



new analytical codes that inductively emerged from the data. For the coding and codes' categorization phases, we used the concept of the "theory of change" as an analytical tool, which can be understood as a "particular approach for making underlying assumptions in a change project explicit, and using the desired outcomes of the project as a mechanism to guide project planning, implementation, and evaluation" (Reinholz and Andrews, 2020: 2). In this sense, we understand the theory of change as the reform program ontology. Its constituent elements are: (a) assumptions, that is, implicit knowledge of how change works; (b) interventions, that is, actions required to achieve the desired outcomes; (c) outcomes/rationales, namely, what is to be achieved and why; (d) context of intervention, that is, conditions under which change mechanisms are activated and work; (e) measurement of outcomes, that is, an evaluation of whether and to what extent desired outcomes are being achieved (Reinholz and Andrews, 2020). These elements have been used to structure the findings subsections, as can be appreciated here below.

## Findings

The complex relationship between autonomy, accountability, and innovation:  
Underlying assumptions and intrinsic contradictions

A major feature of any theory of change is the articulation of the underlying assumptions of how change occurs. The interviews provide rich information on the assumptions regarding the relationship between autonomy, accountability and innovation and how they are combined to generate the intended change. Interestingly, however, many aspects also emerge that reveal contradictions with regard to certain assumptions underlying the SNV apparatus.

*Assumption 1: If schools have autonomy in curricular and organizational matters, they will innovate and introduce more change*

The first assumption is that school innovation relies on the presence of a substantial degree of autonomy, where schools have "margins of freedom in defining their own timetables and programmes" (Representative of INDIRE1) and where teachers have the capacity to define their

own curricular and pedagogical offer (2007 WP). The belief is that “autonomy gives schools and teachers the possibility of responding to the needs of the local context and to those of their students” (School leaders’ association), as well as being flexible enough to implement structural changes and methodological innovations (Law 107/2015).

Although innovation clearly constitutes a central element of the SNV, it is interesting to observe that this concept is not defined in the same way by the actors interviewed, who more often refer to what innovation is not, rather than what it actually is. There is a relative consensus that “innovation is not a frontal teaching model,” “it is not the mere transmission of knowledge,” “it is not the centrality of the classroom as a space for learning,” “it is not rigid timetables and programmes,” “it is not content and programmes” (Technical Document INDIRE; Representative of INDIRE1). However, when trying to define what innovation actually is, our interviewees refer to very different pedagogical and organizational features, such as, for instance, a “student’s critical capacity and protagonism” (Representative of INVALSI; Technical Document INDIRE), “wellbeing and motivation” (Representative of INDIRE2; Academic scholar2), “the acquisition of transversal competences required by the job market” (Representative of INDIRE2), “teachers’ critical use of technologies” (School leaders’ association; Representative of INDIRE1), a “school’s capacity to change” (Member of Ministry2), “collaboration amongst staff” (Academic scholar2) through “a distributed leadership model” (Representative of INDIRE2). More than a precise objective to be reached, innovation thus seems to act as a buzzword of sorts, characterized by a lack of conceptual clarity and consensus. Rather, the innovation concept invokes a mix of different methods, objectives and conditions, which vary according to the interlocutor.

*Assumption 2: If schools are evaluated externally through standardized tests, they will be more effective in using school autonomy*

The benefits of school autonomy and its strict relationship with TBA is something which emerges clearly from the analyzed data, according to which:

“Autonomy increases the efficacy of the education system, only in the presence of robust national systems of standardized assessment” (2007 WP).

Schools are in fact expected to plan their actions according to the external results obtained and to do so within the limits of school autonomy.

What is also clear from the interviews is that the need to introduce instruments for an external evaluation of school actions derives from the strengthening of school autonomy and state deregulation, to avoid the risk of schools being self-referential:

“It is necessary to have national (evaluation) instruments which are part of any autonomous country (. . .) The less centralism, the more schools’ activity has to be evaluated through standardized means” (Member of Ministry1).

The interplay between centralism and decentralism in the characteristics of Italian governance is also reflected in the whole Italian accountability model, which is described as a self-evaluation model “guided” and “controlled” by the Ministry through centrally defined standards, items and tools (Member of Ministry1).

The introduction of external TBA is thus justified by the promotion of school autonomy:

“Standardized tests are implemented in Italy with the introduction of school autonomy (. . .) The more schools are autonomous, the more they need to be accountable for what they do” (Representative of INVALSI).

Although school autonomy is considered central in sustaining innovative processes and in motivating the implementation of a national system of accountability, the limited effects of autonomy and devolution reforms clearly emerge. Several interviewees describe the real autonomy of Italian schools as “extremely limited” (Academic scholar1), “incomplete” (Member of Ministry1) or “fake” (Private foundation1) since “[schools] can’t actually change anything, if not the least important things” (Academic scholar1). Indeed, the interviewees report that schools “do not have the power to hire teachers,” they “are not responsible for managing their buildings and spaces” and “their autonomy is also limited in terms of governance and administrative relationships” (Member of Ministry1; Representative of INDIRE1). These observations confirm how in Italy, notwithstanding the high degree of pedagogical autonomy and the limited autonomy in financial terms, internal governance has remained untouched (Serpieri, 2009). Indeed, curriculum and funding have remained centrally defined and

headteachers have no power to recruit, determine wage levels or decide on infrastructure or renovation works (Colombo and Desideri, 2018). They also reflect the peculiar and hybrid form of “centralized decentralism” (Karlsen, 2000) of the Italian educational system, in which the state still exerts regulative power over schools and local authorities, and where the great discretion of teachers is limited by adherence to the formal rules imposed by the Ministry of Education (Mattei, 2012). Such ambiguity might derive from a ministerial instability (Colombo and Desideri, 2018) and from the “limited effects” of both the 1997/1999 reforms (Grimaldi and Serpieri, 2010: 84) and the of the subsequent 2015 reform in governance/autonomy matters (Barone and Argentin, 2016). The aforementioned issues challenge the *raison d'être* of the SNV apparatus. With insufficient school autonomy, it indeed becomes problematic to make school actors externally accountable for many of their outcomes or to support innovative processes.

*Assumption 3 & 4: If schools use external test results for formative feedback, they will innovate and improve, and if schools innovate in their teaching strategies, they will obtain better results in external tests*

The relationship between accountability and innovation materializes in multiple ways. Firstly, the SNV foresees that, on the basis of standardized external test results and self-analysis, schools should define their improvement plan (*Piani di Miglioramento*). This means that schools should choose which goals they wish to reach over a 3-year period and which pedagogical and organizational actions to undertake in order to reach them. In this sense, external tests are considered not only as an informative tool, but also as a formative one, meaning that they would “help improve teacher’s pedagogical practices” (Academic scholar<sup>1</sup>) and “guide schools in improving their curricular and pedagogical offer, and way of planning” (Member of Ministry<sup>2</sup>). It is believed that, once schools are able to identify their problems, on the basis of evaluative feedback, innovation processes are also possible: “Innovation comes from the solution to problem, and is only possible where you have a clear idea of what your own problems are and what to do with them” (Private foundation<sup>1</sup>). In this sense, the SNV foresees a significant integration of the triad, “external evaluation, self-evaluation and improvement” (Poliandri, 2018). The latter is closely linked with innovation (Faggioli and Mori, 2018) and, at this point, the role of INDIRE comes into play, “helping schools, especially those which find themselves in greater

difficulty, to insert elements in their improvement plans which can give them better chances of succeeding” (Member of Ministry1).

Secondly, according to certain interviewees, external standardized tests are considered to be “instruments of high pedagogical reflexivity” (Academic scholar1), specifically designed to improve student reasoning, since they do not require memorization capacities but rather “require students to understand what they have learnt and to apply it in a new situation” (Representative of INVALSI). Since external tests require students’ reasoning capacities, teachers ideally should take test results as “a stimulus to understand what does not work in their pedagogy and to discuss and reflect on it with other teachers” (Representative of INVALSI). Test results would therefore “guide teachers in changing their teaching strategies” (Academic scholar1). As a counterpart, it is also believed that “innovation practices increase positive results in standardized tests” (Representative of INVALSI). It thus emerges how the SNV and, in particular, the external tests and their features are considered as a change of the current paradigm, as an innovation. The discourse regarding an Italian traditional school model, based on a structured classroom setting, transmissive knowledge, mnemonic exercise and knowledge-based teaching is, in fact, frequently articulated to justify the need to reform the system and introduce external testing:

“The INVALSI tests have certainly been an element of innovation because they have forced us to at least imagine a somewhat different type of learning assessment. (. . .) tests are always more computer based, so they have changed the ritual and changing the ritual has also brought a little innovation” (Private foundation2).

Nonetheless, as the same interviewee points out, the lack of any follow-up given to students is at variance to the goal of fostering student reflection and reasoning (Private foundation2):

“I find it strange that feedback is not given to students (. . .) A person is asked to engage in doing something which is intellectually very complex, without acknowledging the preparation done for it, and no feedback is given. Since we spend time on it and it costs money, let it really be a way to foster deep reflection for every single student engaged in the tests” (Private foundation2).

According to another interviewee, the fact that the INVALSI tests do not measure transversal competencies or subjects other than Italian, Mathematics and English, is also considered to

“undermine the intended effect on increasing students’ critical reasoning and/or stimulating other transversal competencies.” Similarly, the characteristics of the INVALSI tests as being mainly constituted by close-ended questions<sup>7</sup> is also considered as “less stimulating for student’s reasoning” (Academic scholar1).

### Intervention: An equilibrated “three-legged system”?

“As a result of a long policy and trial process” (Member of Ministry2), the interviewees describe the SNV as an articulated and comprehensive intervention, aimed at the achievement of specific outcomes. The SNV is defined by many of the actors as a “three-legged” model (Member of Ministry2; Representative of INDIRE1; Union leader2), referring to the fact that there are three main actors at its forefront: (1) INVALSI—in charge of the coordination of the entire SNV, the definition of evaluation indicators and frameworks, and the production of external standardized tests; (2) INDIRE—responsible for accompanying schools in their improvement actions and innovative practices, and (3) the autonomous external ministerial inspectorate, which is in charge of carrying out a sample-based, external evaluation of schools.

The SNV is made up of different steps, constituting a cyclical process and an evaluation mechanism lasting 3 years, in which the three aforementioned actors are involved. At the first stage, all schools produce a self-evaluation report (*Rapporto di Autovalutazione*, RAV) on the basis of students’ final school results, results in INVALSI standardized tests, as well as a self-analysis, based on a set of items and standards centrally defined by the Ministry. Afterward, schools are expected to develop an improvement plan by identifying both organizational and pedagogical actions, according to the priorities and targets previously pointed out in the self-evaluation report. This step is supported by INDIRE, which helps and supports schools, especially those most in difficulty, to define their improvement plans and actions, and to take care of innovation processes (Presidential Decree 80/2013). In a third step, an evaluation, based on school observations, interviews and an evaluation of reports is conducted by ministerial inspectors within a random sample of schools, with the aim of avoiding the risk of schools being “self-referential.” The last step is represented by so-called “social accountability,” which involves

---

<sup>7</sup> The majority of INVALSI questions are closed answers, even though there are some which require a short or articulated textual answer (Robasto, 2017).

publicizing results, amongst which are the INVALSI standardized test results, to the wider community in a logic of transparency and public responsibility. The system is thus thought of as a cycle, whereby the process of innovation constitutes the results of external evaluation, self-evaluation, improvement plans and social accountability (Faggioli and Mori, 2018).

Although the SNV has been designed as a “three-legged system” (Faggioli and Mori, 2018: 93), the interviews highlight the fact that the INVALSI seems to have much more weight than the other two “legs” of the system. In fact, notwithstanding the “close collaboration” between the INVALSI and INDIRE institutes, since they “both participate in the cabin of the SNV; talking to each other all the time” (Representative of INDIRE1) and do “some research together on evaluation and improvement” (Representative of INVALSI), the role of INDIRE in supporting schools in light of data coming from the SNV is viewed as less central. This is because at INDIRE “work on many things is not done as a direct result of evidence coming from INVALSI tests” (Representative of INDIRE1). Furthermore, the number of external inspectors has been gradually reduced over the years, representing an important challenge, because “a few inspectors cannot guarantee the coverage needed to evaluate all Italian schools” (Member of Ministry2). The fact that “basically the external inspection in Italy does not exist, has clearly made the realization of that [policy] design extremely difficult” (Representative of INDIRE1). The predominance of INVALSI and its standardized tests is also reflected in the fact that according to ministry exponents, when compiling their self-evaluation report, schools often decide to prioritize their results in external standardized tests over pedagogical processes:

“When they [schools] need to identify their different priorities, they choose to look at data coming from the INVALSI tests, because data culture has also grown inside schools, and the more reliable data are clearly that of standardized tests (. . .)” (Member of Ministry2).

The close relationship of the INVALSI with the Ministry of Education is further criticized by the unions because of “the repercussions this can have on the impartiality of the institute and its work” (Union leader2). In light of what has been said, an interviewed union leader representative argues:

“The National Evaluation System was born with three legs (. . .): the only one that has really been working over these thirty years is INVALSI” (Union leader2)

## Rationales behind the SNV: A solution in search of many problems

From the analyzed data, different rationales emerge behind the adoption of the SNV, which refer to the long-term goals that are to be achieved through the intervention and why.

As in other Napoleonic states, the SNV appears to have been adopted as a way of modernizing the governance of the education system under a NPM logic (Verger et al., 2019). In a context of high centralization, this translates into the “decentralization of governance as a way of increasing the efficacy and efficiency of the public administration’s action and bringing it closer to citizens” (Member of Ministry<sup>2</sup>). Following this logic, the SNV has also been adopted as a means of enhancing the “efficiency” of the education system, and external accountability was introduced to “see whether educational efforts were going in the right direction” (Private foundation<sup>2</sup>), thus “reducing public waste” (2007 WP). The adoption of standardized tests, in particular, seems to have derived from an “always increasing preoccupation with levels of school productivity and quality of results” (2007 WP). As in other Southern European countries, this preoccupation with efficiency and quality seems closely linked to the willingness to adhere to international norms and discourses on educational governance (Verger et al., 2019). Indeed, the interviews provide a glimpse into how international pressures and data from international organizations were crucial in determining the adoption of a standardized testing system. In particular, poor Italian results in international PISA tests justify the adoption of the external accountability system (2007 WP). External accountability is in this sense viewed as a way to “improve and harmonize the quality of the education system, with the goal of evaluating its efficiency and efficacy, framing (national) evaluation in the international context” (Legislative Decree 286/2004). The interviews also highlight how solicitations at European level from documents such as *Education at a Glance* (2008) or OECD reports, had been particularly harsh toward Italy, highlighting its “abnormality” because, as opposed to other countries, Italy did not yet have “an essential external assessment system in place to counteract school autonomy” (Member of Ministry<sup>1</sup>). Certainly, these external pressures, together with the socio-political context characterized by the need to contain expenditure due to the economic crisis (Peruzzo et al., 2022), have been relevant in influencing domestic education policy under the center-right government, which started to pilot external accountability through national standardized tests. In this sense, Bordogna (2016) mentioned how the European Union and the European Central Bank sent an official letter to the Berlusconi government during the 2011 financial crisis, soliciting Italy to introduce evaluation, merit and



performance management in education as a way to avoid future default (Grimaldi and Barzanò, 2014).

Nonetheless, from our analysis, other rationales have been put forward which seem to align with findings in other countries (cf. Verger et al., 2019). Firstly, equity and transparency arguments also emerge to justify the adoption of the SNV, as in certain Nordic countries (cf. Camphuijsen and Levatino, 2022). Indeed, the interviews underline that the system serves as a means of diagnosis of critical areas, the identification of which should be followed by interventions in schools, aimed at systematizing practices and reducing learning inequities. The accountability system is, therefore, described as a “photograph, which shows which difficulties schools have, for example, related to their socio-economic and organizational fragility, or educational poverty” (Representative of INDIRE2). In a context characterized by “strong territorial disparities in competences” (2007 WP), this rationale is also linked, according to ministry exponents, to the identification of “the geographical areas which have major difficulties,” with the intention of “reducing the severe geographical and learning gap of the education system” (Member of Ministry2). From the interviews, it also appears that the SNV and its social accountability component, in particular, adheres to a transparency logic and is viewed as a means of empowering citizens’ and parents’ voices to:

“Provide families, students and the local area with tools which enable them to more consciously screen quality improvement and raise the quality of their relationship with the school and teachers” (2007 WP).

Some interviewees, however, place particular emphasis on clarifying how the transparency goal is far from being inserted in a marketized, merit-based logic, neither is it aimed at generating ranking or punitive consequences:

“. . .For us as a Ministry what was important was to provide an evaluation tool, not a tool for judging the level of schools” (Member of Ministry1).

“We use the external evaluation which comes from INVALSI as an instrument to conduct an analysis, rather than a punitive instrument or a classification to understand if we are first or last in national or international rankings” (Representative of INDIRE2).

The system is instead described by the actors as being, by choice, a “reflexive” self-evaluation model, linked to an internal school improvement (Member of Ministry1), as well as a “powerful informative tool,” which might help schools, by informing them of the level of students’ competences (Representative of INVALSI). In this sense, the TBA system is referred to by many actors as solely a means of highlighting in which aspects schools have major difficulties, as a “thermometer” which serves to diagnose school problems:

“Tests are a thermometer, exclusively a thermometer to monitor the temperature of schools (. . .) who has ever said that a thermometer is bad for your health? That’s it, you have to do the tests, then if the result is not good, it’s okay if you take it into account, but at least I have a photograph and photographs do not hurt, X-rays do not hurt” (School leaders’ association).

This emphasis on underlying the harmlessness of the INVALSI standardized test might reflect an internalized way of automatically defending the policy adopted from criticism and opposition, particularly with regard to merit-based awards and the ranking mechanism (Fondazione Agnelli, 2014), which have accompanied the introduction of external, standardized testing in the Italian context for many years (cf. Barzanò and Grimaldi, 2013; Grimaldi and Serpieri, 2012). Specifically, according to the interviewees, the issue of data devolution to schools and their eventual publication has been at the center of the political debate. As one of our interviewees remarks, one of the main points of the criticism received was directed at the government for “being willing to make the learning evaluation public, so that everyone can then rank schools based on that data” (Member of Ministry1). On this topic, the divergent views of the various stakeholders interviewed are evident. On the one hand, a private foundation representative, who was interviewed, highlights the useful role that “[visibilizing test results] brings to school improvement and decision-making” “to allow a more conscious parental school choice” (Private foundation2). On the other hand, from the perspective of the unions, the risk of “[using] INVALSI test data to create school rankings and justify neoliberal policies” is underlined (Union leader2). As the interviewee from the Ministry explains, criticism of the publication of results caused the government to shelve the decision as to whether (and which) results should be made available to the schools themselves on the national web portal (*Scuola in Chiaro*), during “a delicate phase of political mediation with unions” (Member of Ministry1):

“Schools are given the option of making learning outcomes public or not [. . .] it was a wise political choice to avoid the initial prejudices coming to the surface, so we have

given the schools freedom to manage this element, this information regarding their evaluations in terms of transparency for users, particularly parents” (Member of Ministry1).

From the above analysis, it seems that the articulation of the SNV at policy level constitutes a tailored solution to the various problems and characteristics of the Italian educational system (i.e. centralism, bureaucracy, inefficiency, geographical disparities in achievements). Yet, a deeper analysis of the documents and the interviews shows how its adoption has also been largely influenced by globalizing ideas, which are influencing the educational agenda worldwide (Ball, 1998), and seem to have constituted equally important drivers, motivating the adoption of an accountability system in Italy. On the one hand, the justification underpinning the adoption of the SNV is linked to globalization discourses and international competition. A variety of stakeholders acknowledge that, positioned in a global context, externally evaluating and comparing student learning outcomes is “necessary” (Representative of INDIRE1) and is taken for granted (e.g., Private foundation1). On the other hand, beliefs regarding the benefits of datafication in the governance of education seem to form the basis of the whole external accountability apparatus. The main idea is that there is a need for objective, standardized and longitudinal data, and that external experts know how to provide these. According to some interviewees, the legitimation given to the reforms is strengthened by the fact that standardized tests are conducted with “statistically controlled criteria” and that they are “statistically well made” (Representative of INVALSI). Data from an external evaluation are in fact considered an objective and reliable instrument, which provide an “accurate” and “fair” measurement (Technical Document, INVALSI). This would allow schools to identify “what is difficult for schools to see by themselves” (Representative of INVALSI), “avoiding the risk of being self-referential” (Technical Document, INVALSI). External test results are therefore portrayed as:

“...anchors for schools because they are something external. It is not a teacher who proposes to do things in his own way, but an external perspective that is common to everyone” (Representative of INVALSI).

Discourses underpinning such reasoning echo a fetishism for numbers, measurements and comparisons in the educational field, which has been already identified by previous research (e.g. Ball, 2015; Ozga, 2008) and suggests that statisticians, economists or external experts as such

know better than schools or teachers, providing them with “more reliable” data (Representative of INVALSI):

“We [INVALSI] often find ourselves explaining to teachers how to design a test, what it looks like. It’s paradoxical!” (Representative of INVALSI).

This “trust in numbers” also seems to be at the forefront of a criticism, made by a representative of a private foundation, regarding the “lack of objectivity of the school internal self-evaluation report.” According to the interviewee, this instrument would have been introduced as a mere “compromise with trade unions to lower their discontent” (Private foundation2).

Therefore, it remains unclear to what extent the SNV constitutes a means of addressing country-specific problems or whether the equity and school improvement rationales have been mobilized to justify and create consensus around the “idea” of introducing external standardized assessments in a context of welfarist legacy and unions’ contestations, as in Italy. Unions’ power in terms of influencing the decision-making process in education policymaking is however considered to have weakened, as remarked by one of the union leaders:

“There has not been mediation anymore... maybe because in these last few years three/four ministers have changed. We ask for interventions, dialogue on the use of INVALSI data, but paradoxically, on these issues, we have more relationships and dialogue with INVALSI or the Agnelli Foundation than with the Ministry itself” (Union leader2).

