Background paper prepared for the 2017/8 Global Education Monitoring Report

Accountability in education: Meeting our commitments

Accountability and education in the post-2015 scenario: International trends, enactment dynamics and socio-educational effects

A n t o n i   V e r g e r   a n d   L l u í s   P a r c e r i s a .   U n i v e r s i t a t   A u t o n ò m a   d e   B a r c e l o n a
2 0 1 7

Outline

1. Introduction

In democratic societies, accountability is a fundamental aspect in the good governance of any public service, including education. Although some forms of accountability have been present in education systems for some time, currently, accountability in education is more central than ever within the global education debate. Accountability policies are pivotal in the policy recommendations of numerous international organizations and other key players in education. In the 2015 Incheon Declaration, agreed upon in the World Education Forum 2015, UNESCO and other international agencies, such as the World Bank, UNICEF and the UNDP, governments and civil society organizations manifested that:

> We are determined to establish legal and policy frameworks that promote accountability and transparency as well as participatory governance and coordinated partnerships at all levels and across sectors, and to uphold the right to participation of all stakeholders (World Education Forum, 2015, p.3)

and

> We resolve to develop comprehensive national monitoring and evaluation systems in order to generate sound evidence for policy formulation and the management of education systems as well as to ensure accountability. (World Education Forum, 2015, p.4)

As the Incheon Declaration makes clear, key stakeholders perceive accountability as an important instrument in the governance of education at multiple scales. The most common reasons given for adopting accountability systems in education are **efficiency** (i.e., better alignment between governmental or societal aspirations and the purposes schools try to achieve), **effectiveness** (improved performance on the part of schools, teachers and students), **equity** (guarantee that all students reach a minimum level of competence in core subjects) and/or **good governance** (accountability as a mechanism that can promote transparency and democratic control in educational services).

At the same time, accountability is a policy proposal whose effective implementation is complex and requires important technical, administrative and economic resources. Accountability is a slippery and elusive concept that can be used in different ways to mean different things to different people (Bovens, 2007; West et al., 2011). The mechanisms through which accountability policies operate may vary substantially, as the directions of accountability can be diverse: from governments to the international community, from the international community to civil society, from schools to governments, from governments to schools, etc. In fact, as we present in this paper, scholars tend to differentiate between several types or approaches to accountability: political, market, professional and managerial, among others. Despite the different modalities of accountability that may exist, we are currently witnessing a greater push for **accountability systems of an administrative nature that are linked to large-scale evaluations** (and in particular, to the evaluation of learning outcomes through standardized testing) and that hold some level of consequence for those being monitored, mainly schools, principals and teachers (Kamens and Benavot, 2011; Tobin et al., 2015).

Nonetheless, despite global agreement on accountability as a desirable policy approach, evidence regarding the effects of accountability policies in *real* situations remains far from conclusive. As we argue in this paper, the

reception and impact of accountability systems in education will be strongly conditioned by variables related to the specific policy design of these systems, as well as by mediating factors of a contextual and institutional nature.

1.1. About the paper

This policy paper addresses accountability in education from three different perspectives, namely, international dissemination, policy enactment and policy effects. After providing a general definition of accountability systems and presenting the **main models of accountability in education** (section 2), we will address key questions regarding the three mentioned perspectives. In the section on **international dissemination** (section 3), we focus on the main drivers in the promotion of accountability policies in education, including the key role of international organisations in this respect. By doing so, we will also discuss particular models of education accountability that are becoming dominant in the global education agenda and the reasons for their popularity. In the section on **enactment** (section 4), we focus on how accountability policies are perceived and received at the school level and, specifically, on how these policies are being put into practice by key educational actors. In the section on the **effects** of accountability (section 5), we reflect on the main implications of accountability policies in different dimensions, namely, learning outcomes, curriculum and evaluation, professional autonomy, and teacher satisfaction, among others. In the final section (section 6), we close this paper with general conclusions and by providing research and policy recommendations.

1.2. Methodology

This paper is based on the so-called scoping review method. This review method, in contrast to a systematic literature approach, does not depart from a very specific research question, nor does it aim to test a particular theory. This approach instead focuses on identifying the primary 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, within a relatively short period of time (Alegre, 2015).[1]

This review employed two primary scientific databases: SCOPUS and Web of Science. In addition, we also conducted hand-searching in key journals, books and materials of international organizations. The aim of this think-piece is to elaborate a framework through which to analyse accountability in education from multiple angles. The research conducted allowed us to identify the main debates, tensions and agreements that the adoption, implementation and impact of accountability policies generate in the education research field. The scoping review was conducted at three different levels (dissemination, enactment and effects) and included articles published since the year 1995. For each level of analysis, several keywords and search criteria were used (see Box A1 in the Appendix). In total, 277 documents were identified and for the purpose of elaboration in this paper, 150 documents were reviewed in more depth.

The scoping review allowed us to identify important gaps and biases in existing research on the accountability of education that, to a great extent, affect the nature of the data and sources of evidence that we were able to

---

[1] We introduce some procedures of the systematic literature review approach in this paper to render our review more rigorous and transparent (see Arksey and O'Malley 2005).

review. Specifically, we observed existing literature on accountability in education to be clearly biased in favour of: a) research conducted in industrialized Anglo-Saxon countries, with a very low presence of studies conducted in other global locations, particularly in Southern Hemisphere countries; b) research that focuses on learning outcomes as the main dependent variable, as opposed to studies that focus on the effects of accountability in educational processes, social relations and policy enactment; c) research that focuses on forms of performance and market accountability, as opposed to research focusing on other modalities of accountability such as social, network and/or mutual accountability; d) a school effectiveness approach that generally does not pay sufficient attention to the social conditions in which accountability reforms are being implemented and in which they operate.

Most of these absences regarding accountability in education literature will be highlighted in the following sections of this paper, as we develop our review.

2. Accountability in education: main concepts and categories

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 publics (Bovens, 2007; Hatch, 2013).

The concept of accountability is generally tied to the concepts of evaluation and transparency; however, these are not interchangeable concepts. Accountability requires some form of evaluation or assessment (for instance, of the inputs, processes and/or results of education); however, for accountability to happen, not all types of evaluation systems are valid. For example, an evaluation that is conducted in a small sample of schools for diagnostic reasons will not necessarily imply an accountability relationship. Accountability means that there is a legal and/or moral obligation to provide an account, and implies some level of responsiveness on behalf of the subjects being evaluated; for instance, it might require that subjects react or change their behaviour according to the results of the specific scrutiny. The fact that both obligation and responsiveness are inherent characteristics in accountability systems is well captured in a definition by 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 (Bovens 2007, p.450).

At the same time, we need to distinguish accountability from transparency. The promotion of transparency can be one of the potential outcomes of any accountability system. However, the enactment of accountability does not necessarily mean that all types of data on, for example, education systems' performance (and on the performance of schools and teachers) automatically become available to citizens, political groups and/or the research community. In fact, this type of data is generally sensitive, due to an inclination on the part of the media to elaborate rankings and league tables of schools.

Accountability systems can be significantly different according to who is expected to provide the account, to whom is the account owed, what is to be accounted for and what the consequences are of providing an account (see Leithwood and Earl, 2000). Several scholars have created typologies of accountability systems in a range of policy fields according to these and other variables. In Table 1, we present some of the most well-established taxonomies of accountability in education. The table highlights many of the accountability categories as fairly recurrent, particularly the professional, political and market categories; others are labelled differently, despite having similar meanings (for example, the hierarchical and the bureaucratic, the contractual and the legal, or the market and consumer accountabilities).

Of all of the taxonomies, West et al.'s (2011) unpacks the accountability modalities further, in part because it places greater emphasis on different forms of participative and public accountability. Maroy and Voisin's (2013) taxonomy is the only one in the table that is not guided by the nature of the forum, but to a great extent, by the nature of the obligation. Another reason for why their taxonomy is innovative is that it makes explicit what the theory of regulation – or policy ontology – is that serves as the foundation for the different accountability approaches. For example, on the one hand, assuming that educational actors are *homo economicus* and mainly sensitive to their own interests will give rise to an accountability approach grounded in extrinsic incentives. On the other hand, assuming that educational actors tend rather to respond to social norms and obligations will derive very different approaches to accountability (such as reflective or soft forms of accountability).

