

How to cite:

Verger, A., & Parcerisa, L. (2017). A Difficult Relationship: Accountability Policies and Teachers. International Evidence and Key Premises For Future Research. In M. Akiba & G. LeTendre (eds.). International Handbook of Teacher Quality and Policy (pp.241-254). New York: Routledge.

15 A Difficult Relationship

Accountability Policies and Teachers: International Evidence and Key Premises for Future Research

Antoni Verger and Lluís Parcerisa

Universitat Autònoma de Barcelona

Accountability is becoming an increasingly central norm in the global education agenda. Despite the fact that accountability is far from being new in the educational sector, the forms and the main aims of accountability have changed substantially in the last years. Traditionally, political and bureaucratic forms of accountability have aimed to increase the democratic control of education, especially in industrialized countries. However, what we are witnessing more recently is that accountability is being conceived as a central policy solution to the most important problems and challenges educational systems face, both in industrialized and developing societies. Specifically, accountability in education is increasingly perceived as a source of efficiency (i.e., a way to promote better alignment between governmental aspirations and the school purposes), academic excellence (improved performance on the part of schools, teachers, and students), and equity (a way to guarantee that all students reach a minimum level of competence in core subjects).

The most influential international organizations in the education policy field have fueled such high expectations with proposed accountability solutions. Accountability policies are absolutely pivotal in the policy recommendations of the Organisation for Economic Co-operation and Development (OECD) reports on education, including the influential Programme for International Student Assessment (PISA). According to a recent study, 29 OECD country representatives (out of 37) admitted that PISA/OECD recommendations on accountability have influenced policy decisions on the matter at the national level (Breakspear, 2012). For the World Bank, accountability at the school level is a necessary condition for improving teacher quality and learning “because they align teacher and parent incentives” (World Bank, 2015, p. 2). UNESCO has also placed accountability at the center of its Education 2030 Action Framework, and will dedicate the 2017 edition of its influential Global Education Monitoring Report to this theme. For UNESCO, the introduction of accountability mechanisms in national education systems is a necessary condition for countries to meet the internationally agreed sustainable development goals on education (UNESCO, 2015).

In policy terms, accountability is a concept with multiple meanings that covers a broad range of policy options and models (including political, legal, bureaucratic, and market forms of accountability). Nonetheless, the model of accountability that is currently gaining more international attention is strongly grounded in the administrative (or external) evaluation of learning outcomes, and focuses on schools and teachers as the actors that give the account. This type of accountability is usually known as managerial or testing-based accountability (Kamens & Benavot, 2011; Tobin et al., 2015).

The more these procedures are being internationally adopted, the more accountability policies—and, in particular, testing-based accountability policies—become the focus of academic research. Nonetheless, despite the fact that research on the matter is increasingly available, there is still an insufficient level of understanding of under what circumstances and

through what particular mechanisms which types of accountability policies may achieve the expected results. In fact, as this chapter shows, evidence on the effects of accountability policies on learning outcomes, but also on other important aspects of education such as teachers' autonomy, schools' organization, and pedagogical options, is still inconclusive and shows different (and even contradictory) effects in different places.

This chapter reviews international evidence on the effects of accountability policies in education, with a particular focus on teachers' work and behavior. Specifically, on the basis of existing evidence, we aim to show how the different components of accountability schemes (namely their policy design, enactment processes, contextual contingencies, and impact) interact in complex and multiple ways.

With this goal in mind, the chapter is organized as follows. In the first section, we present the main models of accountability in education, and explain why testing-based accountability is becoming the predominant model internationally. In the second section we show how testing-based accountability can crystallize in very different policy designs, and why decisions on design matters are so important to understanding the trajectory of accountability reforms. Third, we reflect on the different impact dimensions of accountability policies in education, including unexpected or non-desired results at the school and teaching levels. Fourth, we focus on the enactment dynamics of accountability policies—including how teachers receive, resist, and/or transform accountability through their daily practices. Fifth, we look at how and to what extent the institutional and socio-economic contexts where accountability policies are enacted also shape the processes these policies go through.

In the final section, on the basis of our review, we derive key premises and directions for future research. All in all, the chapter shows that teachers interact with accountability schemes in different ways: as enactors, as conditions, and as results. To a great extent, the way teachers are conceived and treated within accountability schemes, but also the way teachers perceive, interact, and perform within such schemes, is key to understanding the differential and, at points, contradictory outcomes of accountability in education.

Methodologically, the chapter is based on a *scoping review* approach (Alegre, 2015). This literature review methodology, in contrast to systematic literature reviews, does not depart from a very specific research question or aim at testing a particular theory. The scoping review approach is particularly well equipped to identify the main areas of agreement and dissent within a particular field of research (in our case accountability in education), as well as the main gaps in the existing corpus of literature in such a field, in a relatively short period. The literature search was done in the SCOPUS and Web of Science data-bases. In total, 150 documents were reviewed, although for the purpose of the elaboration of this chapter, not all of them have been used.

DIFFERENT APPROACHES TO ACCOUNTABILITY IN EDUCATION: THE EMERGENCE OF TESTING-BASED ACCOUNTABILITY

Generally speaking, accountability refers to the processes, mechanisms, and/or instruments that make institutions (but also individuals and groups within institutions) meet their obligations and become more responsive to their particular audiences (Bovens, 2007; Hatch, 2013). Both obligation and responsiveness are inherent characteristics in accountability systems. According to Bovens (2007):

Accountability is a relationship between an actor and a forum, in which the actor has an obligation to explain and to justify his or her conduct, the forum can pose questions and pass judgement, and the actor may face consequences.