### The context of intervention: A theory of no-change?

In a theory of change, change is the result of specific mechanisms activated under specific circumstances, meaning that certain contexts support change, while others hamper it. In our analysis, several contextual aspects have emerged that are believed to challenge the effective realization of the SNV programmatic idea. These are related to (a) the structural features of the Italian educational context; (b) school actors’ characteristics and competences, and (c) cultural features.

### *Structural features of the Italian educational context*

According to our interviewees, “the precarious conditions of school buildings,” “inadequate spaces” (Private foundation1), “obsolete material” and “rigid timetables” are particularly responsible for “rendering innovative processes difficult” (Representative of INDIRE1), especially as they challenge the advancement towards a “non-traditional” frontal teaching model:

“There is still a fragmented timetable, even the school environment is built around a frontal lesson model: teacher’s desk, blackboards, teachers speak and the others listen; like assessments, they often evaluate a taught knowledge” (Representative of INDIRE1).

Furthermore, the economic precariousness (and thus low attractiveness) of the teaching profession in Italy is also viewed as “an obstacle to realizing good teaching” (School leaders’ association), because “if you want to have positive personalities that work for the future in a constructive and innovative manner, in my opinion, you need to have an education system which values teachers” (Academic scholar1). The lack of compulsory or “adequate” teacher training is another factor that hampers the successful realization of policy expectations:

“If we had to change something and no minister has had the courage to do so yet, we would make teacher training compulsory because it is not possible to repeat the same things year after year, let alone if one started teaching 30 years ago and thinks that with those same methods good results can be achieved” (Representative of INDIRE2).

A “lack of human resources,” such as “middle management,” which should support principals in their work, is also considered by some of our interviewees to “limit the effectiveness of implementing real autonomy in schools and promoting innovative processes” (Private foundation1).

### *School actors’ personal and professional characteristics*

School actors, specifically teachers and principals, are blamed for lacking the professional competences and personal characteristics to facilitate school innovation. On the one hand, principals are blamed for lacking the “right personality” and “pedagogical vision” (Representative of INVALSI), or “the capacity to read data and know how to use it” (Private foundation1).

Teachers, on the other hand, are portrayed as “lacking the energy needed to trigger the mechanisms of change” (Representative of INVALSI), “lazy” and “unwilling to change” (School leaders’ association). They have no desire “to improve their teaching strategies or to better train” (Representative of INDIRE1):

“A problem that we have encountered every time we propose training, is the fact that the training of teachers and workers, in general, tends to be voluntary (. . .) Therefore, if you organize a training course, say, ‘à la carte’, everyone orders the dish they prefer” (Private foundation2).

The “inadequate attitude toward change and training” attributed to teachers has an impact on the enactment of school autonomy, because “a teacher who wants to do things superficially, does not even feel the need to change, and asks students to adapt to his/her teaching model” (Representative of INDIRE1). Teachers are also viewed as “being inadequately prepared” (Representative of INDIRE2, Private foundation1), “lacking knowledge or competences about learning processes” (Academic scholar2) and “an international vision” (School leaders’ association). Such aspects are considered central in explaining why pedagogical improvement, project capacity, change or innovative processes fail:

“The vast majority of teachers at all levels and grades in school are completely unaware of how learning processes take place, they do not know (. . .) the problem lies in the fact that it is impossible to find 10 school staff members, who are truly capable of managing an active, interactive, collaborative and dynamic classroom and collaborate with other teachers” (Academic scholar2).

“If we look at the Italian school today, the aspect that is most surprising and depressing is the general impoverishment, a sharply lowered level. . . a widespread disinterest, an inability of teachers not only to fascinate but also to operate the new techniques” (Union leader1).

Beyond limiting school autonomy and innovation, a teacher’s “lack of substantial preparation” is also considered to negatively influence the use of testing as a tool to generate student reasoning and pedagogical change, which is at the basis of the policy’s theory of change, as exemplified by the following quote:

“There is a strong gap between what they [teachers] think and what they do, that is, they think of themselves as teachers who promote students’ cognitive activity, but in practice they deliver a lesson, ask questions and give out homework” (Representative of INVALSI).

### *Cultural aspects and ingrained beliefs*

The last set of contextual issues, identified as undermining the successful realization of the policy expectation, is linked to an embedded traditional, knowledge-based culture, which is considered as preventing schools and the educational system from changing substantially. As one of the interviewees notes, “a culture based on knowledge cannot suddenly certify competences” (Representative of INDIRE1). In addition, a school model which, for years, has relied on a theoretical, frontal teaching method is seen as “blocking other innovative ways of working and organizing teaching” (Representative of INVALSI). Moreover, the traditional way of primarily utilizing and relying on textbooks for teaching is seen by the actors as “leveling down” and impeding change in pedagogical practices, for instance, preventing teachers from analyzing students’ difficulties or connecting disciplines (Academic scholar1).

Many different actors also highlight the lack of a culture of evaluation in terms of preventing schools and teachers from considering the SNV and standardized tests positively. A lack of evaluation culture in Italy is associated with “not believing in evaluations at all” (Representative of INDIRE1), “a catholic vision which makes us believe that someone is constantly judging and punishing us” (School leaders’ association), “a lack of self-analysis and evaluation which should be interiorized” (School leaders’ association):

“Italy does not believe in evaluation. Basically, there is a cultural problem related to this. We do not believe in the evaluation system, because it is always viewed as a means of condemnation rather than a means of improvement” (School leaders’ association).

Such a lack of evaluation culture is also seen as the reason for the increasing opposition and resistance of teachers to standardized tests and to policy evaluations, in general, because “no-

one likes to be evaluated,” thus sustaining schools’ capacity to develop “antibodies” to such a testing approach (Private foundation2).

Consequently, contextual, systemic and cultural aspects, which are at the basis of the aspired mechanisms of change, seem to be paradoxically considered as significant obstacles rather than enabling factors. As already underlined by Barone and Argentin (2016: 146) in relation to the 2015 governance/ autonomy reform, it also seems that the system is supposed to operate in a scenario that assumes “a rather unrealistic view of key actors who are called to implement it.” For this reason, rather than fostering a concrete change, many aspects of the SNV system seem to have an aspirational status.

### The achievement of the SNV outcomes: Misunderstanding and misuse

The interviews provide information regarding the extent to which the SNV intentions are being met, and shed light on various aspects that, according to our informants, undermine the achievement of the SNV outcomes.

The inadequacy regarding the way in which the SNV and INVALSI test results are interpreted and used by school actors is one of the main challenges identified by the interviewees. According to them, schools often “misunderstand the purpose and usefulness of the accountability system, and wrongly perceive it as a way to judge and rank them” (Representative of INVALSI). It is believed that criticism of the SNV derives from the fact that the real purpose of the policy, which is “merely diagnostic and informative” (School leaders’ association) and “solely aimed at fostering school improvement” (Member of Ministry2), is misunderstood:

“A misconception is that the tests were a way of judging schools as good or bad, a way of judging teachers as good and bad, a way of judging students in the best class” (School leaders' association).

In relation to this, apart from the aforementioned lack of evaluation culture, interviewed actors admit that “the external communication deriving from INVALSI and/or the ministry is not sufficiently successful in terms of sustaining this culture of evaluation and promoting a positive and clear message regarding the usefulness of national standardized tests” (Member of Ministry2).



Misunderstanding related to poor communication is considered by interviewees to lead schools to discount negative test results, rather than taking advantage of them to review and improve their practices:

“In some cases, they [schools] break the thermometer, and they say that the tests are wrong, that their students are much better, and so on” (Representative of INDIRE1).

This also leads to “undesired effects such as cheating or distorting test results” (Private foundation2). By misinterpreting the policy, schools often use the accountability instruments in a “superficial and automatic way” (Academic scholar2), without adequately documenting and reporting what they do, as foreseen in the SNV:

“This is the greatest difficulty, encountered by schools, is how to provide evidence of the achieved results. Maybe because they are not used to document what they do (. . .) [documenting and reporting] is often perceived as compiling papers, but actually, [the real aim of documenting and reporting] is to make evident the causal links of what I did, the results I got and what I need to do. This is still critical, and we are still working on it” (Member of Ministry2).

To a great extent, this use of accountability instruments as a bureaucratic requirement contradicts the efficiency and de-bureaucratization goals of policy intervention.

According to some interviewees, the misuse of data is also associated with schools’ incapacity to adequately use data from test results, for instance, schools often do not know how to use “all of that data we give them” (Representative of INVALSI), and they “have difficulty in transforming such results into action and improvement plans” (Member of Ministry1). The interviewed actors believe that schools have “difficulty prioritizing and programming their actions within an overarching framework of three years and in relation to the test results, instead, they do so in a fragmented and chaotic manner” (Member of Ministry2). Such “incapacity to plan in advance” is therefore considered to undermine the whole SNV machinery to the extent it “negatively impacts their [schools’] capacity to change and innovate, and to effectively use the autonomy given” (Private foundation1).

According to the key actors interviewed, both the misunderstanding and the misuse of standardized testing have been fueled by publishers opportunistically selling books of “poor

quality,” aiming at helping teachers prepare students for the INVALSI tests, but “spreading an erroneous message that students need to be specifically and intensively trained for this” (Member of Ministry2). Related to this aspect, several actors also recognize and criticize the risk of teachers “teaching to the test” (Member Ministry2; Union leader2; Private foundation1):

“If those who should support didactic improvement interpret the test as a multiple-choice test, they will interpret this as ‘training’ in relation to an operational procedure but will not proceed in the direction of achieving competence through learning outcomes” (Member of Ministry2).

The relatively low policy impact of the INVALSI test results is another aspect, which is perceived as preventing data being used in policy actions or influencing political decisions, contrary to expectations. Moreover, the lack of any support provided to the schools with a negative evaluation is seen as undermining the diagnostic and equity goals of the SNV. Indeed,

“What is missing today - and we are also working with the Ministry regarding this - is specific action in light of the INVALSI tests. ‘You are a school that has a problem, and you need help’. I think this is a little bit lacking” (Private foundation1).

Finally, from the perspective of INVALSI representatives, the fact that standardized tests are not compulsory for students, as they are for teachers<sup>8</sup>, is considered problematic since this results in students and their families not perceiving the standardized tests as important, therefore boycotting them more easily. Nevertheless, the obligatory nature of INVALSI tests for schools is seen by some actors as a measure that has strengthened schools’ resistance and skepticism toward the tests, as illustrated in the following quote:

“The obligation is a bit like the vaccine. So, if they offer it to you then ok, but if they oblige you to do it, it’s completely bad” (Academic scholar1).

---

<sup>8</sup> The implementation of INVALSI tests are part of teachers’ service obligations (“obblighi di servizio”).

## Conclusions

The present study has analyzed the program ontology of the SNV in its current configuration, with a special focus on the relationship between school autonomy, accountability and innovation. The study has also explored the existence of pitfalls and tensions that might hamper the achievement of the declared policy intentions of the SNV. Based on key policy and technical documents, as well as interviews with key educational actors, the analysis was guided by the concept of the “theory of change” and its constituent elements.

The findings show how school autonomy, innovation and (external and internal) school accountability have been articulated and constructed together as a powerful *dispositif*. The Italian case seems to be rather unique, to the extent that TBA is claimed to be an educational innovation device. This is also an example of how already existing themes can be framed as an innovation within the discursive promotion of the “managerial recipe” in education (as highlighted by Serpieri et al., 2015 for the case of school self-evaluation). In the official discourse, external standardized tests are described as a means that foster reflexivity and change in pedagogical practices. In this sense, they are often considered to be an innovative tool in an educational context, as in Italy, characterized by a teacher-centered theoretical didactic culture. Nonetheless, although at first glance, the assumptions seem to be well articulated, a deeper analysis reveals tensions in the arguments used by promoters of the reform to legitimize and justify it. School autonomy, which is considered a primary reason for the introduction of TBA, seems to be lacking in practice, especially in relation to certain managerial and financial aspects. At the same time, innovation, which constitutes a crucial goal of the policy, is not uniformly defined. Similarly, external standardized tests, the design of which is assumed to measure and foster students’ reasoning and innovative practices, are at the same time criticized by certain actors for limiting students’ reasoning and for not testing transversal competencies.

The current Italian SNV is the result of a long process of reforms, piloted and implemented under different governments over the last 20 years, and promoted by both left and right-wing parties. Solicitations from international organizations (EU, OECD), the domestic economic crisis and the consequent austerity measures that characterized the Italian context during those years, have emerged as contextual conditions that acted as a “window of opportunity” (Kingdon, 1984) for the adoption of the reforms. The role and realm of different actors involved in their design and implementation—ranging from public-ministerial, research institutes, professional associations and private foundations—and the struggles between different interest groups appear crucial and

seem to have influenced the aforementioned tensions. According to our findings, the SNV is supported by the two in-house agencies (INDIRE and INVALSI) in charge of its implementation, is defended and taken for granted by exponents of private foundations and is criticized by labor unions, especially regarding certain specific aspects. At the same time, the role of the education ministry in facilitating the mediation process clearly appears. The criticism of teacher unions in relation to ranking mechanisms, discontent and boycotts (Barone and Argentin, 2016) seems to have played a key role in determining the final result of the policy and its specificities (e.g. the non-autonomic publicization of INVALSI test data), including the way in which the intervention is currently being defended and communicated by policy actors (e.g. the emphasis on the harmlessness of INVALSI standardized tests, on internal evaluation and on the formative and reflexive components of the SNV). Nonetheless, according to the union leaders interviewed, the unions' power to influence decision-making processes in education has been decreasing. This confirms the findings highlighted by recent research on the weakening of labor unions as a space for collective bargaining, which results in fewer mediation opportunities and unilateral decisions taken by the Italian government in light of the economic crisis and the increased importance of new private actors (Peruzzo et al., 2022; Sorensen et al., 2021).

The SNV has been referred to as a “three legged” model (Faggioli and Mori, 2018), whereby three main bodies (INVALSI, INDIRE and the ministerial inspectors) are in charge of guaranteeing the successful implementation of the system and its expected outcomes. However, findings show that the insufficient number of external inspectors, the less central role of INDIRE compared to INVALSI and the fact that schools are giving more weight to the test results in their internal report seem to render it “crippled” to a certain extent, as one of its components (INVALSI and its standardized tests) overshadows the others.

In line with the findings of Verger et al. (2019), as in other Napoleonic states, the school autonomy with accountability reform package in Italy has been adopted with the aim to modernize, de-bureaucratize and improve the quality and efficiency of the educational system, and to adhere to international norms and discourses on educational governance. At the same time, however, the analysis has highlighted the role played by other rationales. Similar to what has been found in Nordic countries (cf. Camphuijsen and Levatino, 2022), equity and transparency discourses are used to justify the adoption of the SNV. Globalizing ideas (for instance, the idea of learning achievement as an element of international competition and belief in the benefits of datafication) also seem to have acted as important drivers of the SNV reform. It can be

questioned to what extent these different rationales have played a major or minor role, or whether the weight of each of the rationales has changed over time. In particular, it remains unclear whether the emphasis placed on the self-evaluation report and its school improvement logic and on the equity rationales constitutes an adaptation of global norms and trends (Steiner-Khamsi, 2014) or, as some interviewees argue, whether this was part of a political maneuver to increase acceptance, gain legitimacy and “coat” external accountability with socially desirable arguments, in a context characterized by harsh contestations and protests (e.g. Kickert, 2007).

The analysis also reveals how, according to the key actors interviewed, the context of intervention is not entirely conducive to triggering the expected change mechanisms. Indeed, they claim that a set of challenges related to the features of the education system and to the professional and personal characteristics of school actors limit the possibility to implement real autonomy and promote innovative processes. In this sense, a clear opposition emerges between the supposed virtues of external testing, external experts, reliable data on the one hand and the incompetence of school actors on the other. The continuous blaming of school actors, by almost all of the interviewed actors, union leaders included, reveals a strong sense of distrust in teachers, who are not only seen as requiring external guidance and control, but are also incapable of taking advantage of the benefits of the SNV, even hampering its correct implementation. The recurrent use of metaphors plucked from the medical sphere (such as the idea of curing, “thermometer” and diagnosis) also suggests the idea of a vicious, pathological education system, which is clearly opposed to the virtues of external experts. This last point might explain the weight of non-education actors and private consultancies in terms of education knowledge production in Italy (cf. Grimaldi and Serpieri, 2013; Serpieri et al., 2015).

Ingrained beliefs relating to learning, evaluation and knowledge-transmission are also considered to render the aspired substantial change difficult. For the interviewed experts, in particular, a contingent “lack of culture of evaluation” in the Italian context is considered the main reason behind school actors’ misinterpretation of the SNV and its aims. Certainly, the interviewees also admit that an unsuccessful institutional communication around the SNV has contributed to the spread of negative beliefs, which potentially leads to resistance and opposition, as well as to undesirable practices and the superficial use of data. All this is seen by the interviewees as undermining certain policy objectives and seems to correspond to a frustrating aspiration to change “culture” through policy instruments.

To the extent that the present analysis explores the program ontology and degree of success of a relatively recent policy intervention, which only completed its first cycle of accountability in 2019, it contributes to fostering a reflection regarding the coherence of the SNV premises and the realization of its goals. Furthermore, the paper underlines the benefits of using the concept of the theory of change as an analytical tool to deepen the understanding of how a policy is expected to work, as well as to explore its weaknesses and contradictions. The neo-institutionalist approach has also been regarded as a useful, theoretical lens in terms of understanding how global models of reform are created and shaped by institutional contexts, even though our findings seem to indicate that more nuanced categorizations are needed. The analysis finally confirms the relevance of considering not only the context of text production (Ball, 1993), but also the way in which different actors, involved in the whole process, make sense of and interpret policy expectations and limitations, so as to better understand the discrepancy between stated policy goals and policy realization.

Certainly, it is also fundamental to explore the “context of enactment” of policies (Ball et al., 2012). To identify the role of school actors, an analysis of their beliefs and practices is, therefore, an interesting and potential future line of enquiry. The analysis of how the expected relationship between accountability, autonomy and innovation concretely deploys within different Italian schools thus constitutes a promising avenue for future research. This would indeed facilitate an understanding of the way in which schools deal with and respond to such policy expectations in different ways and to what extent and under which circumstances, the challenges identified by key actors hamper the realization of policy expectations.

## References

- Andreas B, Fernie S and Dainty A (2022) Understanding policy and change: Using a political economy analysis framework. *Construction Management and Economics* 40: 865–883.
- Ball S (1993) What is policy? Texts, Trajectories and Toolboxes, *Discourse: Studies in the Cultural Politics of Education* 13(2): 10–17.
- Ball SJ (1994) *Education reform – A critical and post-structural approach*. Buckingham: Open University Press.
- Ball SJ (1998) Big policies/small world: An introduction to International perspectives in education policy. *Comparative Education* 34(2): 119–130.

- Ball SJ (2015) Education, governance and the tyranny of numbers. *Journal of Education Policy* 30(3): 299–301.
- Ball S, Maguire M and Braun A (2012) *How Schools Do Policy: Policy Enactments in Secondary Schools*. New York: Routledge.
- Barone C and Argentin G (2016) School Reform: Innovation and the rhetoric of Change. *Italian Politics* 31(1): 135–154.
- Barzanò G and Grimaldi E (2012) “Policy” valutativa e contesti di applicazione. caratteristiche procedurali. *ECPS - Educational, Cultural and Psychological Studies* 06: 159–189.
- Barzanò G and Grimaldi E (2013) Discourses of merit. The hot potato of teacher evaluation in Italy. *Journal of Education Policy* 28(6): 767–791.
- Bifulco L, Monteleone R, Mozzana C, et al. (2010) *The institutional mapping of Italy’s education system: Europeanization, centralization and regionalization*. In: WorkAble 3rd Workshop - FP7 EU research project “Making Capabilities Work,” Lausanne, 17–19 November.
- Bordogna L (2016) Italy: The uncertainties of endless reform. public service employment relations confronting financial sustainability constraints. In: Bach S and Bordogna L (eds) *Public Service Management and Employment Relations in Europe: Emerging From the Crisis*. London: Routledge, pp.84–111.
- Boyatzis RE (1998) *Transforming Qualitative Information: Thematic Analysis and Code Development*. Thousand Oaks, CA: SAGE.
- Cairney P (2012) *Understanding Public Policy: Theories and Issues*. New York: Palgrave Macmillan.
- Camphuijsen MK and Levatino A (2022) Schools in the media: framing national standardized testing in the Norwegian press, 2004–2018. *Discourse Studies in the Cultural Politics of Education* 43: 601–616.
- Christensen S and Molin J (1995) Origin and transformation of organizations: Institutional Analysis of the Danish red Cross. In: Scott WR and Christensen S (eds) *The Institutional Construction of Organizations*. Thousand Oaks, CA: SAGE Publications, pp.67–90.
- Colombo M and Desideri A (2018) Italy. The Italian education system and school autonomy. In: Martins S, Capucha L and Sebastião J (eds) *School Autonomy Organization and Performance in Europe. A Comparative Analysis for the Period From 2000 to 2015*. Lisbon: CIES – ISCTE (Centre for Research and Studies in Sociology), pp.101–113.