Table 1. Taxonomies of accountability in education
Source: authors

Leithwood and Earl (2000)	Ranson (2003)	Darling-Hammod (2004)	Anderson (2005)	West at al (2011)	Maroy and Voisin (2013)
• Professional • Managerial • Decentralisation • Market	• Professional • Consumer • Contract • Performative • Corporate	• Political • Legal • Bureaucratic • Professional • Market	• Legal • Professional • Results-driven	• Professional • Hierarchical • Market • Contractual • Network • Participative	• Strong accountability • Neo-bureaucratic • Reflective • Soft accountability

The most traditional forms of accountability are the political and the legal. **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). The assumption here is that if, for example, a government has not fulfilled its promises on education, it will not be re-elected or at least, will lose the votes of many disappointed citizens. In this type of accountability, civil society, Parliament and the media play a key role as watchdogs or forums to which the account is to be given.

Legal accountability means that schools and other educational actors (teachers, education management organisations (EMOs), policy-makers, etc.) are expected to operate in accord with legislation, and any citizen can bring to the courts complaints about the violation of laws by these educational actors (Darling-Hammond, 2004). What some call **contract** accountability derives from legal accountability and is particularly important in the context of forms of public-private partnerships (PPPs) such as charter schools or academies. In these cases, the

private sector needs to fulfil the conditions established in the PPP contract with the state if it wants this contract to be renovated, or not to be fined for breaching the contract (West et al., 2011).

Despite the fact that legal and political forms of accountability are still important, there is currently a greater emphasis on *administrative* forms of accountability, in which individual schools and/or teachers are generally those that provide the account to the government. Nonetheless, administrative accountability is far from being a monolithic category; it includes different models that can vary in terms of hierarchization, managerial style and the role played by professional bodies, the state, the market and/or the community.

Bureaucratic or **hierarchical** accountability is the most conventional modality of administrative accountability. In this modality, public authorities promulgate rules and procedures on how education should be delivered, and school inspectors (and/or other local authorities) try to ensure that schooling takes place according to such rules and procedures (Darling-Hammond, 2004). The first systems of bureaucratic accountability adopted an inputs-approach and, accordingly, focused on discovering whether the necessary resources were in place in schools for the appropriate delivery of the national curriculum (Falabella and de la Vega, 2016). However, currently, in many places inspection services go beyond the inputs approach and are also focusing on educational processes and on educational results as main domains of school monitoring and improvement.

Managerial or **neo-bureaucratic** accountability "includes systematic efforts to create more goal-oriented, efficient, and effective schools by introducing more rational administrative procedures" (Leithwood and Earl, 2000, p.14). This approach is also hierarchical in the sense that the state, through its multiple agencies, is the 'principal' in the accountability process and defines the goals that teachers and schools are expected to achieve, as well as the incentives and sanctions that will be associated with the process (Anderson, 2005). The most distinctive features of managerial accountability are a clear focus on outcomes and results (usually learning outcomes) and the generation of data through large-scale standardized evaluation tests. For 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., 2002).

In countries with higher levels of teacher professionalization, forms of **professional** accountability tend to have a longstanding trajectory. In these cases, "teachers and other staff are expected to acquire specialized knowledge, meet standards for entry, and uphold professional standards of practice in their work" (Darling-Hammond, 2004, p.1050). Within this modality, educators are accountable for adherence to professional standards and to their peers (Anderson, 2005). In contrast to the previous two modalities, in which schools and teachers are evaluated externally and have to respond to an external authority, professional accountability can be viewed as a form of horizontal accountability (West et al., 2011).

The **market** model involves accountability to the consumer (i.e., families) and is promoted through pro-school choice policies and competition dynamics. Here, the role of the state is apparently secondary, because it is assumed that demand-side forces will make schools more responsive to the needs of families and will put pressure on schools to strengthen the quality of their services. Nonetheless, the publication of school rankings and the enactment of performance-based funding formulas are some of the governmental measures that can contribute to strengthening the market accountability approach (Leithwood and Earl, 2000).

Participative or social accountability emphasizes the importance of participation by parents and the community in the education process (Ranson, 2003). In participative accountability, schools are accountable to families not via the school choice and exit mechanisms, but via voice and dialogue within school governing bodies (West et al., 2007). In some cases, particularly in developing countries, school-governing bodies are given important responsibilities in terms of the management of schools, e.g., deciding on future investments and even on the dismissal and hiring of teachers (Patrinos and Fasih, 2009).

Network accountability can be viewed as a form of participative and horizontal accountability in the sense that the participants in a network (or a partnership) are mutually accountable to one another. This includes networks between different schools, between community actors and schools, etc. In this instance, the key mechanism of accountability "is likely to be dialogue" (West et al., 2011, p.53). In the context of international aid and development the logic of network accountability is also applied, although at a different level (i.e., between donors and recipients) under the label of mutual accountability (see Eyben, 2008).

In the table below (Table 2), we summarize the main models of accountability according to the source of legitimacy, who the actor is and the forum in the accountability relationship, aspects of conduct that are focused on, the mechanisms of accountability and the consequences. In most countries of the world, more than one of these modalities may co-exist. At the same time, the same accountability mechanism can be used to promote more than one accountability approach. For example, standardized testing can be used to promote both market and managerial forms of accountability.

Table 2. Modalities of accountability: a synthesis
Source: authors

Source of legitimacy	Type of acc.	Actor/agent	Forum/principal	Aspect of the conduct	Mechanism	Consequences
Political	Political	Politicians School board	Parliament Voters Civil society Teachers organisations Media	Fulfilment of promises	Elections Strikes Public denunciation	Re-election/ No re-election
Legal	Legal	Government Schools	Courts Civil society Teachers' organisations	Behave according with legislation	Trial	Fines, dismissal
	Contract	The private sector, usually under the form of EMOs	Government	Fulfilment of the terms in the contract	Audits Performance evaluations (to find out about the achievement of the expected results defined in the contract)	Contract not renewed Fines
Administrative	Bureaucratic	Schools	Government	Inputs,	Conventional inspection services	Feedback to schools Schools' intervention

			Regulations, government guidelines		
Managerial	Schools, principals, teachers	Government	Learning outcomes and other results	Performance evaluation	Feedback to schools Incentives (sanctions/rewards)
Professional	Teachers	Teachers Teachers organisations	Professional standards, processes	Peer-evaluation	Feedback Mentorship Teachers' training
Market	Schools	Families	Results, responsiveness to parents' demands	School choice (and exit) Rankings (on the basis of performance evaluations)	Resources (fees, vouchers) Reputation and position in the schools' hierarchy (as perceived by parents and schools)
Social/ participative	Schools, teachers	Families, community	Responsiveness to parents and community demands	Voice, dialogue	Loyalty from families Level of internal democracy
Network	Schools	Schools, community partners	Participation in networks Quality of the exchange	Dialogue	Reputation Social capital Exchange of ideas and practices

3. International dissemination of accountability in education

Accountability has become a central norm in the global education reform agenda. However, globalization of accountability at the normative level does not automatically translate into specific and homogeneous changes on the ground (Martinsson, 2011). As we have seen from the above, there are numerous modalities of accountability, but not all of them have gained the same popularity in today's education policy arena. Political and bureaucratic approaches to accountability are longstanding, particularly in democratic and industrialized societies, whereas other modalities are more emergent, including managerial, market and social forms of accountability. Specifically, we are witnessing a gradual shift of emphasis toward a form of accountability focusing on the measurement of learning-outcomes through large-scale testing (Lingard et al., 2016), which we have labelled above as **managerial or test-based accountability**. In this section we focus first on the global drivers of this particular form of accountability in education and then on the particular role of international organizations in the dissemination of accountability measures.