(p. 450)

Nonetheless, despite these very general features, accountability is a polysemic concept that covers a broad range of policy approaches and typologies. The most traditional typologies of accountability in education are the political, legal, and bureaucratic. *Political accountability* means that politicians, legislators, and/or school board members “must regularly stand for election and answer for their decisions” (Darling-Hammond, 2004, p. 1050) in front of civil society, Parliament, and/or the media, whereas *legal accountability* means that educational actors (including teachers, principals, schools, education management organizations, policy-makers, and so on) are expected to operate in accord with legislation; accordingly, any citizen can bring to the courts complaints about the potential violation of laws by any of these actors (Darling-Hammond, 2004). In *bureaucratic accountability* systems, public authorities promulgate rules and procedures on how education should be delivered, and school inspectors (or local authorities) try to ensure that schooling takes place according to such rules and procedures (Darling-Hammond, 2004).

Currently, there is a greater emphasis on other forms of accountability in education, including professional, market, participative, and managerial accountability. *Professional accountability* means that teachers are accountable for adherence to professional standards and give the account to their peers—or to professional organizations (Anderson, 2005); *market accountability* means accountability to the consumer (i.e., families) and is promoted through pro-school choice and pro-school competition policies (West et al., 2011); *participative (or social) accountability* means that schools and their teachers are accountable to families not via the choice and exit mechanisms, but via voice and dialogue within school governing bodies (West et al., 2011); finally, *managerial accountability* “includes systematic efforts to create more goal-oriented, efficient, and effective schools by introducing more rational administrative procedures” (Leithwood & Earl, 2000, p. 14). The most distinctive feature of managerial accountability is its clear focus on results (usually learning outcomes), instead of processes and inputs. The enactment of this approach to accountability involves the generation of data through large-scale standardized evaluation instruments. Because of this reason, this approach to accountability is also known as performative accountability (Ranson, 2003), results-driven accountability (Anderson, 2005), or test-based accountability (Hamilton et al., 2007).

The latter model, managerial or testing-based accountability, has recently become more predominant and globally spread. It has gained international popularity and attention for different reasons. To start with, improving learning outcomes (and not only inputs or access figures) has become the main focus of education reform agendas in both the north and the south. Accordingly, measuring learning outcomes is a necessary tool to test whether reforms achieve the expected results. In this terrain, the No Child Left Behind reform adopted in the US in 2001 has become one of the most emblematic initiatives internationally. Nonetheless, beyond national legislation, international standardized tests such as PISA have become a key driver of testing-based accountability reforms. PISA has not only contributed to introduce competitive pressure to countries for better learning outcomes, but has also become an instrumental device to transfer the technology that allows testing learning skills at the national level. The emergence of a school testing industry globally has also pushed in a similar direction (see Hogan et al., 2016).

The expansion of the test-based accountability regime is also related to the consolidation of new public management (NPM) as a public sector reform paradigm. NPM has contributed to make education reformers more inclined to promote managerial governance styles and educational services oriented toward the achievement of tangible and measurable results (Kalimullah, Ashraf, & Ashaduzzaman, 2012). NPM, but also related global education reforms, such as common core standards or school autonomy, put testing-based accountability mechanisms at the center of education systems since learning measures and benchmarks become the most important tool in the management of education systems and in the promotion of school improvement.

TESTING-BASED ACCOUNTABILITIES: REFLECTING ON MULTIPLE DESIGN OPTIONS

Independently of the specific general approach adopted, policy design variables need to be considered in unpacking the complexity and internal diversity of accountability typologies. Specifically, the same accountability approach can adopt very different forms according to who is expected to provide the account, to whom is the account owed, what is to be accounted for, and what are the consequences of providing an account (Leithwood & Earl, 2000). In this section, we show how these variables (the who, the what, to whom, etc.) can interact and can be differently combined within the model of accountability that is becoming increasingly predominant nowadays: test-based accountability.

First, in relation to who is expected to provide the account in testing-based accountability regimes, the locus could be on individual teachers, on groups of teachers (for instance, teachers from the same grade), on the principal, on the whole school, or on networks of schools. This is an important variable in the sense that individual approaches are expected to undermine cooperation and collegial work between teachers, whereas collective approaches do not necessarily have such an effect (Jones & Egley, 2004).

Second, the account could be owed to a broad range of actors including inspection services, an external evaluation agency, families, professional bodies, or a combination of these agents. Nonetheless, in learning-based accountability, the participation of an education evaluation agency at some point of the accountability process, whether governmental or not, tends to be indispensable. Here, an important variable is the level of publicness of the school evaluations, (i.e. whether the results of the evaluations are made public, and how; for instance, taking into account school socio-economic composition or not). The publication of school rankings is a controversial accountability measure that, according to some authors, could improve academic performance (Boarini & Lüdemann, 2009), but according to others could undermine the potential of cooperation dynamics between schools, or contribute to stigmatizing underperforming schools (Jones & Egley, 2004).

Third, in testing-based accountability regimes, students' learning outcomes are, by definition, the main focus of the accountability system. However, the learning outcomes component could be combined with evaluations of other components, such as drop-outs, graduation rates, school resources and facilities, and/or educational processes (through, for instance, observation in class, revision of teachers' portfolios, etc.). Accordingly, accountability systems could be more or less comprehensive in terms of the education components they cover. More comprehensive accountability designs are expected to prevent schools focusing on specific types of results or parts of the curriculum because they contemplate a broader range of indicators, but could consume more resources and time from teachers and public authorities.