- DiMaggio PJ and Powell WW (1991) Introduction. In: Powell WW and DiMaggio PJ (eds) *The New Institutionalism in Organizational Analysis*. Chicago, IL: University of Chicago Press, pp.1–38.
- European Commission, Directorate-General for Education Sport and Culture (2018) *Study on supporting school innovation across Europe: final report*. Report, European Commission: Publications Office.
- Faggioli M and Mori S (2018) Valutare l'innovazione scolastica: Vincoli e opportunità del Sistema Nazionale di valutazione in. In: Freddano M and Pastore S (eds) *PER UNA VALUTAZIONE DELLE SCUOLE OLTRE L'ADEMPIMENTO. Riflessioni e Pratiche Sui Processi Valutativi*. Milano: Franco Angeli s.r.l, pp.88–100.
- Falabella A (2014) The performing school: The effects of market & accountability policies. *Education Policy Analysis Archives* 22(70): 70.
- Falabella A (2020) The ethics of competition: Accountability policy enactment in Chilean schools' everyday life. *Journal of Education Policy* 35(1): 23–45.
- Ferrer-Esteban G (2016) *Deconstructing the [Italian] Black-Box. Gli Approcci Didattici Degli Insegnanti Italiani e Le Performance Degli Studenti*. Torino: Fondazione Giovanni Agnelli.
- Fondazione Agnelli (2014) *La Valutazione Della Scuola. A Che Cosa Serve e Perché È Necessaria all'Italia*. Roma-Bari: Laterza.
- Grek S and Ozga J (2010) Re-inventing public education: The new role of knowledge in education policy making. *Public Policy And Administration* 25(3): 271–288.
- Grimaldi E and Barzanò G (2014) Making sense of the educational present: Problematising the “merit turn” in the Italian eduscape. *European Educational Research Journal* 13(1): 26–46.
- Grimaldi E and Serpieri R (2010) The reforming trajectory of the Italian educational system. Site-based management and decentralisation as a challenge for democratic discourse. *Journal of Educational Administration and History* 42(1): 75–95.
- Grimaldi E and Serpieri R (2012) The transformation of the Education State in Italy: A critical policy historiography from 1944 to 2011. *Italian Journal of Sociology of Education* 4(1): 146–180.
- Grimaldi E and Serpieri R (2013) Jigsawing education evaluation. Pieces from the Italian new public management puzzle. *Journal of Educational Administration and History* 45(4): 306–335
- Grimaldi E and Serpieri R (2014) Italian education beyond hierarchy: Governance, evaluation and headship. *Educational Management Administration & Leadership* 42(4\_suppl): 119–138.



- Hall D, Grimaldi E, Gunter HM, et al. (2015) Educational reform and modernisation in Europe: The role of national contexts in mediating the new public management. *European Educational Research Journal* 14(6): 487–507.
- Hall PA and Taylor RCR (1996) Political science and the three new institutionalisms. *Political Studies* 44(5): 936–957.
- Hay C (2006) Constructivist institutionalism. In: Rhodes R, Binder S and Rockman B (eds) *The Oxford Handbook of Political Institutions*. Oxford: Oxford University Press, pp. 56–74.
- Hood C and Peters BG (2004) The Middle Aging of new public management: Into the age of Paradox? *Journal of Public Administration Research and Theory* 14(3): 267–282.
- Hursh D (2005) The growth of high-stakes testing in the USA: Accountability, markets and the decline in educational equality. *British Educational Research Journal* 31(5): 605–622.
- Karlsen GE (2000) Decentralized centralism: Framework for a better understanding of governance in the field of education. *Journal of Education Policy* 15(5): 525–538.
- Kickert W (2007) Public management reforms in countries with a Napoleonic state model. In: Pollitt C, Van Thiel S and Homburg V (eds) *New Public Management in Europe: Adaptation and Alternatives*. Basingstoke: Palgrave/Macmillan, pp.26–51
- Kingdon JW (1984) *Agendas, Alternatives, and Public Policies*. Boston, MA: Little, Brown
- Knight R (2020) The tensions of innovation: Experiences of teachers during a whole school pedagogical shift. *Research Papers in Education* 35(2): 205–227.
- Krasner SD (1988) ‘Sovereignty: An institutional perspective’. *Comparative Political Studies* 21(1): 66–94.
- Landri P (2014) Governing by standards: The fabrication of austerity in the Italian education system. *Education Inquiry* 5(1): 25–41.
- Looney JW (2009) *Assessment and innovation in education*. OECD Education Working Papers, No 24.
- Lubienski C (2009) *Do Quasi-markets Foster Innovation in Education? A comparative perspective*. Report, OECD Education Working Papers, No. 25. Paris: OECD Publishing.
- Malen B, Croninger R, Muncey D, et al. (2002) Reconstituting schools: “Testing” the “theory of action.” *Educational Evaluation and Policy Analysis* 24(2): 113–132.
- Martens K, Nagel A-K, Windzio M, et al. (2010) *Transformation of Education Policy*. Basingstoke: Palgrave Macmillan.

- Mattei P (2012) Market accountability in schools: Policy reforms in England, Germany, France and Italy. *Oxford Review of Education* 38(3): 247–266.
- Meyer JW (2008) Reflections on institutional theories of organizations. In: Greenwood R, Oliver C, Sahlin K, et al. (eds) *The Sage Handbook of Organizational Institutionalism*. Thousand Oaks, CA: SAGE, pp.790–812.
- Meyer JW and Rowan B (1977) Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology* 83(2): 340–363.
- OECD (2013a) *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*. Report, Paris: OECD Publishing.
- OECD (2013b) *TALIS 2013 Results: An International Perspective on Teaching and Learning*. Report. Paris: OECD Publishing.
- Ozga J (2008) Governing knowledge: Research steering and research quality. *European Educational Research Journal* 7(3): 261–272.
- Pensiero N, Giancola O and Barone C (2019) Socioeconomic inequality and student outcomes in Italy. In: Volante L, Schnepf S, Jerrim J, et al. (eds) *Socioeconomic Inequality and Student Outcomes: Cross-National Trends, Policies, and Practices*. Singapore: Springer, pp.81–94.
- Peruzzo F, Grimaldi E, Arienzo A, et al. (2022) New public management reforms and industrial relations in the Italian education system. A cultural political economy approach. *Journal of Educational Administration and History*. 1–19.
- Poliandri D (2018) Per un sistema nazionale di valutazione come strumento per il miglioramento delle scuole. In: Freddano M and Pastore S (eds) *PER UNA VALUTAZIONE DELLE SCUOLE OLTRE L'ADEMPIMENTO. Riflessioni e Pratiche Sui Processi Valutativi*. Milano: Franco Angeli s.r.l, pp.67–87.
- Pollitt C (2007) Convergence or divergence: What has been happening in Europe? In: Pollitt C, Van Thiel S and Homburg V (eds) *New Public Management in Europe*. London: Palgrave Macmillan UK, pp.10–25.
- Pollitt C and Bouckaert G (2011) *Public Management Reform. A Comparative Analysis—New Public Management, Governance, and the Neo-Weberian State*, 3rd edn. Oxford: Oxford University Press.
- Reinholz DL and Andrews TC (2020) Change theory and theory of change: what's the difference anyway? *International Journal of Stem Education* 7(1): 2–12.

- Robasto D (2017) La sfida del miglioramento. Tra modelli ed esiti dei processi autovalutativi the challenge of improvement. Between models and outcomes of self-assessment processes. *Form@re - Open Journal per La Formazione in Rete* 17(3): 205–215.
- Sahlberg P (2006) Education reform for raising economic competitiveness. *Journal of Educational Change* 7(4): 259–287.
- Schmidt VA (2008) Discursive institutionalism: The explanatory power of ideas and discourse. *Annual Review of Political Science* 11: 303–326.
- Serpieri R (2009) A ‘war’ of discourses. The formation of educational headship in Italy. *Italian Journal of Sociology of Education* 4(1): 121–142.
- Serpieri R, Grimaldi E and Vatrella S (2015) School evaluation and consultancy in Italy. Sliding doors towards privatisation? *Journal of Educational Administration and History* 47(3): 294–314.
- Sorensen TB, Grimaldi E and Gajderowicz T (2021) *Rhetoric or Game Changer: Social Dialogue and Industrial Relations in Education Midst EU Governance and Privatization in Europe*. Brussels: European Trade Union Committee for Education (ETUCE-CSEE).
- Steiner-Khamsi G (2014) Cross-national policy borrowing: Understanding reception and translation. *Asia Pacific Journal of Education* 34(2): 153–167.
- Stone D (2012) Transfer and translation of policy. *Policy Studies* 33(6): 483–499.
- Thelen K and Steinmo S (1992) Historical institutionalism in comparative politics. In: Steinmo S, Thelen K and Longstreth F (eds) *Structuring Politics: Historical Institutionalism in Comparative Analysis*. Cambridge: Cambridge University Press, pp. 1–32.
- Verger A, Fontdevila C and Parcerisa L (2019) Reforming governance through policy instruments: How and to what extent standards, tests and accountability in education spread worldwide. *Discourse Studies in the Cultural Politics of Education* 40(2): 248–270.
- Verger A and Parcerisa L (2017) *Accountability and education in the post-2015 period: International trends, enactment dynamics and socio-educational effects*. Report, commissioned for the 2017/8 Monitoring Report, Accountability in Education: Meeting our Commitments. Paris: UNESCO GERM.
- Zahariadis N (2003) *Ambiguity and Choice in Public Policy*. Washington, DC: Georgetown University Press.
- Zahariadis N (2016) *Handbook of Public Policy Agenda Setting*. Cheltenham and Northampton, MA: Edward Elgar.

## CHAPTER 4: NEGOTIATING BETWEEN THE ACCOUNTABILITY AND THE INNOVATION MANDATES: EVIDENCE FROM ITALIAN SCHOOLS

### Introduction

Over the past few decades, numerous reform efforts have been directed towards cultivating innovation processes in schools<sup>9</sup>. Educational innovation, intended as a significant change in educational practices (Vincent-Lancrin, 2019), despite its pedagogic focus<sup>10</sup>, is also a process that relies on school governance reforms. Specifically, the promotion of educational innovation is often accompanied by an increased autonomy granted to schools for the implementation of transformative changes. Schools with greater autonomy are expected to have more freedom to innovate, and to enhance their understanding of effective practices tailored to specific contexts and individuals (Looney, 2009). Yet, an increased autonomy for schools is frequently coupled with heightened external control tied to students' performance in national large-scale assessments (NLSA). Governments are willing to delegate more autonomy to schools to the extent to which schools are willing to undergo more intensive monitoring based on their outcomes (Fahey & Koester, 2019; Verger & Parcerisa, 2017). These policy instruments typically manifest in configurations of test-based (TBA) or performance-based accountability (PBA), encompassing the delineation of learning standards, students' performance measured through external standardized tests, the specification of school autonomy, and the implementation of accountability mechanisms (Verger et al., 2019).

---

<sup>9</sup> 1. *Process or product innovations* in schools typically involve changes in teaching and learning (OECD - *Oslo Manual for Measuring innovation*).

<sup>10</sup> Pedagogical innovations, intended as new changed processes for delivering services, such as new pedagogies and teaching practices (European Commission, 2018) usually includes learner-centred education and interactive instruction (Burns & Paniagua, 2018), active methodologies, student-centred learning and curricular integration (SITES, 2009), the application of skills and knowledge to real-life challenges, in opposition to a traditional teacher-centred classroom (Sahlberg, 2006).

The relationship between the two mandates, namely, PBA on the one side, and innovation and autonomy on the other, is however not clear. Research exploring the impact of PBA on educational practices presents conflicting outcomes (Mittleman & Jennings, 2018), also shaped by the interpretations and adaptations of policy environments within schools by educational actors (Diehl & Golann's, 2023; Dulude & Milley, 2021). On the one hand, PBA pressures may serve to communicate good models of teaching and learning (Herman, 2004), motivating school actors to challenge conventional practices (Fahey & Koester, 2019). Conversely, these pressures can influence instructional practices (Avalos et al., 2020), by regulating the structure of curricular content, encouraging curriculum narrowing and endorsing teacher-centered transmissive instruction (Au, 2007; Barrett, 2007). Moreover, beyond *process innovations* in education, *organizational or administrative school innovations*<sup>11</sup> may have less direct relevance for the classroom changes but affect the position of a school within a school market (Lubienski, 2009). Additionally, the pressure to obtain favorable results in large scale assessments can have an adverse impact on the creativity of teachers (Appel, 2020) and their autonomy in educational planning (Farvis & Hay, 2020), by dissuading teachers from taking risks and limiting the time available for innovative or creative practices (Falabella, 2020; Sahlberg, 2009; Knight, 2020).

At the same time, there has been a gradual shift from traditional forms of accountability to more school-owned and school-driven models (MacBeth, 2008). This transformation is manifested through internal forms of accountability, where schools hold themselves accountable for their performance, often through internal processes and assessments, and where crucial elements of accountability are generated largely within school staff, fostering strong professional communities linked to enhanced student performance (Firestone, 2004)<sup>12</sup>. Schools with established internal accountability tend to be more coherent and effective as organizations and more responsive to external accountability demands, since they are more skillful in choosing the curricular areas to prioritize, determine the instructional strategies related to performance measures, and learn how to handle external pressures in alignment with their own core values (Elmore, 2005a). However,

---

<sup>11</sup> Process innovations in education occur at the classroom level, involving teaching and learning, curriculum changes, teaching methods or other programmatic options. Administrative innovations (including marketing or other organisational innovations) instead typically occur in the areas of management, administration, governance of a school and involve substantive changes in the structures or organizational behaviour of schools (Lubienski, 2009).

<sup>12</sup> Internal accountability is intended as the alignment of individual values with collective expectations, reinforced by the processes of accountability (Elmore, 2005a). Internal accountability, including forms of teacher and school self-evaluation and assessment, precedes and determines all school responses to their external environment (Elmore, 2005b).

research also indicates that the interaction between internal and external accountability is not always clear-cut. Without the appropriate internal standards, strong leadership, professional trust and staff capacity to collaborate, external accountability may not consistently promote school responses in alignment with external demands (Elmore, 2005a; Firestone 2004).

Furthermore, the impact of PBA and innovation policies may not be uniform across all schools. In accountability contexts, schools on probation often adopt strategic instructional approaches, tailoring teaching to match test content or narrowing the curriculum to boost test results while avoiding significant changes to their methods or principles (Mittleman & Jennings, 2018). The socio-economic school contexts may also influence teacher practices, such as using test data to categorize students (Hardy et al., 2019), tailoring instruction based on performance, and prioritizing practical-oriented teaching in lower-performing schools (Diamond, 2007).

Whilst previous research focused on school enactment in relation to PBA policies in various contexts (e.g., Falabella, 2020; Gunnulfson et al., 2018; Landri, 2021; Paletta et al., 2020), specific studies exploring the intricate relationship between PBA policies and pedagogical innovation are scarce. In this context, Italy serves as an understudied and relevant case to comprehend how school actors interpret and enact the dual accountability-innovation mandate. Recent reforms in Italy have expanded school autonomy and PBA to incentivize pedagogical innovation and teaching flexibility (Checchi & Mattei, 2021; Paletta et al., 2020). Innovation is inherently part of the National Evaluation System (SNV) (Faggioli & Mori, 2018), an accountability model combining internal forms of accountability (through school self-evaluation reports and improvement plans) and external accountability components (through national student assessments and ministerial inspections) to foster reflexivity and change in pedagogical practices (Paletta et al., 2020). However, despite accountability and innovation being considered equally important dimensions in the Italian policy framework, implementation challenges emerge and the data-intensive external accountability seems to carry more weight than the innovation component (Mentini & Levatino, 2024).

The paper aims to analyse a) how the tension between PBA and innovation unfolds at the school level by examining how school actors understand and adapt their educational practices to the policy mandates, and b) how the local school context influences these dynamics.

The paper follows this introduction with an overview of the Italian reform context. Subsequently, it presents the theoretical concepts, and outlines the methods and data. Results are presented through an ideal case analysis (Stapley et al., 2022), beginning with an overview of the local school context and then illustrating school actor's enactment strategies. The conclusion and discussion section includes reflections on the similarities and differences between schools, policy recommendations and potential future lines of research.

## **The Italian context and local meaning of educational reforms**

The Italian educational system has been long characterized by a centralized and bureaucratic structure (Grimaldi & Serpieri, 2012; Mattei, 2012), and traditionally recognized as a late adopter of educational reforms focused on evaluation, accountability and innovation (Kickert, 2007; Barzanò & Grimaldi, 2012). Starting from the late 1990s, a sequence of performance-based reforms has been initiated to modernize the educational system (Barzanò & Grimaldi, 2014; Checchi & Mattei, 2021).

### **Assessment and accountability policies**

In 2013, a National Evaluation system (SNV), was introduced, encompassing different policy instruments and integrating both external (national assessments and a sample-based ministerial inspection) as well as internal performance-based accountability mechanisms (through school's self-evaluation reports and improvement plans) with a specific focus on school improvement and innovation (Presidential Decree 80/2013). In this comprehensive evaluation system, there are two distinct but complementary national organizations which are responsible for assessment and accountability on the one side (INVALSI), and improvement and innovation on the other (INDIRE). The system has been therefore defined as a "three-legged model" where innovation is one of its main components (Mentini & Levatino, 2024).

The PBA process involves schools' compiling of a self-evaluation report (RAV), which includes a self-analysis of the schools' resources, the organizational and pedagogical processes adopted and

the results achieved, including external INVALSI assessments<sup>13</sup>. This is followed by external inspections that visit the schools to assess the reliability of the report and the stated goals, the development of improvement plans, indicating how the school is expected to reach the goals by defining strategic actions and priorities, and the publication of the documents in school websites or ministerial portals for transparency purposes. The national student assessment (INVALSI), as one component of the accountability system, evaluates students' competencies in Math, Italian and English language at primary, lower-secondary and secondary education. These national tests aim to measure students' "fundamental competences, knowledge and abilities" and are expected to lead to a change in pedagogy, because they are not tied to the evaluation of "simple knowledge" acquisition, but rather to students' problem-solving skills, their ability to apply knowledge, connect it to other domains and reasoning capacities (INVALSI, 2023)<sup>14</sup>.

The PBA model therefore emphasizes diagnostic, equity and innovation logics (Mentini & Levatino, 2024), since it puts emphasis on changing professional practices and the ability of schools to adapt to their specific contexts (Paletta et al., 2020). The system serves to diagnose critical teaching and learning areas, the identification of which should be followed by interventions in schools, aimed at systematizing practices and reducing learning inequities (Mentini & Levatino, 2024). The apparatus is therefore expected to lead to a process of continuous reflection and improvement in organizational and pedagogical aspects (Faggioli & Mori, 2018), with no material consequences attached to the performance results (Paletta et al., 2020). On the contrary, the external standardized tests, fostering student reasoning and transversal competences, are often considered an innovative tool, in an educational context, such as the Italian one, characterized by a teacher-centered theoretical legacy (Mentini & Levatino, 2024). In fact, according to the official discourse, to prepare students for the tests "a slightly different, more engaging pedagogy may be needed, which stimulates students to think about what they are studying and make it theirs" (INVALSI, 2023).

---

<sup>13</sup> The results include both internal student assessment results as well as results achieved in INVALSI national standardized tests, reporting the level of learning achieved in Italian, mathematics and English tests, in relation to schools with similar socio-economic and cultural backgrounds.

<sup>14</sup> In literacy, INVALSI tests measure the ability to comprehend a text; in mathematics, the ability to solve a problem applied to real-life problems and the ability to argue; in English they are tied to listening and reading exercises based on real-life situations, and thus are less focused on the acquisition of knowledge, as for instance the grammar.



## The Italian policy approach to innovation

Contemporary reforms in Italy also showcase increased interest in pedagogical and organizational innovation. Decentralization reforms (law 59/1999 and 107/2015), gave power to lower government levels and expanded the autonomy of schools especially in teaching, curricular and pedagogical matters. The innovation mandate in Italy is based on schools' autonomy in defining of a three-year educational plan (PTOF) where teachers are expected to choose the most effective teaching strategies<sup>15</sup> and plan their curriculum based on the European life-long learning skills, and a teacher evaluation based on innovative teaching methods and student performance (Checchi & Mattei, 2021).

The 107/2015 law ("The Good School Reform") initiated spaces for pedagogical innovation and teaching flexibility. Education innovation became associated with pedagogical changes and innovative teaching methodologies such as cooperative learning, competence-based education (lifelong learning skills), project-based and hands-on workshops, in the direction of overcoming a frontal teaching model (Biondi et al., 2009). At the same time, schools are granted higher autonomy to create partnerships with external stakeholders, and this type of network-based innovation should also lead to higher pedagogical innovation and digital innovation (law 107/2015).

In Italy, the innovation mandate is therefore composed of multiple and diverse pedagogical and organizational features, lacking conceptual clarity and consensus (Mentini & Levatino, 2024). In general, the concept of innovation in Italy is composed of four main dimensions: 1) *Teaching and learning*, which is related to transformations in teaching and evaluation methods (e.g., adopting problem-based teaching, debate, flipped classrooms methods, continuous formative assessment methods), flexible learning spaces and timetable, and in the choice of (digital) educational materials, 2) *Leadership and organizational development*, related to a form of distributed leadership, collaboration between leadership and staff, and amongst staff, modernization of infrastructure and flexibilization of school calendar 3) *Openness and relationship with the external community*, as in schools being open up to the local community (including universities, industry and cultural institutions), collaborating and

---

<sup>15</sup> Teacher autonomy is disciplined through the 59/1999 law, which guarantees teachers with a high degree of flexibility in the timing, modalities and activities they consider more appropriate with respect to each student's individual learning. It establishes that teaching practices are based on the school's pedagogical plan (POF, Piano d'Offerta Formativa), elaborated by each individual school and encompassing diverse methodological options (art 4 law 59/1999).

involving families and collaborating with external partners (such as agencies, institutions, local entities), and 4) *propensity to change*, related to strategic planning, including the use of data for school improvement purposes, technological devices, professional development and curricular change (Nardi et al., 2022). ICT based, and e-learning activities are also part of the innovation mandate (National Digital School plan, 2015).

### School actors' reception of the double-mandate

In Italian schools, teacher and school evaluations were never seen positively by teachers and unions given the intrusion felt in the professional autonomous field, leading to teacher resistance through forms of protests, strikes and oppositions (Barzanò & Grimaldi, 2013). According to previous research, the great majority of teachers was found critical and contrary to the INVALSI assessments (Martini & Papini, 2015), while over half of the Italian teachers expressed preference for internal school self-evaluations (De Angelis et al., 2015). Arguments against the implementation of external tests are related to having introduced competitive dynamics amongst schools (Peruzzo et al., 2022), the perception of being disconnected from school subjects and practices, having generated anxiety and doubts, or perceived as a top-down and bureaucratic requirement (Pastori & Pagani, 2016). The perplexities are also related to the aim and purpose of the evaluations, their supposed “objectivity”, and the change produced in the teacher’s autonomy and instruction (Di Cresce, 2019).