3.1. Global drivers of test-based accountability
Recent literature shows how test-based- (learning-outcomes- or performance-) based accountability has penetrated into a broad range of countries and regions such as the UK and continental Europe (Barroso, 2009; Grek et al., 2009; Ozga, 2013; Vesely, 2012; Verger and Curran, 2014), the so-called Nordic countries (Elstad et al., 2009; Moos, 2013; Moller and Skedsmo, 2013), the US (Lipman, 2002; Hursh, 2005), Canada (Leicht et al.,

2009), Israel (Resnik, 2011), Australia and New Zealand (Codd, 2005; Lingard, 2010), and Latin America (Brooke, 2006; Pavon, 2008; Falabella and de la Vega, 2016). According to these and other studies, the factors that have led to this global trend are numerous. They include economic pressures for outputs-based educational reforms, the expansion of new public management within the educational sector, the spread of global education policies that are conducive to accountability adoption, and methodological and technical advances in the evaluation of education domains. Next, we present these global drivers for accountability reform in more detail.

In a global economy, ever more countries face major **economic pressures for educational reform** and governments and economic actors perceive learning, and particularly the acquisition of skills and competencies aligned to new labour market demands as a key strategy for raising their economic competitiveness (Carnoy and Rhoten, 2002). In fact, 'learning achievement' is considered by new human capital theory as the most significant independent variable for economic growth (Hanushek et al., 2003). The economically-driven education reforms that are being adopted worldwide conceive the increase of learning outcomes as a central goal and, accordingly, the measurement of learning outcomes has become a necessary condition for establishing the level of reform success. In a way, learning outcomes have become the main benchmark for discovering which policies "work" (or not) in impact evaluations and school evaluations. Nonetheless, what we wish to highlight here is that test-based accountability measures have become an important component of an educational reform approach that situates schooling "as the venue for increasing the economic competitiveness of the state" (Foster, 2004, cited in Koyama, 2013, p.82). The No Child Left Behind Act adopted in the US at the start of the 21st century is likely the most well-known and paradigmatic example of this reform approach (see Box 1 below)

Box 1. The No Child Left Behind Act (NCLB).

The No Child Left Behind Act (NCLB) is a federal law focused on K-12 education and promoted by United States President George W. Bush. In 2001, the NCLB passed Congress (with the support of both the Democratic and Republican parties) and the Senate. The law was enacted in January 2002.
This federal law promotes test-based accountability reform in education, with a focus on students' performance. Its primary goal is to make all students proficient in two key areas, mathematics and reading, and to close the performance gap between students of different social or ethnic origins. The four main principles of NCLB are (Stecher, Hamilton and Gonzalez, 2003, p.7):

• Greater accountability for student performance
• Increased local control and flexibility
• High-quality teachers using scientifically based practices
• Expanded options for parents

NCLB requires states to establish accountability systems. Through this law, the government puts pressure on schools and teachers and establishes a system of incentives for schools (and also for school districts) related to students' academic performance. At the same time, it promotes increased achievement goals that must be achieved by all students. Specifically, school districts and individual schools must meet the goals settled in an Adequate Yearly Progress plan (AYP). NCLB affects US public schools (and includes charter schools), but some of the measures only apply to the most disadvantaged schools (low-income schools or Title I schools). In this sense, low-income schools can be eligible to receive additional resources. Finally, NCLB aims to increase the transparency of a school's quality and teacher qualifications, and to promote school choice. Thus, in schools that fail to make adequate progress during an amount of consecutive years, parents have the option to move

> to another (more successful) school, or can request additional educational services such as free private tutoring from the school (Stecher, Hamilton and Gonzalez, 2003). After five years of not meeting AYP targets, NCLB contemplates the possibility of the state closing the failed school, or charterizing it.

The emergence of **New Public Management (NPM)** as a paradigm of public sector reform has placed greater emphasis on public services being managed more independently (i.e., through smaller managerial units) and according to the achievement of measurable outcomes (Gunter et al., 2016). In the context of educational reforms informed by an NPM approach, accountability measures are strongly sound and a key component of such reforms (Moller and Skedsmo, 2013). NPM promotes managerial governance styles at the school level and models of school management and leadership oriented toward the achievement of measurable goals (including learning goals). Despite the fact that during the 1980s, the NPM paradigm had been embraced by conservative and new right governments that mainly aimed at promoting public services that were more efficient and cost-effective (see Tolofari, 2005), today, NPM is accepted as a valid reform approach by a broader range of political ideologies. In fact, NPM measures are currently also part of the public sector reform agenda of social democratic governments (Verger and Normand, 2015). Social democratic parties, under the influence of the so-called 'third-way' during the 1990s[2], began adopting NPM reforms – including school autonomy and accountability – as a way to promote not only economic efficiency, but also equity and the diversification of public services.

Nonetheless, beyond the role of political ideologies and policy paradigms such as NPM, accountability policies might also be adopted because doing so is **politically convenient and rewarding** (Pollitt and Bouckaert, 2011). Enacting accountability systems allows politicians to signal to their publics that they are working hard toward educational change and that they are concerned with educational results and the future of children. At the same time, accountability reforms tend to involve low political risk (since, in fact, the reform pressure is put on schools and teachers, rather than on the government) and are "cheaper and quicker than alternative reforms" (Smith et al., 2004, p.50).

The spread of **global education policies** such as school autonomy and standards-based reform has also become conducive to learning-outcomes accountability.[3] In current educational reform packages, school autonomy and accountability tend to be conceived as inseparable. This is due to the fact that governments (in their role as principals) should be willing to give more autonomy to schools in organizational, budgetary and/or curricular terms, to the extent that schools accept stricter supervision and control via accountability measures. As stated by the OECD (2013, p.45), "greater responsibilities assumed by schools imply greater accountability requirements such as external school evaluation and public reporting of student performance". Accountability is also reinforced by the standards-based reform movement, due to the fact that the development of more clear and measurable common core standards in curricular reforms leads to an increasing emphasis on tests, rewards and sanctions (Darling-Hammond, 2004).

---

[2] The third-way is a political ideology that seeks to reconcile socialism and capitalism and, to this purpose, combines egalitarian and individualist policies.

[3] To some extent, these are policies aligned with the NPM paradigm discussed above; however, these policies have also gained their own importance in the education policy discourse, in part due to the fact that they have pedagogical (and not only managerial) implications.

Methodological and technical advances in the definition of learning standards, as well as in the measurement of learning outcomes, have also made possible the expansion of performance-based accountability worldwide. The methodological progress in the field of standardized evaluation is a necessary condition for the development of more sophisticated and precise accountability systems. The Programme for International Students Assessment (PISA) of the OECD and similar international assessments (such as TIMSS or PIRLS) have contributed to politicians and policy makers turning toward a global 'education race' aimed at student achievement. In the context of this international educational race, the adoption of national standardized evaluation systems has become very strategic. Nonetheless, PISA has not only contributed to the introduction of competitive pressure in countries for better learning outcomes, but has also become instrumental in terms of transferring the technology that allows for testing learning skills at the national level (Meyer and Benavot, 2013; Lingard et al., 2016). This transfer of technology happened at a time when the measurement of learning outcomes has become a commonly agreed-on proxy for "education quality". Many practitioners and scholars have conflated quality to the more concrete idea of students' learning, in part because learning outcomes are more concrete, comparable and measurable than other types of education quality components such as, e.g., process variables (Sayed, 2011).

Finally, it is also worth noting that a **testing and measurement industry** has developed within the emergence of accountability reforms and that for market expansion reasons, this industry sector is highly interested in the deepening of these types of reforms. Companies like Pearson specialize in testing preparation services as well as in the evaluation and tracking of children's learning outcomes and, on the basis of these data, sell education improvement services to countries, local governments, schools and/or families (Hogan et al., 2016). Such economic interests in testing and measurement are also behind current education policy changes. According to the OECD, the fact that "standardized student assessment becomes a more profitable industry" means that "companies have strong incentives to lobby for the expansion of student standardised assessment as an educational policy therefore influencing the activities within the evaluation and assessment framework" (OECD, 2013, p.51). In a similar line of reasoning, Carnoy (2016, p.36) considers that "test makers have a vested economic interest to have educational systems and schools change what they define as academic knowledge or even useful knowledge to fit the particular test they sell".

3.2. The role of international organizations

The influence exerted by a range of international organizations in the education policy field is also behind the expansion of learning-based accountability in education. International organizations have the capacity not only to fund the implementation of accountability reforms, but also to promote normative emulation dynamics between member countries.