Last, but not least, the consequences of the evaluations represent a very important aspect in accountability frameworks. Here, the design options include whether accountability is high stakes and the nature of the consequences (economic incentives, teachers' development programs, and sanctions, such as the intervention of the school, school closure, or its conversion into a charter school, etc.). At the same time, economic incentives can be given at the school or individual level (in the form of salary bonuses). The amount of the incentive—whether more symbolic or more substantive—can also make a difference in conditioning the school or teachers' behavior (Escardíbul, 2015). In general, it is well documented that high-stakes evaluations put more pressure on schools than other forms of evaluation, but this is not necessarily positive. Excessive pressure without sufficient pedagogic support could promote the type of undesired behaviors mentioned above, which we explain further in the next section (Ohemeng & McCall-Thomas, 2013).

In a nutshell, the reviewed policy options show that it is not appropriate to talk about accountability in education as a concrete policy program. Scrutinizing the specific design characteristics of accountability schemes is a first and necessary step to promote more rigorous evaluations of the effects of accountability. In this section we have given some examples of how different design options of testing-based accountability might lead to distinct implications. Next, we provide a more in-depth exploration of the multiple effects of accountability in terms of school organization, teachers' professionalism, educational processes, and students' learning outcomes.

THE MULTIPLE EFFECTS OF ACCOUNTABILITY POLICIES IN EDUCATION AND TEACHERS' WORK

Most countries currently adopt testing-based accountability systems as a way to promote academic excellence and the improvement of learning outcomes. However, accountability systems also have effects in other educational dimensions. Some of these effects are desired and contemplated in the design of the policy, whereas others are neither desired nor expected.

The Mixed Nature of Accountability Effects

In relation to impact on learning outcomes, existing research on accountability in education shows quite contradicting results. Several researchers show that there is a *positive relationship* between accountability policies and students' academic results, although many of them alert that, for different reasons, this relationship can be rather weak or mixed. For instance, in research conducted in Florida, Chiang (2009) concludes that the threat of sanctions on low-performing schools generated an improvement in the math results of elementary students that persists through early years in middle school, although the author notes that the evidence on improving reading is not particularly robust. On their part, Roderick et al. (2008), on the basis of administrative data of Chicago public schools students, find a very positive effect in two degrees concerning basic skills (6th and 8th), although this was not the case of students in the 3rd grade.

Other pieces of research are even more explicit about the *mixed* nature of the results found. This is the case of Boarini and Lüdemann (2009), who conclude that although some aspects of the accountability system (e.g. external standardized national exams) may be associated with improved student outcomes, other aspects (such as publishing school results) have small effects on students' results, or did not have any effect. On his part, Ladd (1999) finds positive and robust effects of accountability measures in Dallas (Texas) for white and Hispanic students in the seventh grade, while black students obtain less positive results. Reback et al. (2014) examined the effects of No Child Left Behind (NCLB) on the school staff behavior and students' achievement. Both positive and neutral effects were found in relation to academic performance and enjoyment in the learning process. Also focusing on the effects of NCLB on students' score in national standardized tests, Dee and Jacob (2011) find a statistically significant positive effect on the results of 4th-grade students in math, but no evidence of improvement in reading.

Finally, another group of scholars has found that accountability reforms have negative effects in learning outcomes, especially from an equity perspective. For example, Andersen (2008) finds that the implementation of a performance management reform in Danish schools has had negative effects on the performance of low socio-economic status students. And Ryan (2004) notes that the test-based accountability regime of NCLB creates incentives that work

against the improvement of low social-background students' performance. Similarly, on the basis of evidence from California, Powers (2003) show that test-based accountability encourages school segregation and early school leaving among poor students and students belonging to ethnic minorities. This type of accountability can also create incentives for governments to reduce academic standards and discourage the best teachers to work in schools with high rates of socially vulnerable populations or with poor performance, which is an issue we take up again in the final section of this chapter.

Beyond learning outcomes, other important interactions that are also contemplated in the reviewed literature (although sparsely) are, on one hand, the relationship between schools and, on the other, teachers' identity and professionalism. According to existing research, accountability policies can trigger both collaboration and competition between schools. The school context and the role of the principal seem to influence whether the schools' behavior becomes more cooperative or competitive within accountability regimes. This theme is still quite underexplored but it is an important theme given the potential benefits that cooperative dynamics can have for educational systems (Keddie, 2015; Sahlberg, 2010).

Regarding the professional identity and professional autonomy of teachers, existing research—usually of an ethnographic character—notes that performance-based accountability policies have more cons than benefits. These policies seem to constrain teachers' professional autonomy, generate some levels of stress, or erode social relationships (see Day, 2002; Jeffrey & Woods, 1996; Maxcy, 2009). Despite the obvious implications of these types of effects in the levels of teacher motivation, there is also a body of literature that reports that accountability could have positive effects on teacher motivation (Finnigan, 2010).

Undesired and Non-Expected Outcomes

More and more research documents the non-desired effects of accountability policies in education, especially high-stakes accountability or pro-school competition accountability (for instance, the publication of the schools' results). In certain contexts, these types of accountability can encourage so-called opportunistic behavior on the part of schools, teachers, and principals. The opportunistic behavior concept refers to a wide range of practices, including cream-skimming (Jennings, 2010), teaching to the test (Ohemeng & McCall-Thomas, 2013), educational triage (Booher-Jennings, 2005; Perryman et al., 2011; Reback, 2008), and test cheating (Jacob & Levitt, 2003) that educational actors adopt to maximize their results in the tests. These behaviors usually emerge when schools and teachers are subjected to higher levels of external pressure to achieve better educational outcomes, especially because of the threat of sanctions these systems involve in case of underperformance.