In terms of innovation, although data shows that less than half of the schools implement a transversal competence-based curriculum (Poliandri, 2015), and that Italian teachers were tied to low innovative practices, such as frontal (face-to-face) teaching and the use of “less constructivist strategies”, including project-based, cooperative learning or inquiry-based learning (De Sanctis, 2010; OECD, 2013), Italy is also the country where schools are considered having strong capacity to adopt innovative practices (OECD, 2019), and is renowned for a vibrant teacher union activism and pioneering pedagogical experiences and figures, including Maria Montessori or Reggio Emilia experience.

However, the purportedly “equilibrated” national evaluation model (SNV) integrating both accountability and innovation components, appears limited in practice (see Mentini & Levatino, 2024). Key barriers include incomplete school autonomy, resource deficiencies, a standardized

teaching culture (OECD, 2013) and a supposed misunderstanding and misuse of the policy instruments among school actors (Mentini & Levatino, 2024). Furthermore, given that it is easier to objectify learning outcomes than other educational processes, there is the risk that schools prioritize test results instead of other processes when compiling their self-evaluations, potentially rendering the system as “crippled”, as the external accountability overshadows other components (Mentini & Levatino, 2024). Consequently, analyzing how school actors interpret, receive and enact the institutional messages in their daily practices becomes a crucial avenue for research in this context.

## **Theoretical framework**

To explain school’s policy enactment processes, the theoretical framework integrates interrelated mechanisms of sense-making (Ball et al., 2012; Coburn, 2001) and institutional filtering (Diehl & Golann's, 2023; Dulude & Milley, 2021), with a heuristic approach to the enabling/constraining role of the local environment (Braun et al., 2011). Rather than referring to policy compliance or implementation, the theoretical framework considers the range of responses that schools give to external pressure, understanding the factors that affect those responses and, in turn, shape external pressure (Elmore, 2005a).

### **Making sense of the policy messages by school actors**

School actors experience and interpret policy prerogatives in different ways actively shaping and transforming the resulting policy outcomes (Ball et al., 2012; Braun et al., 2011). The cognitive process of interpretation involves perceptions, opinions, and attitudes towards educational policies. When faced with a new policy mandate, school actors engage in interpreting and trying to make sense of the message, before putting the new mandate into practice (Coburn, 2001). Teachers and school actors actively mediate institutional pressures, a process framed by their pre-existing beliefs, worldviews and practices (Coburn, 2004). This implies that educators adapt, assimilate, and selectively enact policies and initiatives in alignment with other logics they have been socialized into (Thornton et al., 2012). The process includes: a) noticing and selecting information and messages to put into practice b) constructing meaning and understanding of the information and c) negotiating details and acting upon those interpretations (Coburn, 2001; Porac

et al. 1989). The process of meaning-making also includes how individuals think and feel. Thus, teachers' emotions may influence the sense-making and reaction to the accountability reform (Kelchtermans, 2005). For instance, stress and anxiety related to external control and standardization (Shoen & Fusarelli, 2008) and a greater attention to students' pace and in-depth curriculum (slow-teaching) may signify a risk for teachers, whilst standards, structure and direction may represent a desirable safe path for teachers, who are often overburdened (Falabella, 2020).

### Institutional filtering into local realities

New theoretical approaches integrate policy sociology and enactment research into institutional theory to comprehend how schools navigate different institutional pressures (Diehl & Golann's, 2023; Dulude & Milley, 2021). According to these perspectives, schools filter and locally assimilate the external pressures into their organizational reality and educational activities. Filtering encompasses various environmental aspects entering school organizations, while adaptation pertains to the local incorporation of these filtered aspects into daily operations by school actors. Crucial elements such as routines, networks, and sensemaking processes contribute to understanding how and why schools filter external demands, adapting them into their organizations (Diehl & Golann, 2023). Key actors, including principals or frontline workers such as teachers, actively frame external messages, noticing, selecting, and filtering policies in their organization in different ways (Diehl & Golann, 2023). Consequently, when confronted with external demands, especially when these cover multiple priorities and may even be perceived as contradictory, schools may exhibit a broader spectrum of interpretations and enactments to policy pressures, extending beyond resistance or alignment. As exemplified in the findings section of this article, the interplay between sense-making and filtering mechanisms plays a crucial explanatory role in understanding diverse school responses of schools to external institutional demands (Parcerisa & Verger, 2023).

### The mediating role of school's contexts

Previous research indicates that schools exhibit diverse responses to the policy demands, emphasizing the importance of considering school contexts, agencies and their unique characteristics (Landri, 2021; Coburn, 2004). The translation of policies into practice is a negotiated

and intricate process often influenced by external circumstances encompassing institutional and socio-economic factors (Ball, 1994; Braun et al., 2011). The distinct attributes of schools and other contextual factors can thus elucidate and mediate school actors' interpretations and behaviors in relation to institutional messages (Braun et al., 2011).

The socio-economic composition of the school, intricately linked to its intake, geographical site, location and historical background, contributes to the construction of narratives and institutional stories about the school (Braun et al., 2011). Research frequently observes that to respond to accountability pressures schools under probation are more inclined to adopt instructional strategies such as tailoring teaching to align with test content or narrowing the curriculum to focus on assessed subjects. This strategic approach aims to enhance test results, allowing schools to navigate punitive measures without fundamentally altering their teaching methods, educational principles, or organizational procedures (Mittleman & Jennings, 2018). This maneuvering enables schools to maintain a status quo or continue their activities without disruption (Hallett, 2010). Moreover, the adoption of new organizational routines and responses to accountability demands may vary depending on whether schools are under probation, influenced by how school actors experience external pressure, and their adherence to the accountability policies (Verger et al., 2021).

A school's socio-economic context also influences the teaching approaches and pedagogies. Teachers are found to use test data to categorize underperforming students and tailor their instruction accordingly (Hardy et al., 2019). Additionally, teachers tend to prioritize practical-oriented instruction (e.g., lectures, seat work, memorization, recitation) in lower-performing schools, preparing students for manual, clerical, or low-wage service-sector, while providing more challenging instruction to students from socially advantaged groups (Diamond, 2007). This highlights the intricate relationship that occurs between socio-economic inequalities, teachers' expectations with their students' potential, and the adoption of pedagogic practices<sup>16</sup>.

Material contexts encompassing school technology, resources and budget also play a key role, since the conditions of the school buildings, including their layout, quality, and spaciousness, have a considerable impact on policy enactments on the ground, including the capacity to attract and keep 'good' teachers or other staff (Braun et al., 2011).

---

<sup>16</sup> Working class students typically acquire more context-dependent (restricted codes) of meaning and knowledge compared to privileged students (Hoadley, 2006) and teachers tend to hold lower expectations regarding the future careers and learning achievements of students from economically disadvantaged backgrounds (Tarabini, 2012).

Professional contexts, referred to as the values, ethos and culture of a school, as well as the school management and teacher experiences, ultimately shape how schools prioritize accountability demands compared to other societal expectations (Braun et al., 2011). Individuals adapt to new policies based on their interactions with others and the organizational cultural configuration. Existing networks of collaboration and support within schools, especially among teachers, significantly contribute to how schools translate these demands into specific educational practices and organizational routines (Diehl & Golann, 2023). This responsiveness stems from an openness to dialogue among colleagues and a teacher's self-evaluation embedded in the day-to-day practice, mitigating anxiety and pressure (Elmore, 2005b). Additionally, to align internally with the external accountability demands, schools necessitate a shared culture with strong professional norms, a well-developed capacity for teachers to collaborate around shared values, and a “distributed leadership” involving teachers more actively (Firestone, 2004). The principal's actions and leadership style may therefore be crucial in aligning and maintaining congruence between external and internal factors (Ball & Maroy, 2009).

In conclusion, the singularities of schools, manifested in concrete conditions and “professional cultures”, act as ‘frames’ and ‘filters’ to reform ideas, aiding in understanding the likely configuration of the translation (Landri, 2021).

## **Methods**

The methodology employs a case study approach (Yin, 2009), based on qualitative research conducted in 12 lower-secondary schools in Rome, Italy. The selection of Rome is driven by both analytical considerations, such as feasibility of the study and familiarity with the context, and substantial factors, including ongoing accountability and innovation reforms in schools, and the socioeconomic disparities within the city that are taken into account when selecting schools. Rome serves as a compelling and representative case due to the presence of PBA and innovation reforms, along with a policy arena hosting key educational stakeholders, public authorities and institutions. The city encompasses 200 sub-urban zone areas (equivalent to city neighborhoods) and 15 municipalities, displaying significant variations in social and economic indicators across these areas

(Lelo et al., 2019), proving a fertile ground for exploring differences amongst school enactment responses.

Public state schools exclusively form the focus of this study, given they constitute the majority (73%) of schools in the city of Rome as of 2018/2019 (Ministry of Education, 2023)<sup>17</sup>. In Italy, the state is the main provider of mass schooling, with only limited space for private (religious) schooling.

The study concentrates on lower-secondary education (ISCED 2, age 11-14)<sup>18</sup> commonly known as “middle school”, typically included in comprehensive school institutes covering ages 4-13<sup>19</sup>. The decision to focus on this educational level stems from the highly selective nature of the Italian system after lower-secondary schooling (Giancola & Salmieri, 2022), operating along three very hierarchical tracks: technical, vocational, lyceums (high schools). In the final year of low-secondary school, students undergo the INVALSI national test and a final oral and written examination, before transitioning to secondary education, obtaining a “diploma di licenza media”. Furthermore, the choice to emphasize lower secondary education is grounded in the belief that the relationship between accountability and innovation becomes apparent at this level. Primary schools typically have one teacher per class, whereas in lower secondary students exhibit more autonomy in their learning, and the teacher-student relationship is less mediated by families.

## Sampling procedure

The research employed a two-step sampling process. Initially, a randomized sample of schools was chosen. In this phase, 100 public lower-secondary schools in the city (out of 370 total lower-secondary public schools) were selected based on ownership, educational level and further stratified by average income and performance on INVALSI national tests. Subsequently, utilizing the representative sample, schools were selected according to two primary variables: performance

---

<sup>17</sup> <https://dati.istruzione.it/.opendata/opendata/catalogo/elements1/?area=Scuole11>.

<sup>18</sup> There are a total of 283 low-secondary schools in Rome (both public and private) amongst which 72% are public and 27% are private schools (dati.istruzione.it).

<sup>19</sup> Comprehensive institutes are school buildings (*plessi scolastici*) which include all levels of education, with the aim to give educational continuity to students (age 3–14), and are usually located in the same neighbourhood and in proximity to each other. They are also run by the same principal and collegial body.

in standardized tests and the school average income, acting as a proxy of school's socio-economic status<sup>20</sup>. The data on school average income and performance results were divided into quartiles (low performance (Q1), middle performance (Q2 + Q3) and high-performance (Q4); working class (Q1), middle class (Q2 + Q3); and upper class (Q4). By crossing these two variables, a total of 12 schools were selected across four categories. The characteristics of these schools are presented in table 4 below.

---

<sup>20</sup> The average income by neighbourhood (urban area) where the school is located is retrieved from the Ministry of Economics and Finances public data, based on Revenue Agency Data (2019).



Table 4. Characteristics of selected schools, by income and performance (Source: own elaboration)

<i>School</i>	<i>Average Income</i>	<i>Performance category</i>	<i>Interviewed school leaders</i>	<i>Interviewed teachers (also with technical functions)</i>
1	Upper class (Q4)	High Performing (Q4)	2	1
2	Upper class (Q4)	High Performing (Q4)	1	1
3	Upper class (Q4)	High Performing (Q4)		1
4	Upper class (Q4)	High Performing (Q4)		1
5	Upper class (Q4)	Low Performing (Q1)	2	1
6	Medium class (Q3)	Medium Performing (Q2)	2	4
7	Medium class (Q2)	Medium Performing (Q3)		1
8	Medium class (Q2)	Low-performing (Q1)		2
9	Medium class (Q2)	Low Performing (Q1)		2
10	Working class (Q1)	Low Performing (Q1)		2
11	Working class (Q1)	Low Performing (Q1)		1
12	Working class (Q1)	Low Performing (Q1)		3
<b>Total</b>			<b>7</b>	<b>20</b>

## Data collection and participant selection

The study employed a case study approach, drawing on semi-structured interviews to principals, management team and teachers. Additionally, where feasible, documentary analysis was conducted using publicly available documents such as the school's institutional project (PTOF), self-assessment reports (RAV) and fieldnotes.

A purposive sampling strategy was adopted (Patton, 2015). Principals and management teams were interviewed due to their central primary role in the organizational and managerial aspects, especially given recent reforms on school autonomy (law 1997; law 107/2015). For teachers, selection criteria included those responsible for a course or subject assessed in the national test. In some cases, interviews were extended to teachers not directly involved in assessed subjects. Teachers with technical and/or informal leadership positions (i.e., part of evaluation committees, project planning committees, collaborators of the school principal, instrumental functions) were included given their relevant managerial and coordination tasks<sup>21</sup>.

## Access to schools

Initiating contact with schools involved sending emails and subsequent phone follow-ups, leveraging publicly available contact information on school websites. Clear communication of the project aims and interview details were clearly given to the principal, with personal visits in some instances to establish trust and rapport. While some schools, particularly middle-low and working-class areas, presented challenges for interviews (due to high rotation, other school priorities or unresponsiveness), document analysis served as a valuable complement to triangulate responses.

---

<sup>21</sup> From the 107/2015 school autonomy law, a more complex school organization emerges. Some teachers are appointed as referees for some particular aspects and take part in evaluation/technical committees that are responsible for different bureaucratic aspects, such as the formal filling in of documentations required by the accountability system (Self-assessment report (RAV), and improvement plan (known as 'Piano di Miglioramento', PdM).

However, I still considered and took this limitation into account in the analysis and interpretation of results.

## Interview structure and data analysis

The interviews were conducted between November 2022 and January 2023, lasting an average of 40 and 60 minutes. Employing a semi-structured format, interviews followed specific scripts but remained flexible to delve deeper into emerging issues. Interview scripts were tailored for principals/leadership team covering modules that sought to investigate the relationship between school contingencies and enactment processes. Scripts addressed the opinion and translation into practices of the policy expectations on schools, and at the same time capturing schools' contextual characteristics. The modules included: 1) Participant background, the years of experience working within the school and responsibilities 2) School context and resources - socio-economic composition of students, material and economic resources, school reputation, school climate - to investigate the cultural, social, historical and material elements that condition policy enactment 3) Opinions and perception about the accountability system - opinion about national tests and their publications, use and discussion of test results for improvement and innovation; perceived pressure – to capture opinions on the standardized test, the alignment of school goals with the accountability system, and perceptions about accountability pressures 4) Translation of the accountability system to pedagogical and organizational strategies and innovation practices - exploring the school's and teachers' educational approach, pedagogical practices, and innovation strategies 5) Market accountability and family relationship - to explore aspects that characterize the school's reputation, marketing strategies, and relationships with parents with other schools.

For data analysis, an ideal case analysis was employed following a seven steps methodology (Stapley et al., 2022). Ideal types can be defined as “generalizations or mental representations of a social phenomenon that will never be identical with reality, but which will help to make that reality understandable” (Stapley et al., 2022, p.2). The ideal case analysis is based on the process of developing a typology by grouping cases or participants into different types on the basis of their common features (Stapley et al., 2022). First, I familiarized with the data through interviews, transcription and coding. Transcriptions were coded using ATLAS.ti (version 22) software, employing a flexible coding strategy (Deterding & Waters, 2021), which integrates existing codes from the ERC-funded project (see Parcerisa and Verger, 2023) - in the context of which this study

is conducted - and emerging inductive codes. Case reconstructions were then created, chronologically summarizing each participants' narrative through approximately half a page per participant. By comparing and contrasting the school cases, different categories of enactment responses were developed, based on the school's average income and student body population. Then, descriptions and names were assigned to these types and refined accordingly, followed by a detailed exploration of similarities and differences within and between school actors' narratives.

## **Results**

### **(1) Privileged school settings**

Privileged schools, characterized by an upper-class student body, are prestigious and in high demand. These schools face limited competition, boasting oversubscribed intakes and a favorable reputation, reinforced through positive word-of-mouth. The families choosing these institutions typically possess a “a rich cultural background” and are financially invested in the educational offer (School 4, institutional project 2022-2025). Academic excellence takes precedence in privileged schools, reflecting heightened parental expectations for their children’s “brilliant futures”. Despite endorsing a traditional teaching approach, there is demand for project-based activities and extracurricular courses (Principal, School 4). The stability of teaching staff is also considered essential by parents in school choice and selection.

The reputation of privileged schools is based on student performance, emphasizing academic success. A wide offer of extracurricular projects that align with global citizenship themes, including environmental sustainability, wellbeing, and extra-curricular afternoon courses are offered such as music or sports. Innovation involves institutional and organizational strategies such as establishing partnerships with local entities (e.g., non-profit or religious organizations, music schools), participating in school networks (i.e., plastic free schools) and involving families in financial contributions for extra courses.

The professional contexts of these schools also play a crucial role in shaping the school's capacity for change and innovation. They have clear leadership direction and cohesive staff (e.g., School 4),

demonstrate a more positive attitude towards change and innovation, and are able to activate additional training courses for teachers and interdisciplinary projects. Staff members enjoy stable positions that foster an environment where collaborative learning and project support thrives, since working together for many years" represents "a stimulus to learn from each other and support each other's projects" (Teacher, School 4). Administrative staff in these schools also play a crucial role in supporting managerial and financial responsibilities, such as fund-seeking and project management. Despite sharing common characteristics, privileged schools exhibit varied approaches to accountability and innovation mandates, influenced by their performance level. This diversity underscores the nuanced ways in which schools navigate and respond to external demands.

### *High performance results and embracing innovation without tensions*

In privileged contexts with high-performance results, school actors navigate the dual accountability-innovation mandate by favoring innovation over external accountability, with minimal perceived tension. The data obtained from the INVALSI tests functions as a monitoring instrument, but its systematic use for identifying performance gaps or planning educational improvements remains elusive, creating a gap between symbolic and substantive changes. The enactment of external accountability therefore resembles a formal bureaucratic ritual (Landri, 2021) rather than a catalyst for substantial change in teaching practices. Despite setting goals based on assessment data, the absence of significant positive or negative consequences attached to performance results diminishes the perceived pressure, rendering discussions more formal than transformative.

"Maybe ideally, we think it's right to take this into account, but then in practice, we don't change much of teachers' pedagogy based on INVALSI data" (Teacher/Evaluation committee, school 4).

In schools with highly positive INVALSI results, the absence of consequential pressure contributes to a lack of systematic reflection on the data's utility for identifying learning gaps or planning improvements. The positive results create a "happy island" (Principal, school 2) where performance is already high compared to other schools, and the need for introspection is perceived as unnecessary.

“Students have excellent skills, so we don't even think 'let me go and see where I went wrong' or where the critical point is” (Principal, School 2).

However, while in school 4 the external test data is not perceived as important, nor by teachers or the school principal, and disregarded as a valid evaluation instrument because associated with a multiple-choice test, in school 1 and 2, the INVALSI test is perceived as more important by the school leadership team. This is because, in these contexts, it is associated with a “democratic evaluation method, which is the same for everyone in Italy” (principal, school 2) or “a comparison parameter at national, regional and school level”, thus a variable that highlights internal class differences and which guides schools’ planning (principal, school 1).

Moreover, the innovation mandate in privileged schools is associated with various and different pedagogical and organizational features. Innovation in privileged contexts encompasses both educational processes such as the adoption of competence-based teaching and digitalization (school 1), as well as institutional and organizational innovations such as open schooling approaches (school 1 and 4), doing things that were never done before (school 2) and a willingness to adapt to societal changes (school 1). Beyond such features, while in school 4 and 1, innovation is particularly linked to school’s capacity to establish relationships with external entities and local associations, enriching the educational offer with projects, and is intended as a pre-established practice adopted way before other schools, school 2 intends innovation as a “organisational revolution”, exemplified by the adoption of a model called “DADA model” (Didattiche per Ambienti Di Apprendimento), which implies restructuring classroom settings for different subjects, and requiring students to move autonomously from one classroom to another, thus bringing substantial changes in school spaces, timetable and teachers’ professional development. Such inclinations and initiatives reflect a clear administrative organizational school innovation (Lubienski, 2009) and are sustained by the school's pedagogical autonomy in diversifying the educational offer.

“Innovation also means doing new things, which perhaps have never been done before, or which were a bit afraid to try” (Principal, school 2).

In terms of teaching practices, privileged schools emphasize lifelong-learning, citizenship competences (Principal, school 2), and competencies for the future, such as autonomy of learning and responsibility. Active and interdisciplinary learning take precedence over content-based

teaching (Principal, school 1). Student-centered education, focusing on personalized teaching methodologies, is favored to allow students to express their individual characteristics (Principal, school 1). Moreover, in such schools, diversifying, adapting and personalizing the teaching approaches and overcoming a traditional education model is seen as beneficial to enhance students' learning, inclusion and strengthen key competences. However, the role of teachers and teacher training in sustaining innovative pedagogical approaches changes according to the schools. In school 4, teachers are viewed as having a clear attention to pedagogical innovation and they do not adopt an "old school" frontal teaching method (Teacher, evaluation committee, school 4). Teachers are continuously self-developing and are seen as proactive in suggesting new project-based activities, that is "a strength and is also a continuous stimulus for all school staff (...) and the reason why we have such a wide educational offer (...)" (Teacher, evaluation committee, school 4). On the contrary, in school 1, the leadership team believes that some of the teachers are still tied to using "obsolete teaching methods" associated with a passive transmissive educational model, linked to content-based teaching and student's performance (principal, school 1). When this happens, teachers are given specific attention and the methodologies will be "aligned and adapted to that of the school" through specific training courses (Vice-principal, school 1).