Many of the country case studies reviewed in this paper coincide in pointing to **OECD and PISA** in particular as key drivers of accountability reforms at the country level. In fact, after six editions of this influential international evaluation, school autonomy with accountability measures represent one of the policy recommendations that are consistently included in OECD/PISA reports (see OECD, 2011). These OECD recommendations have framed policy change within numerous educational settings. According to a recent study, 29 OECD country representatives (out of 37) admitted that PISA/OECD recommendations on accountability have influenced accountability reforms at the national level (Breakspear, 2012). In a more indirect manner, the *PISA shock* that

several countries have suffered – especially after the first editions of this international assessment – have also promoted the introduction of accountability reforms at the national level (Elstad et al., 2009).

The OECD division of education mainly advocates for managerial and learning-based forms of accountability. However, this organization also emphasizes that accountability systems should be holistic and focus on improving classroom practices, rather than placing excessive emphasis on learning outputs (see OECD, 2012, 2013). The OECD also highlights that the adoption of these types of policies should be agreed on by key education stakeholders:

> To be designed successfully, evaluation and assessment frameworks should draw on informed policy diagnosis and best practice, which may require the use of pilots and experimentation. To be implemented successfully, a substantial effort should be made to build consensus among all stakeholders, who are more likely to accept change if they understand its rationale and potential usefulness (OECD, 2013, p.14).

The OECD, on the basis of PISA data, considers accountability and school autonomy as two policies that fit well together. According to OECD/PISA, schools with greater autonomy in resource allocation show better student results in the context of strong accountability regimes (in which governments post schools' achievement data publicly). The opposite relationship is also true, since "in countries where there are no such accountability arrangements, schools with greater autonomy in resource allocation tend to perform worse" (OECD, 2011, p.1).

The **World Bank** places bigger emphasis on both market accountability and social accountability approaches. The World Bank actively disseminates these different forms of post-bureaucratic accountability through its numerous knowledge products, lending operations and more recently, through the *Systems Approach for Better Education Results* (SABER). According to the SABER framework paper on school autonomy and accountability, "increasing school accountability is a necessary condition for improving teacher quality" (World Bank, 2015, p.4) and "for improved learning because [this policy] *aligns* teacher and parent incentives" (World Bank, 2015, p.2). The World Bank also echoes the OECD message on the importance of school autonomy and accountability policies working together:

> School autonomy must be complemented with school accountability to promote academic excellence. This has been well documented through various impact evaluations. We also know that the highest PISA scores come from countries where autonomy and accountability are implemented together (OECD 2011) (World Bank, 2015, p.36)

For the World Bank, accountability in education is a key factor in the systems approach to education reform that promotes, alongside the 2020 Education Strategy "Learning for All". This document, which defines the World Bank education policy in the 2010-2020 period, considers that "improved performance and measurable outcomes depend on a careful balance between three policy instruments that influence the behavior of local actors: (1) greater autonomy at the local level; (2) enforcing relationships of accountability; and (3) effective assessment systems" (World Bank, 2011, p.33).

This international organization considers what it calls "the shorter route of accountability" (which operates through choice and school-based management) to have advantages over the "long route of accountability" (i.e., the bureaucratic model of accountability), since "the shorter route affords clients the power to more frequently

provide feedback to providers to let them know how they are [performing] and to hold them accountable for good quality services" (World Bank, 2015, p.5). Overall, for this international organization, post-bureaucratic forms of accountability (including market and social forms of accountability) are not only inherently desirable, but the most appropriate solution in the context of low-income countries without the sufficient material and administrative capacity to put other systems of accountability in place (Bruns et al., 2011; World Bank, 2006).

Accountability is also one of the most frequently used concepts in the Framework for Action Education 2030, which was agreed on by the international community, under the lead of **UNESCO**, to advance Sustainable Development Goal (SDG) number 4. This framework conceives the creation of national accountability systems that focus on learning outcomes as an indispensable aspect in the monitoring and achievement of new educational targets:

> Robust monitoring, reporting and evaluation policies, systems and tools are essential for the achievement of Education 2030. Monitoring quality in education requires a multi-dimensional approach, covering system design, inputs, content, processes and outcomes. (UNESCO 2015, p.17)

In contrast to the World Bank, UNESCO does not promote market forms of accountability. In fact, this international organization is sceptic about assumptions concerning the benefits of market mechanisms in education as advocated by the World Bank and like-minded organizations (see, for example, UNESCO, 2009). In the context of the *2000 Dakar Framework for Action* which was also led by UNESCO, the discourse on accountability was present but not sufficiently developed. This document referred to accountability more as a principle than as a concrete policy approach. In the few references to accountability included therein, the Dakar framework advocated a mix of managerial and social forms of accountability. Fifteen years later, in the context of the *Framework for Action Education 2030*, the idea of accountability reappears with greater force and in more concrete terms. This framework for action, co-signed by UNESCO and other agencies, is likely the document that collects the most official view of this international organization as it concerns accountability in education. In this document, both mutual accountability and a model of accountability that is participatory are highlighted. By focusing on the importance of deliberation and participation in the definition of accountability systems, UNESCO and its partners acknowledge the contentious nature of accountability in the education policy field:

> As the primary responsibility for monitoring lies at the country level, countries should build up effective monitoring and accountability mechanisms, adapted to national priorities, in consultation with civil society. This includes building greater consensus as to what specific quality standards and learning outcomes should be achieved across the life course – from early childhood development to adult skills acquisition and how they should be measured (UNESCO 2015, p.17)

UNESCO, through the Education 2030 Framework, promotes accountability systems that are at least partially based on the measurement of learning outcomes. The Sustainable Development Goal number 4 (SDG-4), which is the SDG that focuses on education, places important emphasis on the improvement of learning outcomes. Out of the seven targets included in SDG-4, five focus on learning outcomes and the achievement of particular skills (literacy, numeracy, global citizenship, culture of peace, etc.). This shift from school access to learning is likely the most significant change that can be observed between the EFA agenda and the SDGs/Education 2030 agenda.

To conclude this section, it is important to mention that this global shift toward a learning-outcomes-based accountability has not necessarily lead to policy convergence internationally. As stated by Van Zanten (2002, p.302), "States cannot avoid global pressures to change in specific directions, but they can twist and transform [these pressures] to fit national purposes and opportunities". Thus, depending to how governments regulate and design it, learning outcomes accountability can be attached to high-stakes or low-stakes evaluations, be more or less comprehensive (in terms of the areas of knowledge it covers) and complemented with other forms of professional, social and/or market accountability. Furthermore, as we develop in the next section, even when accountability systems appear to have very similar forms at the regulatory level, they may translate into very different practices according to how key educational actors interpret and enact them at different levels (district, school, classroom, etc.).

4. Enacting accountability policies

In the context of most educational reforms, a gap between policy aspirations and real practices tends to prevail (McDermott, 2007). Good policy intentions and even good policy designs are not a sufficient condition for reform success. The complicity of the agents that end up applying the policy in question (whether teachers, principals and/or families) is a key factor for understanding the successful implementation of educational reform.

In this section we focus on the role of two key players in the enactment of accountability systems at the local level, namely, teachers and principals, as well as on the mediating role of social contexts in the enactment of these systems. The concept of 'policy enactment' refers to the interpretation and translation processes involved in the deployment of policy programmes at the most micro-institutional level, i.e., the school, the classroom, etc. (Ball et al., 2012). According to enactment theory, the performance of a policy is determinately mediated by how the school staff perceives, gives meaning to and takes ownership of the policy in question. This premise applies to most public policies, but particularly to those related to sectors like education, with ample scope for discretion and creativity, as well as to reforms that tend to generate controversy and some levels of resistance, as is generally the case with accountability reforms.

4.1. Teachers as enactors of accountability

One of the conditions for effective policy enactment is that the subjects of the intervention perceive the policy as meaningful and as having positive effects (independently of whether this perception is correct or not). Jones and Egley (2004), on the basis of a survey applied to 708 teachers in the state of Florida, show that teacher perceptions of the high-stakes testing system in this state are quite negative, with most teachers (79.9%) reporting that this system "was *not* taking Florida's public schools in the right direction". According to these authors, such a negative perception affected the responses that teachers from Florida gave to the existing accountability system. In their words, "without the support of teachers, high-stakes testing will likely become just another failed education reform" (Jones and Egley, 2004, p.26). In the European context, Müller and Hernández (2010) also detected a general disbelief and scepticism among teachers towards any form of accountability, but particularly toward performance-oriented accountability.