One type of opportunistic behavior identified by the literature is the so-called educational triage, i.e. “‘triaging out’ students well below and well above grade level and ‘triaging in’ students close to grade level” (Ladd & Lauen, 2010, p. 429). Ladd and Lauen (2010) analyze the distributional effects of policies based on growth-accountability programs and status programs using student-level panel data from North Carolina, and find no evidence of the existence of triaging out of the low-achieving students from schools. The results contrast with the findings of another study conducted in Texas, which shows that teachers respond to incentives to improve scores focusing on the student group that is closest to the approval rate, at the expense of other groups of students (Booher-Jennings, 2005).

Ohemeng and McCall-Thomas (2013) performed a study on undesirable behavior generated by standardized testing in Ontario. They conducted interviews with various educational stakeholders to understand the pressures and incentives emanating from a standardized testing system. Many of the interviewed teachers argue that the pressure to get good results in

the tests led them to adopt teach-to-the-test practices and to focus on those areas of knowledge that would be tested (Ohemeng & McCall-Thomas, 2013). Similarly, Jones and Egley (2004) found that one of the most frequent complaints of teachers (23.3 percent) in relation to the system of high-stakes testing was the need to spend much of their time preparing their students for the test—i.e. teaching skills and content similar to those included in the test. These findings are consistent with other studies (see Au, 2007; Hamilton et al., 2007; Linn, 2000; Menken, 2006; Volante, 2004). These practices are an object of concern because they result in a reduction of the curriculum and undermine the idea of teaching and learning as a comprehensive process.

POLICY ADOPTION AND ENACTMENT

Promoting new forms of accountability in education is usually a contentious policy option. The process behind the adoption and design of accountability systems conditions the reception and enactment of these systems, as well as their final results.

According to McDermott (2007), in educational reforms there is usually a “gap” between the objectives or aspirations of policy-makers and the practices carried out in local contexts. The fact that an educational policy has a good design (or good intentions) is a necessary but not sufficient condition for reform success. In general, the perception and adherence of key educational stakeholders such as principals and teachers to the policy in question is a key factor to understand its correct implementation. In the case of accountability reforms, this is an especially relevant premise since these reforms often generate controversy and differential responses from teachers, including negotiation strategies and overt and covert resistance.

According to Ball et al. (2012), a narrow conception of “policy,” which omits the political processes that take place within schools, is quite frequent in education policy literature. Policy texts are not implemented mechanically, but are interpreted and usually transformed by actors operating at a micro-institutional level. To overcome the limitations of the concept of policy implementation, scholars like Ball et al. (2012) suggested using the concept of “policy enactment,” which serves to capture and understand the contingent and creative process of decoding and recoding through which political programs are put into action in schools.

Therefore, to have a comprehensive and global view of accountability reforms and to better understand their effects, it is necessary to open the black box that exists between the policy design and the results, and to analyze those intermediate steps in which the policy texts are received, negotiated, and translated into concrete practices.

TEACHERS AS ENACTORS OF ACCOUNTABILITY REFORMS

The compliance of educational actors is a necessary condition to ensure that any educational policy is successfully deployed. Thus, to guarantee that accountability reforms will have the compliance of teachers (although something similar could be said about principals), it is necessary that they feel that the reform will have a positive impact and will be meaningful for their professional development, and for school improvement purposes.

In a comparative study carried out in Europe, Müller and Hernández (2010) explored the perceptions of teachers on accountability systems. This research used a mixed-methods design, which included a survey with a sample of 1,100 teachers per country (in Finland, Sweden, and Ireland), 44 life-history interviews of 22 teachers, observations, and focus group discussions. These scholars noted that Europe is dominated by skepticism and disbelief of teachers regarding accountability and, especially, on performance-based accountability systems.

At the same time, they pointed out the risk that an accountability system can become an “empty mechanism” that produces a negative type of pressure that does not motivate teachers. To avoid accountability becoming superfluous or a negative influence on school improvement, these scholars suggested that accountability mechanisms should aim to strengthen teacher professionalism and respond to the real needs of the educational community (Müller & Hernández, 2010).

Similarly, in the context of the state of Florida, scholars like Jones and Egley (2004) explored teachers’ perceptions about a high-stakes accountability system through a survey of 708 teachers. Their analysis shows that most teachers (79.9 percent) believed that accountability policies implemented in Florida were not going in the right direction. According to these scholars, these perceptions have a negative impact on teachers’ responses against the accountability reform, and condemned such reform to failure.

Learning-based accountability policies tend to alter the existing power relations in school governance by centralizing power in decision-making (Maxcy, 2009). These policies provide higher levels of power to the administration or to principals, at the same time that they undermine the ability of teachers to have more influence in decision-making processes. Nonetheless, when teachers perceive that policies disempower them or undermine their professional autonomy (e.g. on curricular decisions), they generally adopt strategies of resistance or negotiation against the implementation of reforms. For example, in South Africa, teachers boycotted the implementation of the new performance-based accountability system through a wide range of individual and collective actions such as being absent in meetings, refusing to send their records, or not performing other tasks stipulated in the program (Mosoge & Pilane, 2014).

Overall, teachers’ responses to accountability policies should not be taken for granted, but seen as a cornerstone of research about the effects of these policies. As we have seen, the responses of teachers to the adoption of accountability systems range from conflict and resistance to consent, and include intermediate responses, such as negotiation strategies, covert conflict, or evasion. For example, in the state of Texas, Palmer and Rangel (2011) found that teachers tried to balance the care of students’ needs with the pressures from the system of accountability (for example, to teach to the test).

On their part, Kostogriz and Doecke (2011) showed how Australian teachers have been able to assert their professional status and autonomy under accountability systems, through a praxis that aimed to achieve the public good and challenged the regulatory framework of standards-based reform. Other investigations have found similar results and have observed how teachers negotiate their professional agency in the context of accountability systems or how they mediate and challenge the constraints imposed on their professional autonomy (see Hardy, 2014; Osborn, 2006; Robinson, 2012).