Moreover, PBA results are perceived as a reflection of innovative pedagogical and educational practices already in place. The conflict between the PBA mandate and innovation only arises when the former contradicts the evaluation of "more significant transversal and citizenship competences" or continuous formative assessment (Vice-principal, school 1). The PBA mandate in these schools is hence adapted to meet the requirements of improvement and innovation.

"INVALSI tests evaluate only basic skills, therefore do not reflect schools' pedagogical priorities" (Teacher, School 2).

Following this reasoning, in such schools, as opposed to the other categories of schools, teachers do not report using instrumental strategies, such as teaching to the test, to enhance students' performance. This specific practice is either completely disregarded and omitted by the interviewees, or considered to be in clear opposition to pedagogical values related to evaluation, which emphasize "open ended questions, reasoning, reflection and understanding as opposed to the utilization of true or false questions to test students' preparation" (Teacher, school 2).

Parental preferences also contribute to prioritizing the innovation mandate in these schools. The absence of pressure on INVALSI test results by parents underscores their major interest in other performative aspects, such as languages, internationalization (School 1), innovative or well-being educational dimensions (School 2). On the contrary, parents who subscribe children in these schools “do not look at INVALSI results” (school 1) or “are actually critical to students taking the tests (...) and do not even understand why students should take them” (Teacher, Evaluation committee, school 4). Only in one school case among such privileged schools, affluent parents also support the achievement of traditional academic goals in addition to innovation and hence are also “worried about INVALSI results” (Principal, school 2), and academic “basics” which are commonly seen as solid, tried-and-true educational practices (Kohn, 1998).

More often in privileged schools, qualitative and reputational instruments, including questionnaires to students and families, play a more prominent role in gauging parental satisfaction (Principal, school2) and innovative practices serve as a marketing strategy (Lubienksi, 2009), attracting families and making schools more efficient and appealing.

“These (innovative) methods have increased the family’s choice of our school... adopted to make the school more efficient but also more attractive” (Principal, School 2).

This interesting result is contrary to the often-assumed hypothesis by which schools with more reputation face lower levels of external pressure due to competitive dynamics and that schools with higher SES student bodies do not have to divert resources for marketing campaigns since they tend to be oversubscribed (Lubienski, 2009). This is attributed to a specific governance model in Italy, associated with a quasi-market model (Benadusi & Consoli, 2004), based on family’s free choice, that creates a sort of demand-driven competition between schools (Colombo & Desideri, 2018)<sup>22</sup>. In addition, organizational types of innovations adopted in privileged schools are typically associated with quasi-market models (Lubienski, 2009). Indeed, in such privileged schools’, although intakes are already high and competition with other schools is not particularly felt, principals feel the need to maintain the positive reputation and believe that diversification<sup>23</sup>

---

<sup>22</sup> Whereas before the choice of a school was only possible according to the same geographical area, since the 1990s the choice is possible in relation to each school, public and private (Pandolfini, 2009).

<sup>23</sup> Diversification (or differentiation) is defined as ‘an increase in the number of options available locally in schooling, usually from the parents’ perspective’. Innovation and diversification are distinct but inextricably related (Lubienski, 2009, p. 19).



(Lubienski, 2009) and a unique pedagogical offer is what attracts parents to the school, making the school being specifically “chosen in comparison with other schools in the same neighborhood, city or out of Rome (Teacher, Evaluation committee, school 4).

“The families are interested to what we offer (...), during the open days they start asking about robotics or a specific language course or subject, so we have a certain educational offer, and all of it, including the three-year educational plan is visible on the school website” (Teacher, Evaluation committee, school 4).

“I present my offer, and if a parent likes it, they come to our school, if not they go to another one (..) because other schools do different things compared to us, but that's okay, otherwise we would all be photocopies, and then the parent chooses rightly” (Principal, school 2).

Another parameter of school choice is students' appreciation of the school. Thus, keeping the school attractive to students also brings higher intake rates: “the innovative school method that we adopt, keeps students happy and makes them self-responsible, and this is highly appreciated by parents” (Principal, school 2). Privileged schools thus leverage their autonomy to diversify and adapt their educational offer in order to meet parents' expectations, thus maintaining high subscriptions. They also leverage their high reputation and high student body to attract stable teachers. Such aspects, together with positive academic results, resemble family's motivations for school choice and preferences, and do not require additional marketing efforts to attract students. These rationales altogether resemble a “universal” and “a-contextual” (Lubienski, 2009: 17) quasi-market mechanism and economist view of education, by which parents have the right to choose the best education for their children and by which choice and competition offer opportunities of innovation and responsiveness to consumers' preferences (Lubienski, 2009).

## (2) Middle - income school settings

Middle-income schools are characterized by an average middle income, but exhibit varying performance results and socio-economic student population. Two schools (school 5 and school 6) cater to a homogenous middle-class student population, where parents typically prefer traditional teaching methods, summative assessments and extensive use of textbooks (Teacher,

school 6). The teaching method remains traditional, emphasizing students' academic success and results, while also incorporating inclusion strategies. In addition, extracurricular subjects and transversal projects related to environmental education are also implemented. Other schools in this category (School 7, 8 and 9) have a more mixed and heterogenous student population. Such schools experienced an influx of students with immigrant backgrounds, reaching up to 50% in some cases (e.g., School 9), contributing to a multiethnic and multicultural student population with a significant percentage of students from a socially disadvantaged background (e.g. School 8).

Heterogenous schools within this category face considerable challenges, including school disorganization, fragmentation, high teacher rotation, and principal turn-over. Teacher stability is compromised due to precarious contracts and budget cuts, leading to a limited personnel pool. The instability of teachers, “who stay in schools only for only a few years before moving to other schools” (Teacher, school 8), makes long term planning and the implementation of long term-changes and innovations more difficult. In these contexts, there is an imperative for schools to proactively engage in self-activation and adapt to frequent personnel changes.

“If there is a cohesive teaching staff with stable contracts, you can work during the years, but if there is continuous change in personnel, it is more difficult to implement an educational strategy” (Teacher, School 8).

### *Homogenous contexts: Balancing Innovation and Performance amid internal conflict*

Within homogenous contexts, characterized by average performance results, school actors tend to align innovation with performance goals, utilizing the external assessment data to drive both innovation and performance processes. However, this alignment gives rise to conflicting opinions amongst school staff.

At the managerial level, a significant emphasis is placed on the accountability system. Principals and management teams view accountability instruments as an objective and crucial measurement of schools' and teachers' quality. They perceive performance-based tools as essential for self-diagnosis, reflection and influencing teaching strategies. INVALSI assessments, seen as an indicator of teaching quality, are employed to understand the skills effectively taught to students.

“Through INVALSI results we have the possibility of understanding on which skills the class has worked on well or on which ones it has not” (Principal collaborator, school 6).

Data from the PBA also serves as a basis for identifying improvement actions and planning. Such schools analyze learning objectives and gaps from INVALSI results, to strategically allocate efforts in the upcoming years, by supporting improvement processes within the school and establishing future priorities in the educational project. The PBA mandate, in this scenario, is embraced, since schools display agreement with the requirements and expectations of the policy message (Landri, 2021).

“What interests us is the improvement beyond the standardized INVALSI tests (...) The improvement should be both at the level of the curricular and methodological skills of the teacher and at the level of the student, therefore the whole school should improve” (Teacher, evaluation committee, school 6).

When INVALSI results fall below expectations<sup>24</sup>, schools take specific improvement actions based on the data such as adjusting the class compositions, providing psychological support, and offering targeted training for teachers with lower performance results. Moreover, extracurricular courses are activated in the three subjects evaluated in the national test and teaching is modulated to the content of the assessments, by strengthening some tested subjects or modifying teaching priorities on the basis of the test results.

“We modify our teaching also on the basis of what are the answers given by the INVALSI tests” (Principal collaborator, school 6).

“There is greater attention to subjects evaluated in the external tests because, for example, remedial courses are activated only and exclusively for the subjects covered by INVALSI tests” (Teacher, school 6).

While some teachers within these schools positively view national tests as objective measurements of student learning, positively evaluating competences which students will need throughout their whole life, others harbor negative perceptions, considering the evaluations as judgments of their

---

<sup>24</sup> The INVALSI results are given back to schools highlighting the difference in average compared to schools with similar economic, Social and Cultural Status (ESCS).

profession or interpreting them as individual assessments detached from their own teaching and educational planning. Some teachers feel distant from the accountability system, believing it does not influence their teaching nor provides opportunities to improve students' performance<sup>25</sup>.

Negative opinions on the external accountability are also reported in the high-income school with a homogenous student population, experiencing low performance results, and high performance-based pressure. Teachers feel judged in their professionalism (Principal, school 5), and the reputational pressure deriving from external test scrutiny is perceived by teachers as a self-judgment, contributing to stress and anxiety (Principal, school 5). This performative pressure and “fear” of being judged from the external assessments negatively influences teachers' instruction, limiting creativity, personalized instruction and alternative teaching practices (Teacher, school 5). The tension between accountability and innovation in this context becomes evident, hindering teachers' engagement with alternative practices or creativity.

“The pressure you feel spills over to the manager, but also to the teacher. Without pressure you would probably have even more creativity in doing things” (Teacher in evaluation committee, school 5).

In homogenous school contexts teachers also report using intensive teaching to the test strategies, to “accelerate and motivate students with additional testing especially during that time of the year” (teacher, school 5), or as a way to train students on the testing modalities, regardless of the testing content (school 6). Such practice, resonating with an item-teaching form of testing preparation (Popham, 2001)<sup>26</sup> is also sustained by making students exercise often on school textbooks, which already incorporate exercises dedicated to the INVALSI tests.

---

<sup>25</sup> One of the reasons is related to the design of the tests, since they are given back too late to schools according to teachers (usually at the beginning of the following academic year), so teachers believe they cannot not work on improving their results with the same students. Another reason is related to the fact they are anonymous evaluations and results are not given back to teachers individually for each student. Finally, since 2018 INVALSI results are not high-stakes for students, hence both students and teachers often do not attribute particular importance to them.

<sup>26</sup> Item-teaching is when teaching is focused directly on test items or on items much like them, while curriculum-teaching requires teachers to direct their instruction towards a specific body of content knowledge or a specific set of cognitive skills represented by a given test, while (Popham, 2001).

In terms of innovation, a combination of both product<sup>27</sup> and organizational innovations (Lubienski, 2009), can be observed in such contexts, such as the adoption of digital innovations (e-learning, digital technologies, coding), upgraded school facilities, (school 6), “a great focus on interdisciplinarity and collaborative projects” (school 5). Schools are well equipped with digital infrastructures, additional classroom spaces and laboratories because “if a teacher has what she needs to work she will work better and students will study more peacefully (Principal, school 6).

“This is a school that is trying to chase the wave of innovation. A school that has many new projects, especially on digitalization” (Principal collaborator, school 6).

Given the differing opinions among school staff on the PBA mandate, the role of the principal is important in aligning the internal culture with the institutional messages. Principals work to align the internal evaluation methods with the external assessment, and foster a shared understanding of the PBA instruments. Principals in such schools actively engage in communicating the purpose of external evaluations to teachers, emphasizing that INVALSI evaluations aim to monitor and recalibrate teaching-learning processes. By addressing the negative perceptions, the principal advocates for teachers' training and support, facilitating a collaborative approach to develop new instructional approaches and internalizing accountability expectations (Firestone, 2004).

“I'm making teachers understand that evaluations are needed to monitor and to recalibrate the teaching-learning processes” (Principal, School 6).

### *Heterogeneous contexts: Divergent approaches and malleability of the double mandate*

In socially mixed schools, with middle to low performance results, teachers exhibit opposing approaches to the double mandates. These responses depend on the alignment between accountability and the teachers' preferred pedagogical approaches. Teachers may either express tensions or integrate performance data as an innovative or improvement device. When INVALSI tests are understood in opposition to a holistic or competence-based teaching, teachers express higher tensions between the two mandates. On the contrary, when interpreted in alignment with

---

<sup>27</sup> Process innovations in education also includes *product innovations* such as new or substantially different service offered to students (Lubienski, 2009).

existing innovative pedagogical preferences, teachers integrate performance data in their teaching and use the external assessment as an innovative or improvement device.

Some teachers view external assessments as unfair and unreliable. The association of national external tests with quiz-type and standardized evaluations (Teacher, school 8) contradicts their preference for personalized assessment and teaching methodologies (Teacher, school 9), believing they reduce students' conceptual learning and critical thinking. The negative opinion results in open resistance and 'opting out' from the regime of accountability (Landri, 2021).

“I'm critical of the INVALSI because they are quizzes that students have to deal with only during a short time; they don't improve the reasoning capacities of students, they are often only quizzes” (Teacher, school 8).

Teachers openly express a desire to boycott the tests, citing fatigue, lack of awareness and the perceived futility of personalized teaching within a standardized framework (Teacher, school 9). This represents an explicit refusal of the testing logic and performative pressure underpinning the external accountability mandate.

“Let's say that I could boycott the INVALSI tests, because it is useless to talk about personalized teaching when you do things that have been the same for everyone” (Teacher, school 9).

Conversely, when teachers associate the accountability mandate with transversal or competence-based teaching, they display more positive attitudes. INVALSI tests, in this context, are seen as stimulants for reasoning capacities, beyond evaluating curriculum knowledge and content.

“INVALSI is the result of when you work in terms of skills, transversally. It's not a matter of curricular programs, it's being able to think about what you know, being able to come up with a reasoning” (Teacher, school 7).

Interestingly, in some cases, teachers undergo significant transformations in assumptions due to INVALSI assessments, not only focusing on the surface level features of the message (Coburn, 2004). They use the national tests to stimulate students' transversal competences, by integrate them in their teaching throughout the whole year in a form of curriculum-teaching (Popham, 2001) or

they use the simulations of the INVALSI tests to broaden students' divergent thinking. Indeed, teachers seldomly use instrumental strategies such as teaching to the test, and this practice is adopted for different reasons and in less intensive ways than in homogenous schools. For instance, teachers report making students exercise on INVALSI tests only at times, but with no specific intention (Teacher, school 9), or with the rationale "to activate other kind of reasoning and make students evaluate their learning in other ways than the ones they are used to" (Teacher, school 8). Interestingly, one teacher also perceives the external evaluations to be an "innovation" that stimulates problem-based teaching, leading to a departure from a more traditional teaching model.

"I must say I appreciate the work done by INVALSI in the sense that I actually see it as an innovation (...) When the abstract calculation is not required, students have to know how to use what you have in front of them (...) so you have to set up the teaching in a whole other way" (Teacher, school 9).

Teachers in heterogeneous contexts adopt socio-emotional strategies, focusing on equity and inequalities in education, for instance offering courses for students with a migration background in the Italian language. They take on an "activist" and missionary approach, defending students' interests and backgrounds, "by taking care of students and giving them as many experiences as possible" (Teacher, school 9). Innovation in heterogeneous contexts is associated with "not doing things in a traditional and passive way" (Teacher school 9), and encompasses active learning, project-based and manipulative teaching, outdoor learning, small group work, and alternative assessment methods.

"Students have a very low concentration level, so we need to get the message across in very different forms, whether it be in a visual way, perhaps through quizzes, sometimes even in the form of a game" (Teacher, School 8).

Some teachers adopt active-based methods to enhance students' interest (teacher, school 9), or manipulative methods to overcome students' learning difficulties or "absences of educational culture" (Teacher, school 7). This approach is often justified by the fact that "maybe students who have more difficulties in Italian language or reading a textbook, will activate other types of skills through manual activities" (Teacher, school 8). As it is argued in the literature, working-class students often receive instruction that is more practically oriented, and prepares them for manual, or low-wage service-sector (Diamond, 2007). In this sense, innovation in heterogenous school

contexts is associated with *process innovations* in education, which usually focus on different and innovative teaching methods and that enhance efficiency, within an improved pedagogical approach (Lubienski, 2009).

Finally, in these school contexts, the pressures on teachers to balance innovation and students' individual development, alongside the demands for standardization and performativity, generates tensions amongst the school actors. Performative pressures lead to negative emotions in teachers, a decreased use of a creative or personalized instruction, and resistance to change and innovation in classrooms. A decline in teaching creativity is associated with mental tiredness, low energy (Teacher, school 7), loss of enthusiasm (Teacher, school 9), and performance-based anxiety and competitiveness" (Teacher, School 12). The tensions between accountability and pedagogical innovation thus become clear.

### (3) Disadvantaged school settings

Unprivileged schools, characterized by a working-class student body and very low performance, face significant challenges in implementing innovative pedagogical and organizational processes. These schools, located in the suburbs of the city, struggle with subscription due to a less attractive educational offer. Parents and the public administration are perceived by schools as absent and not supportive of schools' actions. Schools' ethos mainly focuses on disciplining and controlling students' behavior.

In unprivileged schools, a traditional educational approach prevails, and there are notable difficulties in creating innovative teaching environments. Personnel shortages and overworked leadership teams hinder the implementation of innovative practices. According to teachers, a lack of cohesion between staff members further limits the schools' capacity to introduce laboratory-type or transversal teaching, which demands additional efforts, collaborative planning, and exchanges among teachers.

“The principal works in two schools and only comes once a week in this school, and therefore there is a lack of clear guidelines, there is no cohesion, and there is a strong disorientation which affects the possibilities that the school has” (Teacher, School 10).



School buildings owned by the public administration are also perceived as "obsolete," with inadequate spaces that hinder the stimulation of innovation in teaching ideas and students' creativity (Teacher, school 9). The lack of functionality in terms of digital devices further impedes changes. Innovative educational practices like hands-on teaching, personalized instruction, or student group work are constrained by the learning environment and would also require a different type of classroom setting (Teacher, school 10).

“For me there is a real space problem and there are very few additional classrooms. Sometimes it would be necessary to work with small groups of students and to have personalized lessons. It may seem obvious but it is not obvious in everyday teaching” (Teacher, school 12).

The lack of suitable spaces for collaborative and innovative teaching methods thus becomes a significant barrier to the effective implementation of innovative practices. Teachers working in these school contexts emphasize the real challenge of limited space, particularly the shortage of additional classrooms to implement group work or provide personalized lessons, presenting a daily challenge in delivering a diversified and personalized type of teaching.

### *Experiencing contradictory pressures: Filtering and Decoupling policy*

In terms of responses to the institutional mandates, in unprivileged school settings, teachers select some policy messages in and others out, entailing either engaging or dismissing certain ideas and approaches (Coburn, 2001). The complex and contradictory pressures emanating from the dual mandates create a challenging environment for teachers to manage.

Teachers working in these schools are often against the PBA mandate for different reasons. They view the INVALSI tests as unfair, particularly when perceived as a top-down bureaucratic and statistical requirement that favors high-performing schools without deepening or analyzing the contextual conditions of such performances. The standardized tests are either criticized for not considering the schools' socio-economic background, because the results are too difficult to understand and analyze for teachers, or because they are unrelated to students' interests and inducing stress and anxiety in students.

“We have to lower the emotional filter that is created every time we have a test, we are all quite agitated” (Teacher, school 10).

Teachers in disadvantaged schools also downplay the significance of the PBA model when it does not align with broader school priorities (Teacher, school 12) or lacks an added value (Teacher, school 10). In disadvantaged school contexts, the disconnect between the external assessments and students' life context, coupled with content misalignment with teachers' curriculum, therefore diminishes teachers' perceived utility of the PBA model.

The tension between accountability and pedagogical innovation also emerges clear in such school contexts. Teachers express reservations about standardized testing, citing its failure to stimulate students' reasoning or high-order thinking skills. This conflict is illustrated by the struggle between ticking boxes and fostering deeper understanding and reasoning.

“Through philosophy we learn how to think and reason. But if you ask me to tick the box, then I am against this type of testing (...) you see that students just proceed mechanically they answer without even understanding what is being asked of them” (Teacher, school 12).

However, this negative opinion on the PBA system differs slightly according to the school. While the majority of teachers in disadvantaged schools share such negative perceptions, some others witness the diagnostic role of INVALSI tests, believing that it could and should be more often integrated in teachers' didactics (school 14).

In terms of pedagogical approaches, a traditional teaching model prevails, with face-to-face instruction and a reliance on textbooks being the preferred method. While teachers express ideal preferences for innovative teaching methods (including real-life problems exercises, argumentative teaching and peer learning methodologies), the practical challenges of implementation, such as requiring additional effort or prior preparation, coupled with the perceived need for structured teaching to manage student's energy and lack of autonomy, results in prioritizing and preferring traditional and structured teaching approaches.

“A bit of traditional school must not disappear because otherwise students do not learn the content” (Teacher, school 12).

Yet, the adoption of innovative pedagogical approaches changes according to teachers and their levels of preparation within a school. While some teachers adopt creative teaching practices, the use of digital tools or cooperative learning methodologies, some others, also within the same school, recognize that such practices are not common, and that a traditional method is prevalent especially within the most aged teachers (school 12). In other school contexts, teachers autonomously adopt a combination of both traditional teaching and unstructured dialogic teaching, as well as diversified activities and group work (school 14). These practices are supported by a feeling of being adequately prepared to do so, although this often means “being the only one that makes students work in such a way” (teacher, school 14). Other barriers recognized by teachers towards the adoption of laboratorial or interdisciplinary teaching are systemic conditions, such as high teaching autonomy, which renders teaching a solitary practice, as well as limited time for training and an inadequate support for continuous professional development. Teachers in unprivileged contexts perceive the mandatory training as ineffective and struggle with a defensive climate against continuous institutional changes, impeding a positive relationship with innovation.

“Schools are oppressed by continuous transformations that come from outside and the climate welcoming new things is very defensive. (...) in short, there is a difficult relationship with innovation” (Teacher, school 11).

Consequently, disadvantaged schools often respond symbolically to the institutional pressures, by indicating improvement actions on the internal evaluation reports, but often without translating them into substantial instructional activities. This form of decoupling in schools reflects a pragmatic approach to meet external expectations without necessarily translating them into meaningful changes, but continue doing ‘business as usual’ (Hallett, 2010).