Performance accountability measures generally tend to have a centralizing effect and alter the power relations within the education system (Maxcy, 2009). These measures tend to empower families, the public administration and/or school principals before teachers. The fact that teachers feel that performance accountability reforms disempower them, or make them lose control over curriculum-related decisions, is clearly a source of resistance and negotiation in the context of the deployment of these reforms. In some cases, teachers' responses to demanding accountability systems can even be belligerent. In South Africa, for example, teachers reacted to a new performance-based accountability system "by not submitting their records, not attending meetings and not carrying out the work agreed upon in their Personal Growth Plans" (Mosoge and Pilane, 2014, p.12). This was a reaction to the increasing surveillance that this policy involved, but also to the fact that it was "too time-consuming, too personnel heavy, too bureaucratic and involving too much paperwork" (SADTU, 2011, p.13, quoted in Mosoge and Pilane, 2014).

Nonetheless, the responses of teachers to high-stakes accountability policies tend to go beyond the categories of "resistance and consent" and include intermediate strategies of negotiation and evasion (Perryman et al., 2011). That is, in the context of accountability systems, teachers can negotiate how their interests and views might fit within the accountability demands, or simply avoid those aspects of the policy they dislike. An example of teachers' negotiation in the context of high-stakes accountability reforms can be found in a study conducted in Texas by Palmer and Rangel (2011). These authors found that, despite the accountability system increasingly forcing them to teach to the test, teachers made significant "efforts to engage authentically with students and learning and to balance the pressures to teach to the test with opportunities to address their students' unique needs" (p.623).

On occasion, the enactment of accountability systems forces teachers to vindicate their professional status and judgement. Research conducted in Australia by Kostogriz and Doecke (2011) shows that teachers are reflective actors who, despite being aware of the potential implications of standards-based reforms and performance-based accountability for their profession, continue to experience their professional practice on other terms than those defined by these reforms. According to them:

> Repeatedly in their conversations with us, teachers have described pedagogical events in which they attempt to reclaim the space and time for educative and responsive practice that is orientated towards the public good. Or, at least, so it seems to us – for we are capturing dimensions of their conversations with us that appear to suggest alternative ways of being and becoming, without necessarily constituting a praxis that breaks through the regulations imposed on them by standards-based reforms. (p.407)

Other research on the enactment of performance-based accountability also conducted in Australia found similar results (see Hardy, 2014; Robinson, 2012). For example, according to Robinson (2012, p.243):

> Although the teachers did to some extent behave as technicians and carry out tests, even teaching for the test when they knew it was imminent, they exercised reflection and continued to rely upon their professional judgement throughout their teaching practices. They worked hard to reconcile the requirements of the two policy-making bodies into a report which they could present to the parents and which they felt still demonstrated the development of the student. It is clear that the teachers in this study actively made choices about their actions in respect to the writing of reports.

And, as this same author concludes:

> Professional agency appears to be constructed through the collective actions of the teachers. Despite the strategies of performance, accountability and control mechanisms in conflicting policy text, it is important to note the importance of strong collegial relationships in enabling the teachers to negotiate their professional agency to adapt and adopt policy requirements to fit some practices and reshape others.(p. 244)

Similar results were found in a comparative study carried out in England, Denmark and France by Osborn (2006). According to this author, recent managerial reforms in the denoted countries, including performance accountability measures, made teachers feel restricted in terms of their professional autonomy. However, "some teachers' response to the reforms had been that of 'creative mediation', taking active control of the changes and gaining a new professional discourse… including new professional practices in the process" (Osborn, 2006, p.246).

4.2. The role of principals

The role of principals and their leadership styles importantly mediates how accountability reforms are perceived and received among teachers (Davis et al., 2000; Diamond, 2007). According to Finnigan (2012), principal leadership is key in school accountability policies, as principals' interpretations and responses to these policies can motivate teachers to use the policies for promoting better school performance. The findings of Finnigan's research in the context of low-performing schools in Chicago indicate the importance of trust between principals and teachers, and the presence of shared and instructional leadership styles as a way to motivate teachers to promote school improvement in the context of accountability regimes (Finnigan, 2012).

In a previous piece, Finnigan (2010) compares probation and non-probation schools in Chicago on the basis of data from the Consortium for Chicago School Research's survey conducted in 1997. Her analysis of this survey shows that socio-economic factors importantly influence probation status (with 92% of probation schools in Chicago serving mostly low-income students), but also that school leadership is key for motivating teachers to move away from probation. Specifically, the two areas of school leadership associated with teacher motivation and with infusing teachers with higher levels of expectancy in their own work are: "(1) instructional leadership (having a vision and high expectations for the school and knowing how students learn); (2) principal support for change (providing teachers with the resources they need and supporting them as they try new approaches)" (Finnigan, 2010, p.179).

In Chile, a school self-assessment programme was implemented in 2005 as a way to introduce an evaluation and accountability culture at the school level. An analysis of the implementation of this programme shows that in schools where teachers reported a lack of trust in principals, the programme was more poorly implemented (Montecinos et al., 2014). It has been also documented that in Chile, school principals tend to be more favourable to the national accountability system than teachers. This is partly because principals, as direct interlocutors with the public administration, feel more obliged to engage with and respond to the accountability system (Assaél et al., 2014). On several occasions, principals have even had to discourage their teachers to organize boycotts against the national standardized test, using the argument that this type of action can have negative

consequences for the school (Montecinos et al., 2015).

In the Canton of Bern, Switzerland, low-stake accountability reforms have been put in place alongside decentralization and principals' professionalization policies. Despite the initial resistance to these changes, they subsequently enhanced harmonization and new modes of both cooperation and competition between schools (Hangartner and Svaton, 2013). In the context of these reforms, teachers have apparently lost power in their relationships with principals, since the former are officially held accountable for their work by the latter. However, most principals know that their capacity to guide the pedagogical development that their schools require depends greatly on the support of teachers, reason why the end up adopting more distributed leadership styles. Furthermore, Hangartner and Svaton (2013) observed that ultimately, principals mainly asked their teachers to use self-evaluation tools, which did not put as much external pressure on teachers. These internal evaluations were combined with external evaluations by experts and the combination of these two evaluation approaches contributed to making the education system more transparent (Hangartner and Svaton, 2013).

Finally, in the US context, scholars such as Spillane and Kenney (2012) and Koyama (2013) have shown that school principals are very active in terms of making sense of and negotiating global and official discourses of accountability in education. The researchers' ethnographic studies show that many school principals who are generally located in vulnerable contexts are using the standardized test results to capture the attention of families and to encourage them to become (for the sake of "the future of their children") more involved in their children's education, as well as in the school dynamics.

4.3. School context and status

Most research on the enactment of accountability policies tends to focus on the role of key educational actors (mainly teachers and principals), but neglect the social, institutional and material conditions of schools. We nonetheless found some exceptions in our research that focus on how contextual variables affect accountability implementation, including the way educational actors interpret and use accountability policies. For example, Douglas (2005, p.15) states that "over-crowded classes, inadequate learning support materials, unsafe working environments, and uncertainties caused by looming retrenchments" are important constraints to the enactment of the performance evaluation system adopted in South Africa. Similarly, Keddie (2014) concludes that variables such as "the school's intake and history, its ethos and values, its access to human and economic resources and its status and power as an outstanding school" (p.502) strategically mediate the way in which schools engage with school autonomy and accountability policies.

To capture the role of contextual and institutional variables, comparisons between schools on probation versus high-performing schools occur frequently in the literature on the topic. Diamond and Spillane (2004) conducted an ethnographic study on the reception of accountability measures in four schools from Chicago, two on probation and two well-performing schools, and their results show that the probation status significantly conditioned how schools reacted to accountability pressures:

> incentives in probation schools were focused on sanctions rather than rewards and led to a superficial response based on external threats while high-performing schools responded in ways that were closer to the intentions of policy makers [i.e. instructional improvement] (Diamond and Spillane, 2004, p.1157).

