

A “three-legged model”: (De)constructing school autonomy, accountability, and innovation in the Italian National Evaluation System

European Educational Research Journal

1–26

© The Author(s) 2023

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/14749041221148280

journals.sagepub.com/home/eer**Laura Mentini**

Universitat Autònoma de Barcelona, Spain

Antonina Levatino 

Universitat Autònoma de Barcelona, Spain

Abstract

The widespread adoption of school autonomy with accountability reforms in education has generated debate regarding the relationship between autonomy, innovation and accountability. While at the policy design level, these three elements are highly related, several authors highlight the contradictions among them. By analyzing key documents and interviews, this paper aims to identify the program ontology behind the current Italian National Evaluation System (SNV), with a focus on the way in which autonomy, accountability and innovation have been conceptualized and linked together. The paper also aims to explore whether pitfalls and/or tensions exist that might hamper the achievement of the SNV goals. The findings highlight the peculiarities of the Italian autonomy with accountability system, which has resulted from the involvement of different stakeholders in the design and implementation of the reforms. The findings also reveal contradictions regarding some of its premises. Various rationales (improvement, efficiency, equity and transparency) emerge that seem to have acted as drivers of the reforms, however, the influence of globalizing discourses on international competition and the benefits of datafication also appears significant. A number of contextual aspects are finally considered which hamper the expected change mechanisms, highlighting the discontinuous ground in which such policy *dispositifs* operate.

Keywords

Test-based accountability, innovation, school autonomy, New Public Management, Italy, reform ontology, policy discourses

Corresponding author:

Laura Mentini, Department of Sociology, Universitat Autònoma de Barcelona, Barcelona, Spain.

Email: laura.mentini@uab.cat

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,¹ 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). 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 welfare 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 basis² 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).

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 broad-ranging 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 1.

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 2 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 Foundation³) 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 1. 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

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 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,

Table 2. 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

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.

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 scholar1) and “guide schools in improving their curricular and pedagogical offer, and way of planning” (Member of Ministry2). 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 foundation1). 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 questions⁴ 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 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 Ministry2). 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 foundation2), 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 Ministry1). 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 legitimization 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 “levelling 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 perceive 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 programing 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,⁵ 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).

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, debureaucratize 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 incompetency 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.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the European Research Council under Grant 680172 (REFORMED project); Doctoral Fellowship from La Caixa Foundation (ID100010434), code LCF/BQ/DR19/11740004.

ORCID iD

Antonina Levatino  <https://orcid.org/0000-0001-7245-3592>

Notes

1. 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.
2. 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.
3. 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.”
4. The majority of INVALSI questions are closed answers, even though there are some which require a short or articulated textual answer (Robasto, 2017).
5. The implementation of INVALSI tests are part of teachers’ service obligations (“obblighi di servizio”).

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” valutative 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. DOI: 10.1080/00220620.2022.2094350.
- 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.

Author biographies

Laura Mentini is a PhD student and “La Caixa” fellow at the Sociology Department of the Universitat Autònoma de Barcelona (UAB) (E: laura.mentini@uab.cat). Her thesis deals with education accountability and innovation policies in Italy. She is a member of the ‘REFORMED’ research project supported by the European Research Council.

Antonina Levatino holds a PhD in Social and Political Sciences and is currently a researcher at the Department of Sociology of the Universitat Autònoma de Barcelona. Her research focuses, on the one hand, on internationalization of higher education and, on the other hand, on accountability and autonomy reforms in education.