Finally, in the US context, scholars such as Spillane and Kenney (2012) and Koyama (2013) have shown that schools can even adapt the official discourses of accountability in education to make them work on their own advantage. These authors show that many school principals are using the standardized test results and the economic competitiveness discourses surrounding them in a tactical way to capture the attention of vulnerable families and encourage them to become more involved in their children’s education and in the school dynamics.

THE SOCIAL AND MATERIAL CONDITIONS OF ACCOUNTABILITY REFORMS

The institutional and socio-economic contexts where accountability policies are enacted condition the final results of these policies. The enactment of accountability policies does not take place in a “vacuum,” but occurs in schools that are embedded in specific socio-economic, institutional, and cultural contexts. In the words of Ball et al. (2012):

Policies enter different resource environments; schools have particular histories, buildings and infrastructures, staffing profiles, leadership experiences, budgetary situations and teaching and learning challenges.

(p. 19)

Thus, to understand the effects of accountability reforms, beyond the particular role of teachers' agency (as highlighted in the previous section), we should take into account how material, social, and structural factors shape the dynamics of schools, and act as both enablers and constraints in the deployment of accountability programs (Ball et al., 2012).

In many cases, research on policy impact and enactment tends to omit variables of a contextual or institutional nature, although there are several exceptions. For example, on the basis of a qualitative case study conducted in an English secondary school, Keddie (2014) found that this school responded to accountability pressures in a more expressive (instead of instrumental) way, to a great extent because of "its advantageous situated factors" (Keddie, 2014, p. 515). Following a similar reasoning, but in a totally different context, Douglas (2005) concluded that cuts in the teaching staff, the existence of overcrowded classrooms, and inappropriate materials for teaching importantly constrained the policy enactment of a performance evaluation system in South Africa.

The comparison between schools with different academic performance is a common way of trying to capture the influence of contextual variables and the role of institutional factors in the policy enactment of reforms. On the basis of data collected with an ethnographic study in Chicago, Diamond and Spillane (2004) showed that the school status (on-probation or high-performing) can lead to the adoption of differential responses to the pressures from the accountability system. In this regard, on-probation schools adopted more superficial responses to avoid sanctions in the short term, while high-performing schools aligned their responses with the real aspirations of policy-makers (i.e. the introduction of instructional changes and innovations). Of course, institutional and socio-economic factors tend to be strongly related. In her own comparison between probation and non-probation schools in Chicago, Finnigan (2010) observes that socio-economic factors importantly influence the probation status (with 92 percent of probation schools in Chicago serving mostly low-income students).

On the basis of quantitative longitudinal data of Chicago public schools, Bryk et al. (2010) explore the influence of organizational features in school improvement. These scholars point out that leadership plays a pivotal role in school improvement dynamics, and can drive changes in related areas, such as school climate, instructional guidance, parent–community school ties, and faculty professional capacity. Nevertheless, as Bryk (2010, p. 28) states, social context also matters and, as a consequence, the "organizational development is much harder to initiate and sustain in some community contexts than others."

Educational reforms based on high-stakes accountability tend to lead to the more motivated teachers leaving the schools "on probation." In this sense, accountability policies could contribute in an unintended way to increasing inequalities between schools. These are at least the conclusions of a study conducted in eleven schools on probation by Mintrop (2003). Specifically, this author found that teachers of on-probation schools are skeptical about accountability policies as the most important driver of school improvement. To these teachers, the accountability system does not acknowledge sufficiently their efforts in these difficult contexts and, as a consequence, the most motivated of them look forward to moving to more advantaged schools.

Finally, Falabella (2014) examined how the institutional context of Chilean schools strategically influences teachers' and principals' responses to accountability system. Her research distinguishes three groups of schools (high-performing schools, middle-low-performing schools, and low-performing schools). For the first group, accountability policies serve to strengthen the school's social prestige and its advantaged position in the local

education market. However, the accountability system does not create incentives to introduce pedagogical or managerial changes. The second group of schools has a very loyal demand from families, mainly for symbolic reasons. Again, these schools have little incentives to introduce educational improvements based on the results of standardized tests. Finally, the group of low-performing schools is the most sensitive to the results of the accountability system. Generally, these schools have high percentages of vulnerable students and the accountability system generates a feeling of stress among teachers due to the sanctions and surveillance dynamics that are associated with poor performance (Falabella, 2014).

FINAL REMARKS: FOUR PREMISES FOR FUTURE RESEARCH

Numerous countries are currently adopting learning-based accountability policies as a way to improve the effectiveness of their educational systems, as well as the productivity of their teaching force. This form of accountability is globalizing but there are still many uncertainties around whether and under what conditions it works, and its real effects in education quality and other educational dimensions.