Nonetheless, further research is necessary to discover whether schools on probation lean toward providing superficial responses to accountability systems, due to inherent institutional issues, or because of the external pressure that they receive from the accountability system itself.

Mintrop (2003) conducted research in 11 schools on probation in two US states and found that teachers in these types of schools were particularly sceptic of the accountability system in place and had serious doubts about its fairness. As one of the teachers stated when referring to the factors influencing the educational results, "it's not the teaching as much as it is the children" (p.7). Overall, teachers perceived the accountability system as unfair because puts all the responsibility on teachers (and on teaching practices) and omits the contextual conditions and the social composition of the schools where they are working. Thus, paradoxically, despite this accountability system aiming to motivate teachers via incentives, Mintrop (2003) found that the most motivated teachers in schools on probation did not view their efforts as having been sufficiently acknowledged and were planning to leave their current employment positions for better schools.

Research conducted in Chile also shows how the institutional context matters in a highly segmented education system. According to Falabella (2014), high-performing schools benefit from accountability policies, because these policies corroborate their good position in the schools' hierarchy. Accordingly, the accountability system does not provide these schools with incentives for changing their managerial or pedagogical styles. Falabella identified a second group of middle-to-low-performing schools that tended to ignore the accountability system results, but for different reasons than the elite schools. These were small schools with a very loyal demand coming from families that appreciate aspects such as, proximity, teachers' quality, and the familiar atmosphere that prevailed in these schools – well above their academic excellence. Thus, the incentives for change on the basis of test results were also low in the context of this second type of schools. A third group of schools was however much more responsive to the accountability system. These were low-performing schools that were placed on the bottom of the schools hierarchy, that struggled to attract students and that enrolled the most marginalized population (including those that had been expelled from other types of schools). These schools were highly stressed by the accountability system and placed achieving better results in external tests at the core of their management priorities. These low-performing schools did not place as much importance on the accountability system due to demand-side pressures, but because of the governmental sanctions and interventions attached to bad results (Falabella, 2014). Similarly, Assaél et al.'s (2014) research on probation schools in Chile show that the probation status of schools does not generate the intended school improvement dynamic; rather, it tends to generate a sense of meaninglessness among teachers and principals.

5. The variegated effects of accountability

This section explores the main effects of accountability policies according to different educational dimensions. Based on international evidence, we provide a global view of the effects of accountability policies in areas such as a school's management style, competition between schools, the relationship between parents and teachers, the behaviour of educational actors, curriculum delivery, learning

outcomes and social inequalities. We will focus on the intended effects of accountability policies, but also on the non-intended effects generated by these policies.

5.1. Learning outcomes

Studies analysing the effects of accountability on students' learning outcomes are increasingly frequent, particularly in the field of the economics of education. These are studies that have been conducted primarily in industrialized countries, and mainly in the US[4]. For this reason, we must be cautious with the generalization of their findings to other world regions and to developing contexts in particular.

Several researchers argue that accountability policies – and in particular high-stakes accountability - generate positive effects on students' academic results (Chiang, 2009; Gaddis and Lauen, 2014; Hanushek and Raymond, 2004; Roderick, Jacob and Bryk, 2002; Springer, 2008; Woessmann, 2007). Nonetheless, in many studies, results are inconclusive or show only low levels of association between accountability and learning results.

Woessman (2007) analysed data from four international assessments, i.e., TIMMS, TIMMS-R, PISA and PIRLS, in order to compare the influence of the institutional setting of different educational systems on students' academic performance. On the basis of these databases, Woessmann (2007) concludes that school autonomy, accountability and competition are three key institutional elements that promote high academic performance among students. Nonetheless, the author acknowledges that his analysis can only illustrate general patterns and as such, more research is required to obtain in-depth knowledge on the implementation of these policies.

Similarly, Chiang (2009) applied a discontinuous regression analysis using data from Florida and concludes that the threat of sanctions on low-performing schools generated an improvement in student results in mathematics in elementary education, which persisted through early years in middle-school; however, the evidence on improving reading was not particularly robust.

Springer (2008) examined the gains of accountability policies in student test scores. The data used in this research were taken from the Northwest Association Evaluation's Growth Research Database, which includes longitudinal data results at the student-level from 45 US states. This author suggests that the threat of sanctions on low-performing schools stimulated productivity. At the same time, the accountability system was able to improve the test results of low-performing students.

Similarly, Roderick et al. (2002) used a three-level hierarchical linear model in order to estimate the effect of educational accountability on achievement levels. This author employed administrative data from Chicago public

---

[4] Most of the studies included in this scoping review on the effects of accountability policies on students' performance were conducted in the US. Some analysed the effects of accountability across states (see Braun et al., 2014), using comparative databases such as the National Assessment of Educational Progress (NAEP) (e.g., Hanushek and Raymond, 2004; Dee and Jacob, 2011). NAEP is a low-stakes test-accountability that measures states' performance; however, in many cases, individual states also have high-stakes accountability policies. Indeed, there are also state-centred studies, all of which are located in a high stakes accountability environment (see Chiang, 2009; Figlio and Rouse, 2006; Gaddis and Lauen, 2014; Jennings and Sohn, 2014). Finally, other studies are based on international databases (see Woessman, 2007).

school students and found an extremely positive effect in two degrees (6th and 8th), where students improved their performance in the Iowa Test of Basic Skills. However, they found inconclusive evidence in relation to the results of students in third grade and observed that the effects of this policy differed depending on the school context.

Many researchers have found mixed or inconclusive results in this particular area of research (Braun, Chapman and Vezzu, 2010; Dee and Jacob, 2011; Figlio and Rouse, 2006; Jennings and Sohn, 2014; Ladd, 1999; Reback, Rockoff and Schartz, 2014). Boarini and Lüdemann (2009) conclude that although some aspects of accountability may be associated with improved student outcomes, others are not. In another study using panel data from schools in Texas, Ladd (1999) found positive and robust effects related to the performance of specific subgroups of students, but not for others. Specifically, Ladd (1999) found positive effects for white and Hispanic students in seventh grade, while black students obtained less positive results.

Reback et al. (2014) examined the effects of No Child Left Behind (NCLB) on school staff behaviour and students' achievements. Positive and neutral effects were found in relation to academic performance and enjoyment in the learning process. Dee and Jacob (2011), also focusing on the effects of NCLB on students' scores in national standardized tests, found a statistically significant positive effect on the results of fourth grade students in maths, but no evidence of improvements in reading.

Some Indian states are carrying out a decentralization process of public services in order to promote accountability to local communities. Pandey, Sundararaman and Goyal (2009; 2010) found that this reform has different impacts in different states but that overall, it has resulted only in the limited improvement of academic results. However, significant impacts have been found in relation to inputs and variables associated with educational processes.

Finally, another group of scholars found that accountability reforms had negative effects on learning outcomes (Andersen, 2008; Ryan, 2004; Powers, 2003). This is the case of Andersen (2008), who found that the implementation of a performance management reform in Danish schools had negative effects on the performance of low SES students attending the reformed schools. Ryan (2004) notes that test-based accountability frameworks such as NCLB create incentives that work against the improvement of low social-background students' performance. Similarly, on the basis of evidence from California, Powers (2003) highlights that the academic performance of schools within accountability frameworks is strongly mediated by their social conditions. According to these scholars, test-based accountability encourages school segregation and early school leaving among students with disadvantaged socio-economic conditions and/or who belong to ethnic minorities. It can also create incentives for governments to reduce academic standards and discourage the best teachers to work in schools with high levels of socially vulnerable populations or populations that perform poorly.

5.2. Inter-school dynamics

Recent literature has investigated the influence of accountability policies on the dynamics of collaboration and competition between schools. A group of scholars acknowledge that the publication of test results in league tables or rankings can generate incentives for competition between schools, teachers and students (Meijer, 2007; Jones and Egley, 2004; Keddie, 2014; Ohemeng and McCall-Thomas, 2013). Competition is a desirable

outcome for advocates of markets in education. However, the reviewed research also shows that competition does not always lead to school effectiveness, but to opportunistic behaviours and exclusionary practices on behalf of many schools in their bid to gain a better position in the local education market.