“This year, there was a decline in math results. So, as a priority in the improvement plan, we needed to write that “we will try to improve results in Italian, in mathematics and in a foreign language”. But then, in terms of specific actions, nothing more is done” (Teacher, school 12).

In disadvantaged schools teachers also often adopt instrumental strategies in classrooms, such as extensive preparation focused on test items (see Popham, 2001) during the whole school year and selective use of texts to prepare for INVALSI. Such emphasis on direct and fact-centered

instruction highlights the pragmatic adaptation of teaching methods to meet PBA demands, as witnessed from the following quotes.

“I look for an example of texts that are more of interest to them, for instance selecting some about cyberbullying or drugs” (Teacher, School 10)

“Sometimes, instead of experiential and hands-on activities, we have to take the history textbook and work in a teacher-centered way (...). We do a lot of test exercises on the tests, since INVALSI provides example exercises, and throughout the year, I mainly work on reading and understanding texts” (Teacher, school 11).

Teachers in disadvantaged schools thus more often employ an "instrumental and technical orientation to change" balancing conflicting priorities and aligning classroom structures with different mandates, even when this is in contradiction with educational and pedagogical values related to innovation (Niesz, 2010). However, despite reservations about the fairness or usefulness of INVALSI, some teachers recognize the need for students to excel within the existing evaluation system.

“If the question is if I am against the INVALSI tests as an evaluation system, yes, I am against it and I do not consider it a useful or fair criteria or method of evaluating. But after having said that, these evaluations exist, and since it quantifies student learning, students have to know how to do them as best as possible” (Teacher, School 11).

Moreover, the rationale behind adopting teaching to the test strategies is not always the same and may also resemble educational and pedagogical aims, ranging from “making students understand the approach adopted by the INVALSI tests” (Teacher, school 12) to adopting test preparation as a valid and additional way to evaluate students’ learning (school 14) or as a way to discuss about current issues, hence “practicing for the INVALSI does not become an end in itself to but also something more about civic education” (Teacher, school 10).

Such conflicting and diverse responses reflect the challenges in reconciling pressures between accountability and pedagogical innovation that teachers in underprivileged school settings face and the employment of adaptive strategies to navigate the competing pressures they feel deriving from the two diverse mandates.

## Discussion and conclusions

This study explored the intricate relationship between PBA and innovation in the context of educational reforms, focusing on Italian schools, which is an underexplored case from this perspective. Framed by sensemaking and filtering theories, the paper analyzed the tension between accountability and innovation at the school level. Through an ideal case analysis, it investigated how school actors interpret and adapt the dual policy mandates, emphasizing the influence of the local school context on these dynamics.

While the Italian PBA system aims to promote innovation, and align PBA and innovation (Mentini & Levatino, 2024), in practice this goal is not as straightforward. According to the analysis, at the implementation level, school actors reveal diverse interpretations and adaptations of the PBA and innovation mandates based on their understandings and interpretations. Teacher attitudes towards the PBA system play a central role in shaping their responses. Those teachers harboring negative opinions are more inclined to resist and reject the mandate, adapting teaching-to-the-test and curriculum narrowing practices. Conversely, when teachers understand the PBA as fostering transversal competences and critical thinking, the alignment between policy and practice becomes more apparent. Moreover, given the traditional and content-based educational legacy in Italy, innovation is often associated with a new and “alternative” way of teaching, including active learning, competence-based, project-based activities, hands-on teaching, and pedagogical changes. Thus, when the external assessment is understood in support of such types of approaches, teachers tend to integrate it more easily in their teaching and schools use it as an innovative and improvement device.

The socio-economic school context and performative pressures also emerge as key determinants, influencing how teachers interpret and adapt their practices to the institutional messages. Schools catering to middle and high-class students, partly because they experience less pressure, are also more able to align internal and external accountability to drive innovation, adopting various strategies that meet the expectation of school improvement based on the external test data. Conversely, in low-performing contexts with low-class students, teachers decouple and escape from institutional pressures, employing instrumental strategies to balance conflicting priorities and align with PBA demands. Moreover, ethnically diverse low-income school contexts witness high

tensions between the PBA and innovation mandates. School actors in these contexts struggle to balance performative demands with the capacity for creativity, personalized instruction, and the stimulation of deeper understanding and high-order thinking skills. Undesirable outcomes such as intense student test training and traditional teaching approaches become more pronounced in such settings. Finally, in disadvantaged school contexts, despite expressing a preference for innovation, a perceived need for structured teaching as well as practical challenges, leads teachers to rely on traditional methods, highlighting the tension between their educational values and the demands of PBA. On the contrary, in high or middle-income homogenous contexts, with high performance results, the accountability mandate represents a stimulus for diversification in teaching and a competence-based transversal teaching. Yet, instrumental strategies such as teaching to the test are also found in middle income school contexts, however adopted for diverse rationales and intensity depending on the school student population.

Furthermore, a constant pressure to innovate and improve, coupled with the performative and reputational pressure stemming from the PBA mandate, is felt by the majority of school actors, irrespective of the schools' socio-economic characteristics. Administrative and bureaucratic tasks stemming from the PBA also create a burden in some schools, impact teachers' emotional experiences and influence their practices. Feelings of self-imposed pressure and the negative emotional experiences deriving from the PBA influence the use of creativity and alternative or innovative teaching practices.

The innovation mandate in Italy is however nuanced and ambiguous, associated with diverse pedagogical and organizational features, thus it is understood and adapted to different school agendas and priorities, depending on schools' socio-economic status. Disadvantaged schools tend to understand innovation as *process innovation* focusing on teaching and learning, with a greater focus on socio-emotional skills or manipulative and experiential teaching. In heterogeneous schools, teachers adopt alternative teaching methods to overcome student's learning difficulties and behavioral issues since they are less demanding and respond to more "practically-oriented" learning needs. Schools better positioned use innovative practices to teach high-level order and problem-solving skills or use organizational changes, network collaborations and extra-curricular activities to enrich their educational offer. Interestingly, in privileged school contexts, innovative practices are predominantly used to retain parents and maintain a high reputation, responding to a market- model based on family's free choice (Lubienski, 2009). Indeed, given the Italian quasi-market education system, pressures do not only come from administrative and bureaucratic forms

of accountability, but also from parents' expectations, who play a role in influencing school practices, despite their educational demands and preferences are far from homogenous. Within upper and middle-class families, some groups of parents claim the use of traditional education models and are tied to academic goals, whereas other groups ask for more experiential-oriented and project-based activities in schools.

Beyond the broad socio-economic conditions, it also emerges that in the analyzed dynamics the schools' leadership and the cohesion among staff play a particular role in mitigating school responses. Schools exhibiting cohesive professional communities and visionary leadership, also display a higher adoption of innovative practices and less superficial responses to the PBA mandate. In fact, stable contracts for teachers and continuous professional self-development contribute to higher cohesion and collaborative teaching, supporting the adoption of institutional changes in schools, extracurricular and interdisciplinary teaching. This tends to disadvantage schools serving low-income students, given the fewer and less stable professional resources that can set a clear educational approach or support long-term educational planning. Moreover, a clear leadership direction in sharing a correct understanding of the accountability policy in schools seems to support an internal accountability and positive learning environment by which teachers are less likely to exercise control over student curriculum and performance.

In the study, it finally emerges that material elements, including school buildings and infrastructure, play a crucial role in supporting innovative methodologies, such as collaborative classes, laboratorial teaching, digital education or outdoor education. Policy implications suggest the need to reinforce structures and create conditions for change and innovation especially in more disadvantaged school contexts. Supporting school autonomy through increased financial and managerial resources and personnel is crucial, highlighting the importance of such type of support especially in larger schools, with higher need of organization, or in disadvantaged schools where management teams and teachers are subject to more frequent changes and internal rotations.

These findings highlight potential avenues for future research that could contribute to a more comprehensive understanding of the complex relationship between PBA and innovation in education. Firstly, there is a need to delve deeper into the emotional experiences of teachers under PBA pressures, examining the impact of PBA on teachers' well-being, job satisfaction and teaching practices. Secondly, a closer examination of the interplay between schools' socio-economic contexts and innovation would offer a deeper understanding of how innovation manifests in

different settings and the role of the school socio-economic composition in influencing teaching and organizational school practices. Lastly, further research could adopt a comparative perspective and explore the relation between PBA and educational innovation across contexts, investigating how the discrepancies between policy discourse and practical implementation may differ according to countries, identifying both barriers and enablers and considering cultural, institutional and policy variations.

## References

- Appel, M. (2020). Performativity and the demise of the teaching profession: The need for rebalancing in Australia. *Asia-Pacific Journal of Teacher Education*, 48 (3), 301–315.
- Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher*, 36 (5), 258–267.
- Avalos, M. A., Perez, X., & Thorington, V. (2020). Comparing secondary English teachers' ideal and actual writing practices for diverse learners: Constrained professionalism in figured worlds of high-stakes testing. *Reading & Writing Quarterly*, 36(3), 225–242.
- Ball, S. J., & Maroy, C. (2009). School's logics of action as mediation and compromise between internal dynamics and external constraints and pressures. *Compare*, 39(1), 99–112.
- Ball, S., Maguire, M., & Braun, A. (2012). *How schools do policy: Policy enactments in secondary Schools*. Routledge.
- Barrett, A. M. (2007). Beyond the polarization of pedagogy: Models of classroom practice in Tanzanian primary schools. *Comparative Education*, 43(2), 273–294.
- Barzanò, G., & Grimaldi, E. (2012). “Policy” valutative e contesti di applicazione. Caratteristiche Procedurali.nECPS - *Educational, Cultural and Psychological Studies*,n06(6), 159–189.
- Benadusi, L., & Consoli, F. (Eds.). (2004). *La governance della scuola*. Il Mulino.
- Biondi, G., Mosa, E., & Panzavolta, S. (2009). *Autonomia e innovazione: Scenari possibili tra teoria e pratica, programma education*. FGA Working Paper, 16(2), 1–37.
- 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.



- Burns, T., & Paniagua, A. (2018). *Understanding innovative pedagogies: Key themes to analyse new approaches to teaching and learning*. OECD Education Working Papers, 172(172), 1–135.
- Cecchi, D., & Mattei, P. (2021). Merit pay for schoolteachers in Italy, 2015–2016: A new regime of education accountability? *Comparative Education Review*, 65(3), 445–466.
- Coburn, C. E. (2001). Collective sensemaking about reading: How teachers mediate reading policy in their professional communities. *Educational Evaluation and Policy Analysis*, 23(2), 145–170.
- Coburn, C. E. (2004). Beyond decoupling: Rethinking the relationship between the institutional environment and the classroom. *Sociology of Education*, 77(3), 211–244.
- Colombo, M., & Desideri, A. (2018). Italy. The Italian education system and school autonomy. In S. Martins, L. Capucha, & S. O. J. (Eds.), *School autonomy organization and performance in Europe. A comparative analysis for the period from 2000 to 2015* (pp. 101–113). CIES – ISCTE (Centre for Research and Studies in Sociology).
- De Angelis, M., Marzano, A., & Iannotta, I. S. (2015). *TEACHER EVALUATION in ITALY: PROBLEMATIC NODES and CRITICAL FEATURES*. Proceedings of EDULEARN15 Conference 6th–8th July 2015, Barcelona, Spain.
- De Sanctis, G. (2010). *TALIS. I docenti italiani tra bisogni di crescita professionale e resistenze*. Fondazione Giovanni Agnelli Programma Education FGA Working Paper, 24(2), 1–34.
- Deterding, N. M., & Waters, M. C. (2021). Flexible coding of in-depth interviews: A twenty-first-century approach. *Sociological Methods & Research*, 50(2), 708–739.
- Diamond, J. B. (2007). Where the rubber meets the road: Rethinking the connection between high-stakes testing policy and classroom instruction. *Sociology of Education*, 80(4), 285–313.
- DiCresce, C. (2019). *Cosa si sa e cosa si pensa dell'INVALSI: uno studio esplorativo*, Working paper INVALSI n. 38/2019.
- Diehl, D. K., & Golann, J. W. (2023). An integrated framework for studying how schools respond to external pressures. *Educational Researcher*, 52(5), 296–305.
- Dulude, E., & Milley, P. (2021). Institutional complexity and multiple accountability tensions: A conceptual framework for analyzing school leaders' interpretation of competing demands. *Policy Futures in Education*, 19(1), 84–96.
- Elmore, R. (2005a). The accountable leadership. *The Educational Forum*, 69(2), 134–142.
- Elmore, R. (2005b). *Agency, reciprocity, and accountability in democratic education*. Consortium for Policy Research in Education.

- European Commission, Directorate-General for Education Sport and Culture, Y. & PPMI. (2018). *Study on supporting school innovation across Europe: Final report*. [http://publications.europa.eu/publication/manifestation\\_identifier/PUB\\_NC0117315EN](http://publications.europa.eu/publication/manifestation_identifier/PUB_NC0117315EN) N%0Ahttps://www.schooleducationgateway.eu/downloads/innovation/InnovationStudy.pdf
- Faggioli, M., & Mori, S. (2018). Valutare l'innovazione scolastica: Vincoli ed opportunità nel sistema nazionale di valutazione. In M. Freddano e & S. Pastore (a cura di) *Monografia, Per una valutazione delle scuole, oltre l'adempimento* (pp. 88–99). Milano: Franco Angeli Edizioni.
- Fahey, G., & Koster, F. (2019). *Means, ends and meaning in accountability for strategic education governance*. OECD education working papers, (204).
- Falabella, A. (2020). The ethics of competition: Accountability policy enactment in Chilean schools' everyday life. *Journal of Education Policy*, 35(1), 23–45.
- Farvis, J., & Hay, S. (2020). Undermining teaching: How education consultants view the impact of high-stakes test preparation on teaching. *Policy Futures in Education*, 18(8), 1058–1074.
- Giancola, O., & Salmieri, L. (2022). Cross-national achievement surveys and educational monitoring in Italy, chapter 4. In L. Volante, S. V. Schnepf, & D. A. Klinger (Eds.), *Cross-national achievement surveys for monitoring educational outcomes: Policies, practices and political reforms within the European Union* (pp. 69–90). EUR 30380 EN, Publications Office of the European Union.
- Grimaldi, E., & Barzanò, G. (2014). Making sense of the educational present: Problematising the “merit turn” in the Italian eduscape. *European Educational Research Journal*, 13(1), 26–46.
- Grimaldi, E., & Serpieri, R. (2012). The transformation of the education state in Italy: A critical policy historiography from 1944 to 2011. *Italian Journal of Sociology of Education*, 4(1), 146–180.
- Gunnulfson, A. E., & Roe, A. (2018). Investigating teachers' and school principals' enactments of national testing policies: A Norwegian study. *Journal of Educational Administration*, 56(3), 332–349.
- Hallet, T. (2010). The myth incarnate: Recoupling processes, turmoil, and inhabited institutions in an urban elementary school. *American Sociological Review*, 75(1), 52–74.
- Hardy, I. J., Reyes, V., & Hamid, M. O. (2019). Performative practices and ‘authentic accountabilities’: Targeting students, targeting learning? *The International Education Journal: Comparative Perspectives*, 18(1), 20–33.

- Herman, J. (2004). The effects of testing on instruction. In S. Fuhrman & R. Elmore (Eds.), *Redesigning accountability* (pp. 141–146). Teachers College Press.
- Hoadley, U. (2006). *The reproduction of social class differences through pedagogy: A model for the investigation of pedagogic variation*. University of Cape Town & Human Sciences Research Council, 5692(April), 1–31.
- INVALSI. (2023). *INVALSI Open Sito Ufficiale Area Prove Nazionali*. Retrieved September 15, 2023, from [https:// www.invalsiopen.it/](https://www.invalsiopen.it/)
- Kelchtermans, G. (2005). Teachers' emotions in educational reforms: Self-understanding, vulnerable ment and micropolitical literacy. *Teaching & Teacher Education*, 21(8), 995–1006.
- Kickert, W. (2007). Public management reforms in countries with a Napoleonic state model. In C. Pollitt, S. Van Thiel, & V. Homburg (Eds.), *New public management in Europe: Adaptation and alternatives* (pp. 26–51). Palgrave/Macmillan.
- Knight, R. (2020). The tensions of innovation: Experiences of teachers during a whole school pedagogical shift. *Research Papers in Education*, 35(2), 205–227.
- Kohn, A. (1998). Only for my kid: How privileged parents undermine school reform. *Phi Delta Kappan*, 79(8), 568–578.
- Landri, P. (2021). To resist, or to align? The enactment of data-based school governance in Italy. *Educational Assessment, Evaluation and Accountability*, 33(3), 563–580.
- Lelo, K., Monni, S., & Tomassi, F. (2019). Socio-spatial inequalities and urban transformation. The case of Rome districts. *Socio-Economic Planning Sciences*, 68, 100696. <https://doi.org/https://doi.org/10.1016/j.seps.2019.03.002>
- Looney, J. W. (2009). *Assessment and innovation in education*. OECD Education Working Papers, 24(24), 62.
- Lubienski, C. (2009). *Do quasi-markets foster innovation in education? A comparative perspective*. Report, OECD education working papers, No. 25. OECD Publishing.
- MacBeath, J. (2008). Leading learning in the self-evaluating school. *School Leadership & Management*, 28(4), 385–399.
- Martini, A., & Papini, M. (2015). *Che cosa ne pensano gli insegnanti delle prove INVALSI*. Working paper 24/2015. INVALSI.
- Mattei, P. (2012). Market accountability in schools: Policy reforms in England, Germany, France and Italy. *Oxford Review of Education*, 38(3), 247–266.

- Mentini, L., & Levatino, A. (2024). A “three-legged model”: (De)constructing school autonomy, accountability, and innovation in the Italian National evaluation system. *European Educational Research Journal*, 23(3), 321–346.
- Ministero dell’Economia e delle Finanze - Dipartimento delle Finanze *Dichiarazioni 2020 - Anno d’imposta*. (2019). Retrieved from [https://www1.finanze.gov.it/finanze/analisi\\_stat/public/index.php?tree=2020](https://www1.finanze.gov.it/finanze/analisi_stat/public/index.php?tree=2020)
- Ministry of Education. (2023). *Portale Unico dei Dati della Scuola*. <https://dati.istruzione.it/opendata/opendata/catalogo/elements1/?area=Scuole>
- 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). Springer.
- Nardi, A., Rossi, F., & Toci, V. (2022). Le dimensioni dell’innovazione: Un framework per la valutazione dei processi di innovazione scolastica. *IUL Research*, 1(1), 144–159.
- National plan for digital education (Piano Nazionale Scuola Digitale). (2015). Italy: Ministero dell’Istruzione, dell’Università e della Ricerca.
- Niesz, T. (2010). “That school had become all about show”: Image making and the ironies of constructing a good urban school. *Urban Education*, 45(3), 371–393.
- OECD. (2013). *TALIS 2013 results: An international perspective on teaching and learning*. Report. OECD Publishing.
- OECD. (2019). *TALIS 2018 results (volume I): Teachers and school leaders as lifelong learners*, TALIS. OECD Publishing.
- Paletta, A., Basyte Ferrari, E., & Alimehmeti, G. (2020). How principals use a New accountability system to promote change in Teacher practices: Evidence from Italy. *Educational Administration Quarterly*, 56(1), 123–173.
- Pandolfini, V. (2009). Public or private education? parents’ choices between actual and potential pluralism. *Italian Journal of Sociology of Education*, 5(2), 189–217.
- Parcerisa, L., & Verger, A. (2023). *Researching ‘autonomy with accountability’ in schools: A qualitative approach to policy enactment and practice*. REFORMED Methodological Papers, 3, 1–33.
- Pastori, G., & Pagani, V. (2016). Cosa pensate dei test INVALSI? Dirigenti scolastici, insegnanti e studenti provenienti dalla Lombardia descrivono la loro esperienza. *Journal of Educational, Cultural, and Psychological Studies*, 2016(13), 97–117.

- Patton, M. Q. (2015). *Qualitative research and evaluation methods: Integrating theory and practice* (4th ed.). Sage.
- Peruzzo, F., Grimaldi, E., Arienzo, A., D’Onofrio, G., Franchi, C., & Sebastianelli, P. (2022). New public management reforms and industrial relations in the Italian education system. A cultural political economy approach. *Journal of Educational Administration and History*, 55(4), 381–399.
- Poliandri, D. (2015). *I processi e il funzionamento delle scuole - Dati dal Questionario Scuola INVALSI (RAV) e dalle sperimentazioni. VALES e VM INVALSI.*
- Popham, W. J. (2001). Teaching to the test. *Educational Leadership*, 58(6), 16–20.
- Porac, J. F., Thomas, H., & Baden Fuller, C. (1989). Competitive groups as cognitive communities: The case of Scottish Knitwear manufacturers. *Journal of Management Studies*, 26(4), 397–416.
- Sahlberg, P. (2006). Education reform for raising economic competitiveness. *Journal of Educational Change*, 7, 259–287.
- Sahlberg, P. (2009). The role of education in promoting creativity: Potential barriers and enabling factors school education and creativity. *Measuring Creativity*, 337, 337–344. <http://www.greenschool.org/wp-content/uploads/2012/03/Pasi-Sahlberg.pdf>
- Schoen, L., & Fusarelli, L. D. (2008). Innovation, NCLB, and the fear factor. *Educational Policy*, 22(1), 181–203.
- SITES. (2009). *Second information technology in education study*. SITES 2006 technical report. <https://www.ica.nl/sites-2006>
- Stapley, E., O’Keeffe, S., & Midgley, N. (2022). Developing typologies in qualitative research: The use of ideal-type analysis. *International Journal of Qualitative Methods*, 21, 1–9.
- Tarabini, A., & Rotger, J. M. (Eds.). (2012). *Sociologia del currículum i de la praxi educativa*. In *Sociologia de l’Educació per a professorat de Secundària Barcelona: El Roure* (1st ed., pp. 289–316).
- Thornton, P., & Lounsbury, M. (2012). *The institutional logics perspective: A new approach to culture, structure and process*. Oxford University Press.
- Verger, A., Ferrer-Esteban, G., & Parcerisa, L. (2021). In and out of the ‘pressure cooker’: Schools’ varying responses to accountability and datafication. In S. Grek, C. Maroy, & A. Verger (Eds.), *World yearbook of education 2021: Accountability and datafication in the governance of education* (pp. 219–239). Routledge.