In this chapter, we have reviewed the existing literature on accountability in education through a scoping review approach. The adoption of this methodology has allowed us to dig into the pool of existing research on the topic and to systematize its most important results. More importantly, the scoping review approach has allowed us to reflect on how different accountability policy designs, enactment processes, and contextual contingencies might lead to differential effects and outcomes. In this respect, four basic premises for future research derive from our review. These are:

1. Accountability is not a specific education policy; rather, it covers very different policy options. Different accountability designs might lead to diverse outcomes, but also to different responses and enactment processes.
2. Despite accountability in education mainly being seen as a tool to improve students' learning, it also affects other schooling dimensions. Some of these effects are desired and/or expected, whereas others are not (including narrowing the curriculum, students' selection, triaging, etc.). The fact that learning-based accountability alters core educational processes and teachers' priorities makes this premise especially relevant from the quality education perspective.
3. Accountability is usually a contentious policy option; the process of adoption of accountability systems—as well as the particular characteristics of these systems—conditions their reception and enactment by teachers, as well as their final implications. Although in some contexts, teachers and principals are using—and even gaming—the accountability system in a way that supports educational processes, in many other cases the contentious nature of accountability schemes undermines teachers' attempts to provide high-quality education.
4. Accountability in education, as in any other global policy, is not context resilient. The institutional and socio-economic contexts where accountability policies are enacted shape the trajectories, responses, and final results of these policies. Low-performing schools—which are often located in low-income contexts—feel more improvement pressure than well-performing schools; however, such pressure tends to trigger more instrumental than expressive responses on behalf of schools, which arguably has neutral or even negative implications for the overall educational quality.

It is important to notice that the reviewed studies for the elaboration of this chapter have been conducted mainly in industrialized countries. For this reason, we must be cautious with the

generalization of the findings to other world locations. For the same reason, we do not claim that the premises systematized here have universal validity. Nonetheless, the four identified premises can contribute to lay the foundations of more comprehensive research frameworks that could be relevant in different educational settings where accountability schemes have been adopted. In the end, these four premises mainly tell us about the under-researched elements and absences in existing literature on accountability in education, and about the missing causal links between the different moments and components of the accountability policy process.

Specifically, premise 1 reflects on the importance of scrutinizing the specific design characteristics of accountability schemes as a first and necessary step to promote more rigorous evaluations of the distinct trajectories and effects of accountability in education. Because of their major implications, design options should not be simply seen as technical decisions, but as profoundly political. As we have seen, some of these options have the capacity to alter the most fundamental relations of power within educational systems and generate very different types of responses among teachers and other key educational stakeholders.

From premise 2, it derives that the non-expected/non-desired effects of accountability should also be included in the research frameworks or analytical models of future studies. There is still an overwhelming majority of research on accountability in education that narrowly focuses on learning outcomes as the unique dependent variable. This is something that definitely restricts our understanding of the systemic nature of the effects of accountability in education. In particular, more research on the effects of accountability measures on teachers' work, identities, and behavior, and concerning what kind of accountability systems encourage collaborative or competitive attitudes among educational actors, would be welcome.

Premise 3 tells us about the importance of opening the black box that exists between the policy design and its final outcomes. Even economists of education who try to measure the impact of accountability measures through statistical data have realized the importance of the implementation moment (see Woessmann, 2007). Nonetheless, policy implementation is not a mechanical process, but one strongly shaped by subjective variables (professional identities, perceptions and beliefs, multiple interpretations, teachers' motivation, and so on) that are difficult to measure in quantitative terms. Despite an emerging body of research focusing on several of these variables, investigation into the connections between the enactment of accountability policies (by teachers, principals, and other educational actors) and different types of educational results is still scarce.

Finally, premise 4 reflects on the importance of looking at contextual variables to understand the deployment and impact of accountability policies. Through our review, we have detected that most research on the theme is biased toward a school effectiveness approach that usually pays insufficient attention to how the socio-economic context strategically mediates the way accountability policies operate at the school level. Conducting comparative studies between different countries, states, or regions could shed light on the complex interplay between institutional and socio-economic settings, enactment processes, and the effects of accountability in multiple dimensions, including teachers' identity, autonomy, and motivation.

Overall, the way teachers are conceived and treated within accountability schemes, but also how teachers perceive, interact, and perform within such schemes, is key to understanding the different and sometimes contradictory outcomes of accountability in education. High-stakes accountability tends to generate more resistance than other models because of the issues it raises in defining what part of students' success or failure can be attributed to teachers. As stated by Leithwood and Earl (2000), accountability systems would be more legitimate if, instead of holding teachers accountable mainly for their students' learning outcomes—which depend on many factors—teachers were held accountable for making the most beneficial uses of the available resources in an effort to move toward the broader goals of the educational system. This, of course, should include improving learning outcomes, but also promoting conducive learning environments and quality education experiences for all.