Based on qualitative research, Ohemeng and McCall-Thomas (2013) conclude that the publication of results and rankings, and the misuse of this information, contribute to create an atmosphere of competition between schools that mainly leads to increasing pressure to achieve better results. This external pressure translates into an increase of pressure on teachers, which leads them to adopt a 'teaching to the test' approach and other educational strategies that are questionable from a pedagogical point of view. Other adverse effects identified by Jones and Egley (2004) are the stigmatization of low-performing schools, the blaming of schools and teachers for poor performance, and the degradation of the image of public schools. Nonetheless, according to their research, only 3.8% of the surveyed teachers said that tests activated competitive dynamics between schools and among educators.

Through a qualitative case study conducted in an English secondary school, Keddie (2014) analysed how the school context mediates the potential effects derived from the combination of school autonomy and accountability, such as competition between schools. Her research departs from the premise that competitive dynamics can impact negatively on equity and educational practices, and that competition can instead motivate some schools to adopt exclusionary practices in the student admissions processes, in order to maintain a privileged position within the school market. However, these types of practices were not identified in the school this author studied in-depth. On the contrary, the school presented the possibility of collaboration with other schools, in large part as a result of its self-confidence and advantageous position in the local education context.

In research conducted in New York City, Jennings (2010) analysed whether schools tried to screen students in the context of an accountability system that informed families about school quality. Methodologically, the research was based on participant observation and interviews with 29 teachers and administrators. This author found that inter-organizational networks between schools, particularly among principals, can provide expressive or instrumental resources for better performance. Particularly in competitive environments, where accountability systems are combined with school choice, networking can encourage school behaviours with contradictory consequences in terms of equity. Specifically, Jennings found that two out of three schools played an active role in the school choice process and attempted to recruit high performing students in order to perform better themselves, increase their social prestige, and to guarantee their own survival in a competitive marketplace. The methods used by schools to select the best students were diverse and included "signaling to families during the recruitment process, using the administrative data to take advantage, creating alliances with junior high schools, and learning the ranking preferences of the students" (Jennings, 2010, p.244). Nonetheless, this author also found that the three principals provided different responses to the same incentives. According to her, these responses were mediated by the principals' own biography and their worldview.

5.3. Teachers' professional identity
Accountability policies are also related to changes in the identities of teachers and more specifically, to their professional identity and their perception of self-efficacy as education professionals (Barret, 2005; Buchanan, 2015; Assaf, 2008; Rex and Nelson, 2004). According to Day (2002), identity comprises a mixture of elements

(social, cultural, biographical and institutional) and varies as a function of the specific role and the circumstances in which teaching is exercised. Some authors have pointed out that the introduction of the culture of performativity in schools through accountability measures has generated important changes in the role, identity and social relations between teachers (Santos, 2004).

Using qualitative methods, some authors have found that accountability policies generate rejection, stress or even pressure to leave the teaching profession, as well as the intensification of work (Sloan, 2006; Katsuno, 2012; Troman, 2000; Valli and Buese, 2007). Similarly, others have found that accountability can have negative effects on teachers' motivation, especially among teachers who work in disadvantaged contexts (Finnigan and Gros, 2007), and is likely to erode trust and collegial relationships among school staff (Jeffrey, 2002; Troman, 2000; Maxcy, 2009).

However, in educational environments with high accountability pressure, collaborative practices can be perceived as an approach for maintaining teachers' sense of professionalism. The way in which extrinsic demands are filtered and recontextualized is a key point for understanding how teachers make sense of accountability policies and their perceived effects on professionalism and autonomy (Locke et al., 2005).

The research strand on the culture of performativity and its effects on the identity and role of teachers has been developed primarily in Anglo-Saxon countries, and particularly in the UK. The work of Ball (2003) and Jeffrey and Troman (2012) are important references in this particular field. Generally, their research explores these aspects by using ethnographic approaches and focuses on the effects of accountability policies in teachers' relationships and subjectivity. In research conducted in England, Gewirtz (2003) showed the main effects of the discourse of performativity on teachers' relationships and autonomy. The discourse of performativity tends to diminish teacher autonomy, intensify teaching work and undermine sociability. Nevertheless, Gewirtz also found different impacts on teachers, depending on schools' socio-economic context. Research conducted in Sweden with a similar angle maintains that the intensification of an audit culture in an already marketized environment can erode teachers' professional autonomy (Lundström, 2015). Sloan (2006) indicates that in some cases, the control mechanisms imposed by accountability policies are likely to reduce the agency of teachers, increasing feelings of stress and affecting their instructional practices. However, the evidence found by Sloan (2006) shows mixed effects and varied responses to accountability. In this sense, some investigations indicate that teachers can also adopt strategies with which to mediate and resist the effects of accountability (Robinson, 2012; Webb, 2002; see also section 4.1 in this paper).

Furthermore, some scholars have explored the effects related to the de-skilling and de-professionalization of teaching work (Jeffrey and Woods, 1996; Woods and Jeffrey, 2002). These studies generally suggest that accountability policies constrain the professional autonomy of teachers, who are pushed into playing a more technical and administrative role (see also Apple, 1989). In New Zealand and England, Locke et al. (2005) found that teachers are experiencing constraints on their professional autonomy as a consequence of increasing external accountability. Day (2002) also showed that accountability reforms were limiting teachers' professional autonomy to make decisions within the classroom. His research also observes the impact of accountability measures on the professional identity formation of teachers. In this regard, Day (2002) concludes that accountability reforms strengthen a kind of technical or instrumental rationality among teachers, at the expense of a more expressive identity attached to a comprehensive vision of education.

The relationship between accountability and teachers' professional identity has also been explored through a comparative research lens. Scholars like Osborn (2006) have attempted to compare the impact of accountability mechanisms on the identity of teachers in different contexts. The evidence presented by this author on the basis of data from Denmark, France and England suggests that the professional identity of teachers is mediated by several elements, including the political economy and the different accountability structures governing the teaching work in each country. According to Day (2002,p.688), factors such as leadership, the school's culture and social composition, and teachers' "sense of vocationalism" can also mediate the effects of accountability reforms on the identity of teachers.

5.4. Opportunistic behaviour and instrumental responses

In certain contexts, test-based accountability policies can generate perverse incentives and encourage so-called opportunistic behaviour on the part of schools, teachers and/or principals. The opportunistic behaviour concept refers to a wide range of practices that include cream-skimming (Jennings, 2010), teaching to the test (Ohemeng and McCall-Thomas, 2013), educational triage (Booher-Jennings, 2005; Perryman et al., 2011; Reback, 2008) and even test cheating (Jacob and Levitt, 2013), which educational actors adopt in order to maximize their results in said tests. These behaviours generally arise when high-stakes testing systems are in place. In these cases, schools and teachers are subjected to higher levels of external pressure to achieve better educational outcomes, particularly due to the threat of sanctions that these systems involve in the case of underperformance.

One type of opportunistic behaviour identified in the literature is 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 and Lauen, 2010, p.429). Ladd and Lauen (2010) analysed the distributional effects of policies based on growth-accountability programmes and status programmes using student-level panel data from North Carolina. These authors found no evidence of the existence of the triaging out of low-achieving students from schools. These results contrast the findings of another study conducted in Texas, which shows that teachers responded to incentives for improving scores by 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) studied schools' responses to standardized testing in Ontario. They conducted interviews with various educational stakeholders in order to understand the pressures and incentives emanating from a standardized testing system. Many of the interviewed teachers argued that the pressure to achieve 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 and McCall-Thomas, 2013). Similarly, Jones and Egley (2004) found that one of the most frequent complaints of teachers (23.3%) in relation to the system of high-stakes testing was the need to spend much of their time preparing their students for a test, i.e., teaching skills and content similar to that which would be included in the test (similar findings can be found in Au, 2007; Hamilton et al., 2007; Linn, 2000; Menken, 2006; Volante, 2004). In the next section, these practices are shown to have the potential for resulting in a reduction of the curriculum and undermining the notion of teaching and learning as a comprehensive process.

5.5. Curriculum and evaluation

Accountability policies also affect the pedagogical and curricular decisions and practices that take place in schools (Au, 2007; Chiang, 2009; Diamond, 2007; Jones and Egley, 2004; Klinger, Maggi and D'Angiulli, 2011; Maier, 2010; Mintrop and Trujillo, 2007; Rustique-Forrester, 2005). However, existing evidence on this particular theme indicates divergent results in terms of what is the direction and the substance of the changes identified.