- Verger, A., & Parcerisa, L. (2017). *Accountability and education in the post-2015 scenario: International trends, enactment dynamics and socio-educational effects*. Paper commissioned for the 2017/8 Global Education Monitoring Report, Accountability in education: Meeting our commitments. <http://unesdoc.unesco.org/images/0025/002595/259559e.pdf>
- Verger, A., Parcerisa, L., & Fontdevila, C. (2019). 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.
- Vincent-Lancrin, S., Urgel, J., Kar, S., & Jacotin, G. (2019). *Measuring innovation in education 2019: What has changed in the classroom?*. OECD Publishing.
- Yin, R. (2009). *Case study research: Design and methods* (4th ed.). SAGE.

## CHAPTER 5: SUMMARY OF THE MAIN RESEARCH FINDINGS

This doctoral dissertation has undertaken a comprehensive examination of the relationship between PBA, innovation policies and teachers' emotions in the education sector. Through an analysis of teacher emotional experiences, policy discourses, and school enactment strategies, I shed light on the different aspects which, at different but simultaneous and interconnected levels, influence the selection, adoption and professional experiences of PBA in the educational sector.

The scoping review in **Chapter 2** presented the emotional implications of PBA on teachers' subjective and professional experiences. The review shows how PBA is a crucial part of a changing professional environment that reinforces and/or modifies feelings and emotions already existing in teachers' professional experience. Specifically, several analyzed studies indicate that PBA in education may reinforce expectations on socially desirable emotions and self-regulation that teachers should display (Bodenheimer & Shuster, 2020). The self-regulation of emotions, similar to what happens in other service-oriented/caring occupations, is considered inherent to the teaching profession, but reinforced under PBA policies. Moreover, new norms such as impersonal distance, neutrality, and disengaged objectivity can be reinforced by PBA policies (Steinberg, 2008), neglecting the intrinsic emotional nature of the teaching profession (Chávez et al., 2021; Perold et al., 2012). The increased use of standardized testing also accentuates a range of emotions that are already inherent to the teaching profession such as irritation, stress, shame, disgust, guilt, pride, joy, sympathy (Stenberg, 2008). The review also highlights that the effects of PBA on teachers' emotions are varied and mixed, ranging from positive emotions such as joy, pride and job satisfaction (Holloway & Brass, 2017; Falabella, 2020) to unpleasant and negative emotions, including stress, preoccupation, disillusionment, frustration, fear, dissatisfaction, anxiety and pain. The reviewed literature thus recognizes a nuanced scenario associated with both the «pleasures» and the «terrors» related to PBA reforms (cf. Holloway & Brass, 2017, p.380). Moreover, the analysis highlights a number of explanations for the emotions that teachers experience as a result of PBA reforms. One of the most relevant is that not only the material consequences attached to PBA policies (e.g. material stakes attached to performance, such as hiring and firing decisions or financial incentives or sanctions) can bring a range of negative emotional effects on teachers, but also the risk for teachers of being judged and the potential effect this has on social recognition and

teachers' professional reputation, can discourage teachers from engaging with creative or innovative pedagogical practices and limit their agency (Pierlejewski, 2020).

Moreover, the review highlights that PBA policies can have a profound emotional toll on teachers, as they navigate the demands of accountability while striving to meet the wellbeing and emotional needs of their students. The misalignment between competing mandates and models of good teaching, as well as the intensification of tasks and responsibilities in order to meet the accountability requirements is a key explanatory factor for teachers' negative emotional experiences following PBA reforms. Chapter 2 further sheds light on the role of teacher agency and strategies employed to deal with the tension of conflicting priorities and demands, as well as the implications of PBA pressure on teachers' engagement with alternative practices. In order to avoid adverse emotions following low results (Steinberg, 2008), and risk and feelings of uncertainty deriving from innovative practices requiring time and attention (Falabella, 2020), teachers engage less with innovative or alternative pedagogical practices. Through «pragmatic compliance» (Braun & Maguire, 2020) teachers respond effectively to accountability demands, even when it means to contradict educational and pedagogical values related to innovation (Niesz, 2010). This Chapter finally contributes to illuminate a number of factors, at macro, meso and micro levels that intensify and weaken the emotional impact of PBA. Importantly, systemic aspects, as well as cultural and societal values can influence how teachers perceive and enact PBA, and their feelings towards these reforms. Moreover, the school culture, the socioeconomic background of the student population and the quality of relationships among school actors, are found to explain different teachers' emotional experiences under PBA contexts. For instance, teachers working in lower socioeconomic environments are found to do more emotional labor in order to maintain a balance between trying to meet the accountability demands and students' learning progress (Steinberg, 2013). The role of the principal in mediating teachers' emotional experiences of PBA policies is also recognized. Finally, teachers' working conditions (including subject taught, contract status), their individual sense-making (beliefs, pedagogical identity and task perception), and personal characteristics (such as age and gender) are found to influence how teachers are emotionally affected by PBA. A key finding is that teachers' beliefs about their pedagogical role and task influences how they interpret and respond to the expectations deriving from such changes. For instance, when teachers' educational goals are aligned to accountability demands, teachers experience a lower degree of stress (Shaw, 2016).

**Chapter 3** has delved into the underpinnings of accountability and autonomy reforms on innovation, within the Italian national evaluation system (SNV), shedding light on the assumptions,



rationales and tensions of these two mandates, from the point of view of the actors involved in the ideation and policy design of the reforms. The findings of this Chapter highlight four main assumptions underpinning the relationship between autonomy, accountability and innovation and how they are intricately combined to generate the intended change in the Italian accountability system. These include: 1) the reliance on a substantial degree of autonomy to introduce innovation and change in curriculum and organizational aspects 2) the need to counterbalance school autonomy through external evaluations by means of standardized tests, as a way to enhance the efficacy of the educational system 3) innovation and improvement in schools deriving from schools' use of feedback from external test results and 4) the enhancement of innovation in teaching strategies as a mechanism to boost external test results. The peculiarities of the PBA Italian system are also highlighted in this Chapter, which emerges as an articulated, equilibrated and comprehensive system comprising a cyclical evaluation process, involving three main national actors: (1) The INVALSI institute, in charge of evaluation indicators and frameworks, and the external standardized tests; (2) INDIRE institute, responsible for improvement actions and innovative practices, and (3) the ministerial inspectorate, in charge of the out external evaluation of schools. Still, the findings reveal contradictions in some of the premises, since the datafication coming from the standardized tests and the INVALSI institute overshadows the other two components.

The Chapter further sheds light on various rationales (including improvement, efficiency, equity and transparency) that have acted as drivers of PBA reforms in Italy. Some of these rationales (e.g. preoccupation with efficiency and quality) are in line with other Southern European countries (cf. Verger et al., 2019). Yet, in Italy, equity and transparency arguments also emerge as justification of the PBA reforms, in line with certain Nordic countries (cf. Camphuijsen and Levatino, 2022). It also emerges that equity, transparency and diagnostic rationales are used as an intentional political maneuver to increase consensus around the introduction of external accountability assessments in light of unions' objections and protests (e.g., Kickert, 2007). In addition, the influence of globalizing discourses on international competition and the benefits of datafication appear to have a significant role in justifying the adoption of PBA policies in the Italian context. Finally, the confluence of factors such as solicitations from international organizations, coupled with the domestic economic crisis in Italy during that period, served as contextual conditions acting as a "window of opportunity" (Kingdon, 1984) for the adoption of the PBA reforms.

The Chapter finally illustrates a number of different contextual aspects that are considered by the stakeholders to hamper the expected change mechanisms, thus highlighting the discontinuous ground in which such policy dispositifs operate. These are related to (a) the structural features of

the Italian educational context, that include the conditions of school buildings, the economic precariousness of the teaching profession or the lack of human resources in schools; (b) school actors' characteristics and competences, such as the professional competences (knowledge, skills and competences) and personal characteristics (including personality, vision and willingness to change) of teachers and principals needed to facilitate school innovation and use the PBA data and (c) cultural features such as an embedded traditional, knowledge-based culture and teaching method or a lack of a culture of evaluation. More importantly, according to the interviewed actors, the achievement of the accountability systems' outcomes is limited by teachers' and principals misunderstanding and misuse of the purpose and usefulness of the accountability system, shedding light on the supposed incompetency of school actors as opposed to the virtues of external testing, external experts, and reliable data.

**Chapter 4** has provided empirical evidence of the negotiation of PBA and innovation mandates in Italian schools, emphasizing how school actors interpret, filter, and adapt the policy expectations to their organizational and educational practices. Special attention has been given to the influence of local school contexts and performative pressures on this process.

While the Italian accountability system is intended to promote change and innovation, integrating and operationalizing accountability and innovation in schools proves to be complex. Findings revealed different school responses to the double mandates of PBA and innovation, which are influenced by teachers' attitudes towards PBA and their prior pedagogical beliefs. For instance, when teachers in Italian schools view PBA as promoting transversal competences and critical thinking, there is better alignment between policy and practice, and teachers are more inclined to integrate the data into their teaching.

Moreover, in Italy, although PBA policies emphasize self-evaluation, diagnostic and improvement logics to generate pedagogical change, they can inadvertently stifle creativity and innovation within schools, generating tensions within school actors, who struggle to balance the demands of performativity and standardization with the capacity for creativity, personalized instruction, and the development of high-order thinking skills. Notwithstanding the Italian PBA model being relatively "mild" (Pensiero et al., 2019: 84), not foreseeing punitive or material consequences attached to its outcomes, a performative (self-imposed) pressure is felt by the school actors. This pressure is based on performative and reputational consequences stemming from the PBA mandate, influencing principals and teachers' experience and perceptions of PBA reforms, and often restricting the use of creativity and alternative or innovative teaching practices in some

school contexts. The emotional tool and pressure felt is also strengthened by the burden of administrative and bureaucratic tasks stemming from PBA reforms.

Socio-economic factors and performative pressures significantly influence how teachers interpret the mandates and adapt their practices. Schools with middle and high-class students face less pressure and can better align internal and external accountability to drive innovation. In contrast, low-performing schools with disadvantaged students often experience high tension between PBA and innovation. This often leads to undesirable outcomes like intense test training and a perceived need for structured teaching that pushes teachers towards traditional methods, creating a tension between their educational values and PBA demands.

Moreover, since the innovation mandate in Italy is based on a not so clear and shared definition, educational innovation is understood and adapted in different ways by the school actors, in relation to the schools' socio-economic contexts and performance results. While disadvantaged and heterogeneous schools tend to view innovation as a process innovation, focusing on socio-emotional skills, experiential teaching methods and emphasizing students' practical learning needs, better-positioned schools use innovative practices to teach high-order and problem-solving skills, along with organizational changes, network collaborations, and extracurricular activities, as a way to retain parents and maintain a high reputation.

Beyond the role of school's socio-economic factors and accountability pressures, this Chapter has highlighted the key role of other dimensions in mediating the alignment between PBA expectations and school practices, underscoring how the use of external accountability mechanisms to drive change and improvement in schools is far from being linear or straightforward. These contextual dimensions are related to a) school organizational dynamics and professional contexts, including the role of school leaders and staff cohesion and b) material contexts, such as school infrastructure and material resources. In relation to the former, the organizational complexity of schools leads to unawareness among teachers regarding accountability standards, practices, and processes, resulting in unintended outcomes such as teaching to the test, intensive student test preparation, or curriculum narrowing. Moreover, stable contracts for teachers, a continuous professional development generated among staff and a clear leadership direction, emerge as key levers to the adoption of change and innovative pedagogical practices in schools, in alignment with the PBA expectations. High rates of faculty turnover in complex centers have an impact on the frequency of innovations. With regards to the latter, material elements, including school buildings, infrastructure, and spaces are seen as facilitating the adoption of innovative practices in schools.

Overall, from a social justice standpoint, a relationship emerges between school autonomy, innovation, and the school's socio-economic status. Disadvantaged schools serving low-income

students lack managerial and financial resources and experience a higher tension in juggling the simultaneous demands of PBA and innovation. Such risks highlight the innovation “fatigue” (Pedrò, 2023), stemming from constant requests for innovation and improvement and the mismatch between external pressure and the reality of schools (Quilabert., 2024), which is not accompanied by recognition or support for teachers.

The Chapter finally contributed to shedding light on the crucial role of interpretative lenses of school actors in relation to PBA, to explain interpretation and filtering dynamics. The sense-making given to the reform significantly influences the importance and interpretation of accountability instruments, subsequently impacting their responses to the policy reform and classroom practices adopted. This finding aligns with the review in Chapter 2, by which teachers’ beliefs about the validity of a test influences their experience towards the reform. In the Italian case, characterized by a traditional and content-based educational legacy and extensive criticisms over PBA given the intrusion felt in the teachers’ professional autonomy (cf. Barone & Argentin, 2016; Barzanò & Grimaldi, 2013), teachers who perceive PBA alignment with their educational beliefs and values, such as the promotion transversal competencies and critical thinking, demonstrate synergies between PBA and innovation and better alignment between expected policy outcomes and adopted practices.

## References

- Bodenheimer, G., & Shuster, S. M. (2020). Emotional labour, teaching and burnout: Investigating complex relationships. *Educational Research*, 62(1), pp. 63–76.
- Barone C and Argentin G (2016) School Reform: Innovation and the rhetoric of Change. *Italian Politics* 31(1): 135–154.
- Barzanò G and Grimaldi E (2013) Discourses of merit. The hot potato of teacher evaluation in Italy. *Journal of Education Policy* 28(6): 767–791.
- Braun, A., & Maguire, M. (2020). Doing without believing—enacting policy in the English primary school. *Critical Studies in Education*, 61(4), pp. 433–447.
- Camphuijsen MK and Levatino A (2022) Schools in the media: framing national standardized testing in the Norwegian press, 2004–2018. *Discourse Studies in the Cultural Politics of Education* 43: 601–616.

- Chávez, R. C., Etcheberrigaray, G., Vargas, S., Assaél, J., Araya, R., & Redondo-Rojo, J. (2021). Actividades emocionales del trabajo docente: un estudio de shadowing en Chile. *Quaderns de Psicologia*, 23(1), pp. 1–26.
- Falabella, A. (2020). The ethics of competition: accountability policy enactment in Chilean schools' everyday life. *Journal of Education Policy*, 35(1), pp. 23–45.
- Holloway, J., & Brass, J. (2017). Making accountable teachers: the terrors and pleasures of accountability. *Journal of Education Policy*, 33(84), pp. 1-22.
- Kingdon JW (1984) *Agendas, Alternatives, and Public Policies*. Boston, MA: Little, Brown.
- Kickert W (2007) Public management reforms in countries with a Napoleonic state model. In: Pollitt C, Van Thiel S and Homburg V (eds) *New Public Management in Europe: Adaptation and Alternatives*. Basingstoke: Palgrave/Macmillan, pp.26–51.
- Niesz, T. (2010). “That school had become all about show”: Image making and the ironies of constructing a good urban school. *Urban Education*, 45(3), pp. 371– 393.
- Pensiero N, Giancola O and Barone C (2019) Socioeconomic inequality and student outcomes in Italy. In: Volante L, Schnepf S, Jerrim J, et al. (eds) *Socioeconomic Inequality and Student Outcomes: Cross-National Trends, Policies, and Practices*. Singapore: Springer, pp.81–94.
- Pedro F. (2023), Chapter 10: Where is the school going? International trends in educational innovation, *Sociology, Social Policy and Education*, pp 147–164,
- Perold, M., Oswald, M., & Swart, E. (2012). Care, performance and performativity: Portraits of teachers' lived experiences. *Education as Change*, 16(1), pp. 113– 127.
- Pierlejewski, M. (2020). The data-doppelganger and the cyborg-self: theorising the datafication of education. *Pedagogy, Culture and Society*, 28(3), pp. 463–475.
- Shaw, R. D. (2016). Music teacher stress in the era of accountability. *Arts Education Policy Review*, 117(2), pp. 104–116.
- Steinberg, C. (2008). Assessment as an “emotional practice. *English Teaching: Practice and Critique*, 7(3), pp. 42-64.
- Steinberg, C. (2013). Teachers dealing with failure in a low socio-economic status school in South Africa. In Craig, C.J., Meijer, P.C. & Broeckmans, J. (Ed.) *From Teacher Thinking to Teachers and Teaching: The Evolution of a Research Community (Advances in Research on Teaching)* (pp. 427-442) Bingley, England: Emerald Group Publishing

Verger A, Fontdevila C and Parcerisa L (2019) Reforming governance through policy instruments: How and to what extent standards, tests and accountability in education spread worldwide. *Discourse Studies in the Cultural Politics of Education* 40(2): 248–270.

## CHAPTER 6: CONCLUSIONS

The findings presented in this dissertation offer valuable insights into the complex interplay between PBA, innovation policies and teachers' work in contemporary education, with a focus on a Southern-European country, Italy. The scoping review conducted in Chapter 2 revealed the emotional labor of teachers involved in navigating PBA contexts. Additionally, Chapter 3 provided a nuanced perspective on the policy intentions and challenges underpinning PBA and innovation in Italy. Furthermore, the empirical findings from Italian schools, as discussed in Chapter 4, highlighted the diverse range of understandings and adoptions of the policies in schools and the importance of contextual factors in mediating the enactment of accountability and innovation mandates. By integrating these insights, this dissertation contributes to a more comprehensive understanding of the complexities involved in the adoption and enactment processes of contemporary policies in education.

### **Theoretical insights: integrating cognitive approaches and institutional theories in policy adoption and enactment**

This dissertation adopts a combination of theoretical lens, drawing from new institutional theories and constructivist approaches to study the deployment and enactment of accountability and innovation policies in education. The combination of these approaches, namely, ideational factors on one side, and new institutionalist currents on the other, has enabled to delve into the non-linearity and complexity of the policy-making process (Kingdon, 1984), and has shed light on the role of multiple actors that shape the policy discourse on PBA. As seen, a multiplicity of state institutions (including evaluation, inspection or improvement agencies, research agencies, national leadership or teacher institutions, ministries of education and their staff) but also non-state actors such as teachers' unions shaped the policy process in multiple ways, highlighting the conflictual and ideological policy adoption process in the country and the legitimacy of different interest groups in shaping the policymaking in Italy (Galanti, 2023). At the same time, the relevance of "ideas" in explaining policy change (Hay, 2006; Schmidt, 2008) and wider social norms and priorities, such as those promoted by international organizations, were key in understanding education policy-making (cf. Grek & Ozga, 2010; Martens et al., 2010). Institutional theories have furthermore enabled to highlight the role of administrative traditions and bureaucratic apparatus as key loci of policymaking in Italy (cf. Verger et al. 2024).

Furthermore, new theories combining policy sociology and enactment research with institutional theory (Diehl & Golann's, 2023; Dulude & Milley, 2021) provided valuable tools to understand how school staff actively shape, filter, and adapt PBA to fit their school environments. The combination of these perspectives has been a key contribution to the field, enabling to go beyond policy compliance and implementation, and recognizing the key role of school actors' sense-making, active filtering and adaptation of policies in the daily routines as well as the enabling/constraining role of the local school environment in policy deployment (Braun et al., 2011). This heuristic approach has simultaneously shed light on the cognitive process of interpretation of the actors involved in the policy enactment process (Ball et al., 2012; Coburn, 2001), the variety of responses schools to external pressures, as well as examining the factors that influence these responses and, in turn, shape the external pressure itself (Elmore, 2005a).

### **Methodological insights**

The value of a multi-sited case study (Bartlett & Vavrus, 2014; 2017) to the study of multi-layered educational reforms is a key highlight of this dissertation. By moving between different policy levels (international, national and local) as well as focus/units of analysis (teachers, policy, schools) and analyzing how particular effects and mechanisms of a policy might differ according to diverse socio-economic school contexts within the same city, I had the opportunity to explore how the implementation of global education policies often deviates from a linear and top-down process. Rather, the interplay and dialogue between the different levels and sites of analysis I have gone through - which go from policy formulation to policy enactment - contributes to understanding the ways in which certain policy ideas are formulated, permeated and adopted in a specific context and time. Moreover, by moving between different policy levels through apparently separate research strands, a clearer picture of the expected and effectively generated policy outcomes of a specific reform model is provided. As seen, the assumptions behind the relationship between autonomy, innovation and accountability, that occurs at the policy discursive level, is central to understanding the policy enactment dynamics in schools, and how these are mediated by certain contextual conditions. Similarly, understanding the emotional effects of PBA on teachers in different countries, was important to understanding the attitudes, experiences and practices that school actors in Italy faced when negotiating between accountability and innovation mandates in schools. The transversal comparison of the PBA design in the country across time was also key to



understanding how PBA was adapted and re-configured over time. This highlights the need for increased attention to a transversal and multi-sited methodological approach in the field of education policies.

A realist evaluation approach has also enabled to unpack the policy intentions and elucidate the mechanisms and conditions under which PBA may promote specific consequences on school actors' experiences and practices. This approach shed light on the contexts, mechanisms, and outcomes of PBA and innovation in the Italian case, generating insights into the institutional context in which PBA is implemented, the mechanisms that in the Italian context might trigger certain outcomes, and the deployment and testing of the hypothesis of how PBA interventions are expected to work in the Italian educational system. In addition, the analysis of the "theory of change" components (i.e., assumptions, intervention(s), rationales/ outcomes, context and measurement of outcomes) as conceptualized by Reinholz and Andrews (2020) has proven to be effective to uncover the initial programmatic goals and intentions and the coherence between the policy theory of change and achievement of outcomes at the implementation level.