REFERENCES

- Alegre, M. À. (2015). *Què funciona en educació: la pregunta necessària*. Barcelona: Fundació Jaume Bofill & Ivalua. Retrieved from: http://www.ivalua.cat/documents/1/07_07_2015_15_54_14_Que_funciona_educacio.pdf.
- Andersen, S. C. (2008). The impact of public management reforms on student performance in Danish schools. *Public Administration*, 86(2), 541–558. doi: 10.1111/j.1467-9299.2008.00717.x.
- Anderson, J. A. (2005). *Accountability in education*. Paris: UNESCO-International Institute for Educational Planning.
- Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher*, 36(5), 258–267. doi: 10.3102/0013189X07306523.
- Ball, S. J., Maguire, M., & Braun, A. (2012). *How schools do policy. Policy enactment in secondary schools*. New York, NY: Routledge.
- Boarini, R., & Lüdemann, E. (2009). The role of teacher compensation and selected accountability policies for learning outcomes: An empirical analysis for OECD countries. *OECD Journal: Economic Studies*, 1, 1–20. doi: 10.1787/eeco_studies-v2009-art10-en.
- Booher-Jennings, J. (2005). Below the bubble: “Educational triage” and the Texas accountability system. *American Educational Research Journal*, 42(2), 231–268. doi: 10.3102/00028312042002231.
- Bovens, M. (2007). Analyzing and assessing accountability: A conceptual framework. *European Law Journal*, 13(4), 447–468. doi: 10.1111/j.1468-0386.2007.00378.x.
- Breakspear, S. (2012). The policy impact of PISA: An exploration of the normative effects of international benchmarking in school system performance. OECD Education Working Papers, 71. *OECD Publishing*. doi: <http://dx.doi.org/10.1787/5k9fdffr28-en>.
- Bryk, A. S. (2010). Organizing schools for improvement. *Phi Delta Kappan*, 91(7), 23–30.
- Bryk, A. S., Sebring, P. B., Allensworth, E., Easton, J. Q., & Luppescu, S. (2010). *Organizing schools for improvement: Lessons from Chicago*. Chicago, IL: University of Chicago Press.
- Chiang, H. (2009). How accountability pressure on failing schools affects student achievement. *Journal of Public Economics*, 93(9), 1045–1057. doi: 10.1016/j.jpubeco.2009.06.002.
- Darling-Hammond, L. (2004). Standards, accountability, and school reform. *The Teachers College Record*, 106(6), 1047–1085.
- Day, C. (2002). School reform and transitions in teacher professionalism and identity. *International Journal of Educational Research*, 37(8), 677–692. doi: 10.1016/S0883-0355(03)00065-X.
- Dec, T. S., & Jacob, B. (2011). The impact of No Child Left Behind on student achievement. *Journal of Policy Analysis and Management*, 30(3), 418–446. doi: 10.1002/pam.20586.
- Diamond, J., & Spillane, J. (2004). High-stakes accountability in urban elementary schools: Challenging or reproducing inequality?. *The Teachers College Record*, 106(6), 1145–1176.
- Douglas J. (2005). *Teacher professionalism and education transformation*. Braamfontein: Centre for Education Policy Development (CEPD).
- Escardíbul, O. (2015). Is the introduction of pay-for performance salary incentives for teachers linked to students’ academic performance advisable? *What Works in Education Series*, 1, Barcelona: Fundació Jaume Bofill & Ivalua. Retrieved from: http://www.ivalua.cat/documents/1/27_10_2015_13_44_00_WhatWorksinEducation_english.pdf.
- Falabella, A. (2014). Do national test scores and quality labels trigger school self-assessment and accountability? A critical analysis in the Chilean context. *British Journal of Sociology of Education*, 37(5), 1–18. doi: 10.1080/01425692.2014.976698.
- Finnigan, K.S. (2010). Principal leadership and teacher motivation under high-stakes accountability policies. *Leadership and Policy in Schools*, 9(2), 161–189. doi: 10.1080/15700760903216174.
- Hamilton, L. S., Stecher, B. M., Marsh, J. A., McCombs, J. S., & Robyn, A. (2007). *Standards-based accountability under No Child Left Behind: Experiences of teachers and administrators in three states*. Santa Monica, CA: Rand Corporation.
- Hardy, I. (2014). A logic of appropriation: Enacting national testing (NAPLAN) in Australia. *Journal of Education Policy*, 29(1), 1–18. doi: 10.1080/02680939.2013.782425.
- Hatch, T. (2013). Beneath the surface of accountability: Answerability, responsibility and capacity-building in recent education reforms in Norway. *Journal of Educational Change*, 14(2), 113–138.
- Hogan, A., Sellar, S., & Lingard, B. (2016). Corporate social responsibility and neo-social accountability in education: The case of Pearson plc. In Verger, A., Lubienski, C., and Steiner-Khamsi, G. (Eds), *World yearbook of education: The global education industry* (pp. 107–124). New York, NY: Routledge.
- Jacob, B. A., & Levitt, S.D. (2003). Rotten apples: An investigation of the prevalence and predictors of teacher cheating. *Quarterly Journal of Economics*, 118(3), 843–877. doi: 10.1162/00335530360698441.

- Jeffrey, B., & Woods, P. (1996). Feeling deprofessionalised: The social construction of emotions during an OFSTED inspection. *Cambridge Journal of Education*, 26(3), 325–343. doi: 10.1080/0305764960260303.
- Jennings, J. L. (2010). School choice or schools' choice? Managing in an era of accountability. *Sociology of Education*, 83(3), 227–247. doi: 10.1177/0038040710375688.
- Jones, B. D., & Egley, R. J. (2004). Voices from the frontlines: Teachers' perceptions of high-stakes testing. *Education Policy Analysis Archives*, 12(39), 1–34.
- Kalimullah, N. A., Ashraf, K. M., & Ashaduzzaman, M. N. (2012). New public management: Emergence and principles. *Bup Journal*, 1(1), 1–22.
- Kamens, D. H., & Benavot, A. (2011). National, regional and international learning assessments: Trends among developing countries, 1960–2009. *Globalisation, societies and education*, 9(2), 285–300. doi: 10.1080/14767724.2011.577337.
- Keddie, A. (2014). 'It's like Spiderman... with great power comes great responsibility': School autonomy, school context and the audit culture. *School Leadership & Management*, 34(5), 502–517. doi: 10.1080/13632434.2014.938040.
- Keddie, A. (2015). School autonomy, accountability and collaboration: A critical review. *Journal of Educational Administration and History*, 47(1), 1–17. doi: 10.1080/00220620.2015.974146.
- Kostogriz, A., & Doecke, B. (2011). Standards-based accountability: Reification, responsibility and the ethical subject. *Teaching Education*, 22(4), 397–412.
- Koyama, J. (2013). Global scare tactics and the call for US schools to be held accountable. *American Journal of Education*, 120(1), 77–99. doi: 10.1086/673122.
- Ladd, H. F. (1999). The Dallas school accountability and incentive program: An evaluation of its impacts on student outcomes. *Economics of Education Review*, 18(1), 1–16.
- Ladd, H. F., & Lauen, D. L. (2010). Status versus growth: The distributional effects of school accountability policies. *Journal of Policy Analysis and Management*, 29(3), 426–450.
- Leithwood, K., & Earl, L. (2000). Educational accountability effects: An international perspective. *Peabody Journal of Education*, 75(4), 1–18.
- Linn, R. (2000). Assessments and accountability. *Educational Researcher*, 29(2), 4–16. doi: 10.3102/0013189X029002004.
- Maxcy, B. (2009). New public management and district reform: Managerialism and deflection of local leadership in a Texas school district. *Urban Education*, 44(5), 489–521. doi: 10.1177/0042085908318778.
- McDermott, K. A. (2007). "Expanding the moral community" or "blaming the victim"? The politics of state education accountability policy. *American Educational Research Journal*, 44(1), 77–111. doi: 10.3102/0002831206299010.
- Menken, K. (2006). Teaching to the test: How No Child Left Behind impacts language policy, curriculum, and instruction for English language learners. *Bilingual Research Journal*, 30(2), 521–546.
- Mintrop, H. (2003). The limits of sanctions in low-performing schools: A study of Maryland and Kentucky schools on probation. *Education Policy Analysis Archives*, 10(3), 1–30.
- Mosoge, M. J., & Pilane, M. W. (2014). Performance management: The neglected imperative of accountability systems in education. *South African Journal of Education*, 34(1), 1–18. doi: 10.15700/201412120947.
- Müller, J., & Hernández, F. (2010). On the geography of accountability: Comparative analysis of teachers' experiences across seven European countries. *Journal of Educational Change*, 11(4), 307–322.
- Ohemeng, F., & McCall-Thomas, E. (2013). Performance management and "undesirable" organizational behaviour: Standardized testing in Ontario schools. *Canadian Public Administration*, 56(3), 456–477. doi: 10.1111/capa.12030.
- Osborn, M. (2006). Changing the context of teachers' work and professional development: A European perspective. *International Journal of Educational Research*, 45(4), 242–253. doi: 10.1016/j.ijer.2007.02.008.
- Palmer, D., & Rangel, V. S. (2011). High stakes accountability and policy implementation: Teacher decision making in bilingual classrooms in Texas. *Educational Policy*, 25(4), 614–647. doi: 10.1177/0895904810374848.
- Perryman, J., Ball, J. S., Maguire, M., & Braun, A. (2011). Life in the pressure cooker – school league tables and English and mathematics teachers' responses to accountability in a results-driven era. *British Journal of Educational Studies*, 59(2), 179–195. doi: 10.1080/00071005.2011.578568.
- Powers, J. M. (2003). An analysis of performance-based accountability: Factors shaping school performance in two urban school districts. *Educational Policy*, 17(5), 558–585. doi: 10.1177/0895904803256789.
- Ranson, S. (2003). Public accountability in the age of neo-liberal governance. *Journal of Education Policy*, 18(5), 459–480. doi: 10.1080/0268093032000124848.

- Reback, R. (2008). Teaching to the rating: School accountability and the distribution of student achievement. *Journal of Public Economics*, 92(5), 1394–1415. doi: 10.1016/j.jpubeco.2007.05.003.
- Reback, R., Rockoff, J., & Schwartz, H. L. (2014). Under pressure: Job security, resource allocation, and productivity in schools under No Child Left Behind. *American Economic Journal: Economic Policy*, 6(3), 207–241. doi: 10.3386/w16745.
- Robinson, S. (2012). Constructing teacher agency in response to the constraints of education policy: Adoption and adaptation. *Curriculum Journal*, 23(2), 231–245. doi: 10.1080/09585176.2012.678702.
- Roderick, M., Jacob, B. A., & Bryk, A. S. (2008). The impact of high-stakes testing in Chicago on student achievement in promotional gate grades. *Educational Evaluation and Policy Analysis*, 24(4), 333–357.
- Ryan, J. E. (2004). Perverse incentives of the No Child Left behind Act. *The NYUL Review*, 79, 932–989.
- Sahlberg, P. (2010). Rethinking accountability in a knowledge society. *Journal of Educational Change*, 11(1), 45–61.
- Spillane, J. P., & Kenney, A. W. (2012). School administration in a changing education sector: The US experience. *Journal of Educational Administration*, 50(5), 541–561. doi: 10.1108/09578231211249817.
- Tobin, M., Lietz, P., Nugroho, D., Vivekanandan, R., & Nyamkhuu, T. (2015). *Using large-scale assessments of students' learning to inform education policy: Insights from the Asia-Pacific region*. Melbourne: Australian Council for Educational Research (ACER).
- UNESCO. (2015). *Framework for Action Education 2030: Towards inclusive and equitable quality education and lifelong learning for all*. Retrieved from: <http://www.uis.unesco.org/Education/Documents/wef-framework-for-action.pdf>.
- Volante, L. (2004). Teaching to the test: What every educator and policy-maker should know. *Canadian Journal of Educational Administration and Policy*, 35. Retrieved from: <http://www.umanitoba.ca/publications/cjeap/articles/volante.html>.
- West, A. P., Mattei, & Roberts, J. (2011). Accountability and sanctions in English schools. *British Journal of Educational Studies*, 59(1), 41–62.
- Woessmann, L. (2007). International evidence on school competition, autonomy, and accountability: A review. *Peabody Journal of Education*, 82(2–3), 473–497. doi: 10.1080/01619560701313176.
- World Bank. (2015). What matters most for school autonomy and accountability: A framework paper. *SABER Working Paper Series*, 9. Retrieved from: https://olc.worldbank.org/sites/default/files/Whatmatters_SAA_framework_paper_Demas_2014_4.pdf.