Moreover, the employment of a purposive sampling to map key stakeholders enabled to consider the perspective of a wide range of stakeholders (e.g., policymakers, research organizations, private foundations, labor unions) in the policymaking process (Chapter 3). The double sampling procedure used for selecting schools (Chapter 4) has enabled to go beyond the categorisation of schools that are usually considered in the Italian literature on the topic (e.g., including middle and low-income schools with diverse performance results).

### **PBA and innovation in a Southern-European setting: Learning about the Italian experience**

Drawing on the Italian case study, this research explores the relationship, challenges and opportunities of PBA on innovation in this national context, as well as the lessons that can be learned from its experience, especially in other similar contexts, with the premise that institutional legacies and politico-administrative regimes mediate the adoption of education reforms (e.g. Pollitt, 2007; Pollitt and Bouckaert, 2011; Verger et al., 2019) as well as their convergence and divergence patterns (cf. Verger et al., 2024).

On the one hand, the adoption and enactment of PBA linked to innovation in Italy, which is rooted in a Napoleonic tradition (Ongaro, 2008), highlights the challenge of assimilating certain innovations. Indeed, in alignment with Verger et al. (2019) and Kickert, (2007), the implementation of the school autonomy with accountability reform package in Italy, like in other Napoleonic states, was adopted with aimed to modernize, and enhance the quality and efficiency of the educational system, while also conforming to international standards and discourses on educational governance. Its implementation has been fragmented and irregular, resulting in incoherent policy frameworks, and meeting obstacles and resistance at the implementation level. Moreover, the awareness of interviewed actors that led and designed the PBA reforms about the contextual issues that would have undermined the successful realization of the policy goals, forces us to ask whether PBA in Italy has been adopted within a widely shared “theory of no-change”. Indeed, the interviewed stakeholders had clear awareness about the challenging conditions that are not conducive to the expected cultural changes in the educational realm, already at the policy design and adoption phase. They were intended to promote an evaluation culture but without expecting it to generate immediate results. In the Italian case, the PBA system, promoted by both center-left and center-right governments, was furthermore driven by globalizing ideas, such as learning achievement as an element of international competitiveness and the belief in the advantages of data-driven decision-making. From a critical perspective, this finding highlights the fertile ground and receptivity to the permeation of globalizing ideas in education (Ball, 1998) in this country, with the risk of not considering country-specific problems in the adoption of education reforms.

On the other hand, the findings of this dissertation also show how cumulative changes in PBA reforms adopted in the country (starting from the late 1990s) have generated a complex picture of how these policies have been selected, adopted and translated at the local level. Firstly, the involvement of various actors in the design and implementation of these reforms, from public-ministerial bodies, research institutes, professional associations, to private foundations, played a significant role. The interplay and conflicts between these diverse interest groups appear as crucial factors influencing the design and adoption of a PBA model in its most recent configuration. In addition, the incremental adoption of different policy instruments related to PBA over the last 20 years generated disparate opinions on key educational actors. More specifically, regardless initial resistance by teachers and unions towards external evaluations, and merit-based and ranking mechanisms more specifically, the dissertation shows how the integration of external and internal self-evaluation mechanisms, and equity, transparency and improvement motives underpinning such policy reforms, has generated heterogenous perceptions and uses of the PBA system by

school actors, ranging from extremely critical to fully supportive and in line with the policy expectations. In fact, at the policy level, the adoption of the improvement and internal evaluation perspective, as well as the political mediation to avoid the automatic publication of school results, was a response to the criticism of too much initial emphasis on external evaluation, especially deriving from labor unions. These features, together with the absence of any prize or merit-based or ranking consequences, were not contemplated in the initial piloted schemes of the PBA system. Such experiences of PBA in schools are also supported by the fact that in the Italian context, PBA reforms are strongly linked and associated, on the one hand, with pedagogical and organizational autonomy of schools, and on the other, with the push towards changing and innovative pedagogies and teaching practices, overcoming a traditional content-based and transmissive legacy. It can be derived that the widespread adoption of PBA in education in Italy is not a standalone reform package, but is always more interlinked to other educational mandates and imperatives, as well as representing different ideological orientations, which support its adoption, legitimacy and implementation dynamics. Moreover, in a context of conflictual and ideological policy solutions (Galanti, 2023), PBA policies have been constantly revisited and recalibrated from their initial phase and piloting until their most recent configurations, through conscious political choices taken to overcome prejudices and accommodate different stakeholder perspectives. This process of layering and accommodation has certainly played a key role in supporting the legitimation and assimilation of the policies in the country, but also contributed to making the policy framework of PBA policies more incoherent and difficult to implement (e.g., Capano & Lepori, 2024), influencing the organizational responses of schools and the perception and experiences of teachers and school leaders, who are at the forefront of their adoption in schools.

### **Understanding the role of discursive ideas and sense-making**

Drawing from discursive/constructivist institutionalism (Hay, 2006; Schmidt, 2008), the dissertation has highlighted the centrality of policy ideas in the definition of the policy agenda and in influencing political actors, who decide to frame them and use them to convince others (Cairney, 2012). Indeed, the various rationales put forward to justify the adoption of PBA reforms in the country, the influence of globalizing ideas, as well as the use of such rationales in the policy mediation and negotiation with labor unions, has contributed to shape the final configuration of PBA in the country. The findings also highlight how, in Italy, policy makers and key stakeholders involved in the policy design, have articulated together school autonomy, innovation and (external

and internal) school accountability as a powerful dispositif, framing, in a unique way, PBA as an innovative tool, in alignment with the discursive promotion of a “managerial recipe” in education (Serpieri et al., 2015).

Moreover, it is essential to consider the multiple, but even sometimes conflicting meanings provided by the key actors involved in the policymaking as well as the enactment process in schools with respect to PBA and innovation policies. The ambiguous and multifaceted meanings given to innovation concept, on the one side, and the often-conflicting meanings and goals given to PBA (and the external assessments in particular) on the other, has largely contributed to shaping the adaptations, acceptances and uses of these two policy mandates in the country. The ingrained ideas and beliefs on learning, evaluation and knowledge transmission are seen as undermining policy objectives and negatively influencing expected uses of the PBA policy instruments by school actors in the country. Also, the misunderstanding of the purpose and usefulness of the PBA system and test results by school actors is seen as a main challenge, which leads to undesired effects, resistance or superficial uses of the PBA instruments, without taking advantage of them to review and improve practices.

However, we should also underscore the key role of the negative idea and expectations that stakeholders hold on teachers in Italy, as opposed to the virtues of the external PBA system. The continuous and widespread negative view on teachers, who are blamed and viewed as incompetent by key educational actors, requiring external guidance and control, and incapable of taking advantage of the goals of the PBA system or the innovation mandate, might have played a key role in shaping teachers’ attitudes, emotions, experiences and responses to PBA and innovation policies in schools. Indeed, often simplistic conceptions of teacher motivation are highly challenging from an implementation point of view (Checchi & Mattei, 2021). In this sense, although associated with wider professional autonomy at the policy level, innovation in Italy appears to be in practice another artifact used to influence and control teachers’ work, as it is also the case in other southern European contexts (e.g., Quilabert et al., 2024).

## **Understanding the role of institutional contingencies and local school contexts**

From an institutionalist theory perspective, it is important to stress that policy ideas do not operate in a vacuum and are context dependent (DiMaggio & Powell, 1991; Meyer, 2008). Political or

economic contexts, together with institutional forces, create the conditions for the behavior of actors, as well as for the development, diffusion and translation of their ideas (Stone, 2012). The analysis conducted in this dissertation has revealed how, according to the key actors interviewed, the context of intervention is not entirely conducive to triggering the expected change mechanisms of PBA policies. It emerged clearly that the contextual aspects identified by key stakeholders (e.g., structural features including school buildings, teacher training and precariousness, the school actors' personal and professional characteristics, as well as cultural aspects and ingrained beliefs) hamper the expected change mechanisms, highlighting the discontinuous ground in which PBA policy dispositifs operate in the Italian context.

From a policy enactment perspective, the analysis revealed that there are several contextual intervening factors which shape school actors' logics of enactment, including socio-economic factors, administrative, market accountability and innovation pressures stemming from the external environment, but also material contexts, and professional ones such as school leadership and the strength of professional communities in schools. This finding is important to consider from a realist evaluation perspective, recognizing the importance of the context in which interventions are implemented, and identifying the way in which socio-economic or political environment exerts an influence on how policy interventions operate (Christensen & Molin, 1995) and how - and which - outcomes are finally achieved.

### **Limitations and future lines of inquiry**

While this dissertation has provided valuable insights into the deployment and effects of PBA and innovation policies in education, some limitations should be acknowledged and addressed in future studies. First, the focus on the Italian context should be taken into consideration when generalizing our findings to other educational contexts. Future research could explore PBA and its relationship to innovation across different national contexts to identify common trends and variations through comparative analysis. A critical discourse analysis of PBA and innovation policy ideas across different contexts is a potential and interesting avenue for research. Second, the reliance on qualitative methods may have constrained the ability to capture broader patterns and trends in policy outcomes. Future studies could employ mixed-methods approaches to provide a more comprehensive understanding of the effects of PBA on educational practice and outcomes in the country. Thirdly, further studies could use in-depth interviews and biographical interviews with

teachers to deepen the narratives and subjective experiences of PBA and innovation, considering the role emotions have in mediating policy enactment within different socio-economic contexts. Indeed, there is insufficient empirical research from the perspective of the sociology of emotions applied to the Italian accountability reform. Moreover, the low number of interviews conducted with school leaders in some middle-low and working-class schools in this dissertation (due to staff turnover, other school priorities, or unresponsiveness) has been accounted for in the analysis and interpretation of the results. Future research could more clearly focus on how this specific category of actors, working in unprivileged contexts, live and negotiate the policy mandates within their daily work, and the mediating role they play with teachers. Finally, in order to give voice to students, who are often underrepresented in research on PBA policies, or educational policies in general, a specific study that takes into account students' perspective and experiences with PBA and innovation might be interesting to deploy. For instance, adopting participatory qualitative approaches such as photovoice methods to assess students' needs (e.g., Wang & Burris, 1997), represent their educational experiences and translate them into actionable knowledge (e.g., Luescher et al., 2021).

## **Policy relevance**

This research has contributed to assessing the coherence between policy goals and outcomes of the Italian PBA system, which is particularly relevant from a policy perspective. By systematically evaluating public education policies (such as the effectiveness of the SNV in promoting innovation), policymakers can identify areas of success, as well as areas requiring improvement or adjustment, making informed decisions to strengthen and refine education policies.

Moving forward, by critically assessing the role of “context” (Ball et al., 2012), the research has attempted to uncover what works, for whom, where and in what contexts, nuancing the evaluation of the effectiveness of the policy reforms by considering the local contexts in which policies are enacted. This approach is essential to avoid a rigid and standardized analysis of policy impact, which only considers the effectiveness of the intervention, without contemplating the specific cultural, systemic or social contexts in which policies are being implemented. Incentivizing policymakers to understand how certain policies have disparate effects in different school contexts becomes crucial to the design of certain policies or concrete interventions aligned with equity and quality goals and the conditions and needs of specific contexts.

Another key policy implication derives from the need to concretely support the conditions for educational innovation, especially under autonomy and PBA education contexts. The innovation mandate, when rarely linked to inclusion or investment in resources in highly complex centers, may imply a risk in equity of the educational system (e.g., Quilabert et al., 2024). This includes supporting teacher autonomy with adequate training, generating opportunities for staff collaboration and exchange, and spaces for co-teaching, particularly in larger schools with greater organizational needs or in disadvantaged schools experiencing frequent changes in management teams and internal rotations. Moreover, the “limited” financial and managerial school autonomy in Italy, expressed by all educational actors interviewed, highlights the need to accompany school autonomy with increased financial and managerial resources and personnel, in order to balance the external accountability demands and foster conditions conducive to change and equity, especially in highly complex school contexts.

Full attention should be also paid to recognise and support teachers’ work in demanding school contexts, in order to avoid unintended consequences of accountability and innovation approaches (e.g., Quilabert et al., 2024). This implies providing teachers with increased trust and support that they require to be able to promote innovative processes and take full advantage of the PBA instruments to promote improvement actions that align with the initial policy goals. Instilling policymakers with a positive view of these educational professionals is also essential to counterbalance the supposed inefficacy of the education system in Italy (e.g., Barzanò and Grimaldi, 2012), as well as the increasing role of new private actors in the public education system (Peruzzo et al., 2022; Sorensen et al., 2021). The full recognition of teachers’s competency, along with the adequate support and training, avoiding blaming attitudes, is essential to this end.

From the study it emerges that the emotional effects of PBA are experienced by teachers located in contexts that differ significantly from one another, suggesting that some emotional experiences might constitute an integral part of accountability reforms. However, also within the same educational system (i.e., Italy), a constant pressure to innovate and improve, coupled with the emotional impact of such mandates, is universally felt by school actors, irrespective of the schools’ socio-economic characteristics. Thus, policies should take this aspect seriously into account, since teachers’ wellbeing, professional satisfaction and motivation influence the adoption of the creative and innovative practices by teachers, which are expected outcomes in the first place. Moreover, the dissertation shows that factors at the macro level (system), such as stable contracts or

professional development, meso-level (school), including leadership and staff cohesion, and micro-level (individual) can intensify or weaken the emotional effects experienced by teachers. Leveraging on such positive variables could be crucial for policymakers to enhance acceptability and teachers' positive experiences with the reforms.

To conclude, the dissertation highlights the need for continuing research on the complex relationship between PBA reforms, teacher experiences, emotions and innovation, and the benefits of bringing together different theoretical perspectives and methodologies on this topic, not only to provide a more comprehensive understanding, but also to make fruitful collaborative research efforts possible.

In light of the findings, policymakers, educators, and researchers should adopt a policy context-sensitive, interdisciplinary and integrated approach to PBA education reforms. Efforts should be made to cultivate a more balanced approach of PBA and innovation that harnesses both frameworks. This may involve reimagining accountability mechanisms as “intelligent” forms of accountabilities that allow for greater teamwork, distributed leadership, and professional learning communities among educators, without excessively emphasizing performance indicators, external inspections and achievement data (Tolo et al., 2020), or that at same time refuse trust in educational evaluations to make informed and independent judgment of results, and communicate those judgements intelligibly, without damaging educational objectives (O’ Neill, 2013). Horizontal accountabilities that address peers, equals and stakeholders outside of the hierarchical relationship between central government and executive agency (Schillemans, 2008), or multiple accountability frameworks that combine different forms or logics of accountability (Hooges et al., 2012; Kim & Yun, 2019), such as professional-based or process-based accountability in addition to a control-based or test-based accountability (see Kim & Yun, 2019), may also provide alternative paradigms and approaches to accountability to explore. Overall, there is a need to reconceptualize accountability measures to better align with the goals of equity, social justice, teacher wellbeing and professionalism in specific contexts. Finally, policymakers should consider the diverse perspectives and experiences of stakeholders in the design and implementation of PBA policies, ensuring that they are responsive and aligned with the perceptions and needs of school actors since, at the end of the day, they are the ones on the front line of providing education.



## References

- Ball SJ (1998) Big policies/small world: An introduction to International perspectives in education policy. *Comparative Education* 34(2): 119–130.
- Ball, S., Maguire, M., & Braun, A. (2012). *How schools do policy: Policy enactments in secondary Schools*. Routledge.
- Bartlett, L., & Vavrus, F. (2017). *Rethinking Case Study Research. A Comparative Approach*, Routledge
- 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.
- Cairney P (2012) *Understanding Public Policy: Theories and Issues*. New York: Palgrave Macmillan.
- Cecchi, D., & Mattei, P. (2021). Merit pay for schoolteachers in Italy, 2015–2016: A new regime of education accountability? *Comparative Education Review*, 65(3), 445–466.
- Christensen S and Molin J (1995) Origin and transformation of organizations: Institutional Analysis of the Danish red Cross. In: Scott WR and Christensen S (eds) *The Institutional Construction of Organizations*. Thousand Oaks, CA: SAGE Publications, pp.67–90.
- Coburn, C. E. (2001). Collective sensemaking about reading: How teachers mediate reading policy in their professional communities. *Educational Evaluation and Policy Analysis*, 23(2), 145–170.
- Diehl, D. K., & Golann, J. W. (2023). An integrated framework for studying how schools respond to external pressures. *Educational Researcher*, 52(5), 296–305.
- DiMaggio PJ and Powell WW (1991) Introduction. In: Powell WW and DiMaggio PJ (eds) *The New Institutionalism in Organizational Analysis*. Chicago, IL: University of Chicago Press, pp.1–38.
- Dulude, E., & Milley, P. (2021). Institutional complexity and multiple accountability tensions: A conceptual framework for analyzing school leaders' interpretation of competing demands. *Policy Futures in Education*, 19(1), 84–96.
- Elmore, R. (2005a). The accountable leadership. *The Educational Forum*, 69(2), 134–142.
- Galanti, M. (2023). Expert legitimacy and competing legitimation in Italian school reforms. *Policy and Society*, 42.

- Grek S and Ozga J (2010) Re-inventing public education: The new role of knowledge in education policy making. *Public Policy And Administration* 25(3): 271–288.
- Hay C (2006) Constructivist institutionalism. In: Rhodes R, Binder S and Rockman B (eds) *The Oxford Handbook of Political Institutions*. Oxford: Oxford University Press, pp. 56–74.
- Hooge, E., T. Burns and H. Wilkoszewski (2012), “*Looking Beyond the Numbers: Stakeholders and Multiple School Accountability*”, OECD Education Working Papers, No. 85, OECD Publishing.
- Kim, T. & Yun, J. T. (2019). Logics of accountability: Cross-national patterns in schoollevel controls. *Education Policy Analysis Archives*, 27(119).
- Kingdon JW (1984) *Agendas, Alternatives, and Public Policies*. Boston, MA: Little, Brown
- Kickert, W. (2007). Public management reforms in countries with a Napoleonic state model. In C. Pollitt, S. Van Thiel, & V. Homburg (Eds.), *New public management in Europe: Adaptation and alternatives* (pp. 26–51). Palgrave/Macmillan.
- Luescher, T. M., Fadiji, A. W., Morwe, K., & Letsoalo, T. S. (2021). Rapid Photovoice as a Close-Up, Emancipatory Methodology in Student Experience Research: The Case of the Student Movement Violence and Wellbeing Study. *International Journal of Qualitative Methods*, 20.
- Martens K, Nagel A-K, Windzio M, et al. (2010) *Transformation of Education Policy*. Basingstoke: Palgrave Macmillan.
- Meyer JW (2008) Reflections on institutional theories of organizations. In: Greenwood R, Oliver C, Sahlin K, et al. (eds) *The Sage Handbook of Organizational Institutionalism*. Thousand Oaks, CA: SAGE, pp.790–812.
- O’Neill O. (2013) Intelligent accountability in education, *Oxford Review of Education*, 39:1, 4-16
- Ongaro E. (2008), "Introduction: the reform of public management in France, Greece, Italy, Portugal and Spain", *International Journal of Public Sector Management*, Vol. 21 Iss: 2 pp. 101 - 117
- Pedró F. (2023), Chapter 10: Where is the school going? International trends in educational innovation, *Sociology, Social Policy and Education*, pp 147–164.
- Peruzzo F, Grimaldi E, Arienzo A, et al. (2022) New public management reforms and industrial relations in the Italian education system. A cultural political economy approach. *Journal of Educational Administration and History*. 1–19.

- Pollitt C (2007) Convergence or divergence: What has been happening in Europe? In: Pollitt C, Van Thiel S and Homburg V (eds) *New Public Management in Europe*. London: Palgrave Macmillan UK, pp.10–25.
- Pollitt C and Bouckaert G (2011) *Public Management Reform. A Comparative Analysis—New Public Management, Governance, and the Neo-Weberian State*, 3rd edn. Oxford: Oxford University Press.
- Quilabert, E., Verger, A., Moschetti, M. C., Ferrer-Esteban, G., & Pagès, M. (2024). The obstacle race to educational improvement: governance, policies, and practices in disadvantaged schools. *British Journal of Sociology of Education*, 1–23.
- Reinholz DL and Andrews TC (2020) Change theory and theory of change: what's the difference anyway? *International Journal of Stem Education* 7(1): 2–12.
- Serpieri R, Grimaldi E and Vatrella S (2015) School evaluation and consultancy in Italy. Sliding doors towards privatisation? *Journal of Educational Administration and History* 47(3): 294–314.
- Schillemans, T. (2008), Accountability in the Shadow of Hierarchy: The Horizontal Accountability of Agencies. *Public Organ Rev* 8, 175–194.
- Schmidt VA (2008) Discursive institutionalism: The explanatory power of ideas and discourse. *Annual Review of Political Science* 11: 303–326.
- Sorensen TB, Grimaldi E and Gajderowicz T (2021) *Rhetoric or Game Changer: Social Dialogue and Industrial Relations in Education Midst EU Governance and Privatization in Europe*. Brussels: European Trade Union Committee for Education (ETUCE-CSEE).
- Stone D (2012) *Transfer and translation of policy*. *Policy Studies* 33(6): 483–499.
- Tolo, A., Lillejord, S., Flórez Petour, M.T. et al. Intelligent accountability in schools: A study of how school leaders work with the implementation of assessment for learning. *J Educ Change* 21, 59–82 (2020).
- Verger A, Fontdevila C and Parcerisa L (2019) Reforming governance through policy instruments: How and to what extent standards, tests and accountability in education spread worldwide. *Discourse Studies in the Cultural Politics of Education* 40(2): 248–270.
- Verger, A., Ferrer-Esteban, G., & Fontdevila, C. (2024). Not everything that moves must converge: evidence from global policy and practice on performance-based accountability, In Edwards Jr, D. B., Verger, A., McKenzie, M., & Takayama (eds), *Researching Global Education Policy: Diverse Approaches to Policy Movement*, Bristol: Policy Press (pp. 39–69).
- Wang C, Burris MA. Photovoice: concept, methodology, and use for participatory needs assessment. *Health Educ Behav*. 1997 Jun;24(3):369-87.