In the US context, Chiang (2009) conducted an econometric analysis of the effects of accountability reforms on educational production. His research design was based on a discontinuous regression analysis model applied to a large database of Florida schools, which included school-level and student-level data. According to this author, a wide range of instruments is available to schools for improving their results, although he acknowledges that 'threatened' schools may be encouraged to manipulate the system of accountability in order to improve outcomes, without necessarily increasing the real knowledge or strengthening the learning experiences of their students. Nonetheless, ultimately, Chiang (2009) found that underperforming schools usually responded to threats of sanctions by increasing school spending in areas such as curriculum development.

Other scholars have found mixed or inconclusive effects of accountability on school practices and particularly, on the curriculum. Mintrop and Trujillo (2007) contrasted the differential effects of high and low-performing schools in California. They explore the differences in organizational effectiveness, the response of schools in front of the external demands coming from the accountability system and the quality of the experiences of students in these schools. Their research design was based on a multiple-case study and a mixed methods approach that uses different data sources; these included teacher and student surveys, classroom observations, pieces of student work and interviews with teachers and district administrators. Paradoxically, the authors found that there were no relevant differences in organizational effectiveness between these two types of schools. However, they did find that in high-performing schools, there is strong alignment between the curriculum and the demands derived from the accountability system. Thus, high-performing schools tend to focus their efforts on aligning classroom practices and curriculum with the content evaluated in external tests.

Similarly, a qualitative meta-synthesis carried out by Au (2007), shows that high-stakes tests can trigger contradictory trends regarding pedagogical practices and curricular changes. Firstly, this study highlights that the predominate effects derived from high stakes testing policies are closely related to the adoption of teacher-centered pedagogies, the fragmentation of knowledge and the reduction of the curricular content. Nevertheless, despite it was only observed in a few cases, Au (2007) also found that particular test-designs can lead to the very different effects such as student-centered pedagogies or the expansion of the curriculum, although sometimes they are expressed in contradictory ways.

Results on the relationship between accountability and instructional changes are also diverging. For instance, Diamond (2007) found that high-stakes evaluations have a bigger influence on the content of instruction than on pedagogy, whereas Maier (2010) concludes that tests have reduced the school influence on decisions related to the development of teaching and instructional practices.

Authors like Jones and Egley (2004), Klinger et al. (2011) and Rustique-Forrester (2005) found negative effects of accountability systems on delivery of the curriculum. Jones and Egley (2004) explored the perception of teachers around the high-stakes testing programme of Florida. The evidence is based on surveys of 708 teachers from 30 school districts in Florida. A total of 610 teachers answered open-ended questions to justify their previous

response and 115 highlighted the negative effects of high-stakes testing on the curriculum. They pointed out the narrowing effects on the curriculum (13.1%), the lack of depth to the curriculum (4.9%) and the timing of the test (4.8%), which forced them to teach curricular contents without enough time to ensure students' learning. According to Jones and Egley (2004), high-stakes accountability approaches generally encourage teachers to invest more time in the subjects or topics that will be tested. In Ontario, Klinger et al. (2011) found similar results. They emphasize that in the context of test-based accountability systems, there is the danger that non-tested subjects end up being marginalized.

In research conducted in England, Rustique-Forrester (2005) examined the relationship between accountability policies and social exclusion. This research, which was based on 44 interviews with teachers and principals of four schools with high rates of exclusion, found that accountability policies and particularly a high-stakes approach to accountability, when embedded in a market regime (with school choice and test pressures), frequently led to a narrowing of the curriculum. Thus, the non-intended effects of accountability policies are more pronounced in higher-excluding schools, where there are no strong internal structures for supporting collaboration and communication between teachers, or for supporting students' educational needs (see also Jones and Egley, 2004).

For the reasons mentioned thus far, teachers, but also more and more families, tend to be critical of learning-outcomes accountability approaches that are too narrow. Accordingly, they wish to see other sources of information be taken into account in accountability systems including, for example, different elements of the curriculum, teaching methods and school environment, among others (West et al., 2007).

5.6. Spending behaviour

Regarding the influence of accountability on schools' spending behaviour, authors such as Chiang (2009) have observed that low-performing schools increase spending on teacher training, technology and curriculum development as a response to accountability pressures. In relation to the spending behaviour of states, authors like Rubenstein et al. (2008) have examined the influence of accountability policies on the financing of schools. These authors found no relationship between accountability and equity in relation to school funding. However, Rubenstein et al. (2008) found that strong accountability systems generated improvements in the adequacy of the resources acquired by schools. In contrast, Falabella (2014) found that Chilean schools, in response to state incentives attached to the results of the national learning assessment, spent an important part of their resources on hiring expensive consultancy firms that primarily assisted them to conduct test simulations (see also Osses, Bellei and Valenzuela, 2015).

5.7. Inequalities

Another controversial issue in the debate about accountability in education concerns its effects on education inequalities. On the basis of evidence collected in the US, scholars like Carnoy and Loeb (2002) conclude that accountability policies can narrow the so-called racial achievement gaps. In a study conducted in Texas, Skrla, Scheurich, Johnson and Koschoreck (2001) explored the connection between accountability and equity. Many of the interviewees in this study perceived that accountability measures could be useful for reducing inequalities. However, the authors acknowledge that the effects of these policies on equality are often contradictory.

Regarding equity matters and, specifically, in terms of resources distribution and the reduction of achievement gaps between minority or working class students, and their peers several researchers suggest that accountability policies are likely to have mixed effects (Darling-Hammond, 2004) or no significant effects (Lee and Wong, 2004). Lee and Wong (2004), in research conducted in the US, indicate that despite not finding a negative effect of accountability on equity, equity does not occupy a central place in accountability reforms. Consequently, accountability reforms do not bring additional educational resources to key areas such as "per-pupil expenditures, class size, and qualified teachers", which can contribute to educational equity (Lee and Wong, 2004, p.821).

Another group of researchers point out that accountability policies may have unintended effects in terms of increasing class and race inequalities (Andersen, 2008; Barnhardt, Karlan and Khemani, 2009; Lipman, 2002; 2004). One of the most recurrent non-intended effects of performance-based accountability, with serious implications for educational equity, is that it challenges the retention of high quality teachers in low performance schools. In research conducted in North Carolina, Clotfelter, Ladd, Vigdor and Diaz (2004) found that accountability pressures led to higher levels of teacher turnover in schools with high concentrations of low-performing students (see also section 4.3 in this paper). Other researchers have also observed that, in the context of accountability frameworks, the gap between the student scores of white and black students (Hanushek and Raymond, 2004) and students with high and low SES (Andersen, 2008) increases.

Finally, other researchers have highlighted the potential contradictions between the goals set by policies such as NCLB in terms of equity. This is due to the fact that the accountability policies included in regulatory schemes such as NCLB (see Box 1) generate important tensions when aiming at simultaneously promoting, on the one hand, competitiveness between schools and, on the other hand, universalism in the provision of education (Aske, Connolly and Corman, 2013).

In summary, some lessons can be derived from the evidence reviewed above. Firstly, the relationship between accountability policies and its local effects is complex and not necessarily straightforward. Such effects are often mediated by variables of a different nature, which makes it difficult to easily attribute concrete results to accountability policies. Nonetheless, the evidence reviewed points to robust associations between performance-based accountability approaches and its effects in some areas (such as the curriculum and the generation of opportunistic behaviours), and to contradictory trends in others (for example, in learning outcomes, social inequalities and spending behaviours). Undoubtedly, the most controversial areas of research are those related to outputs and outcomes, namely, the effects of accountability on academic results and equity. Contextual and methodological reasons may be able to explain these variations, despite both aspects not always being sufficiently contemplated or debated in existing research. To conclude this section, in Table 3, we synthesise the literature on the effects of accountability in different educational dimensions.

Table 3. Effects of accountability. A review
Source: authors

Dimension	Effects	Sources
Learning outcomes	-Positive impact on students' performance -Mixed or inconclusive impact on students' performance -Negative impact on students' performance	Hanushek & Raymond (2004); Springer (2008); Woessmann (2007) Dee & Jacob (2011); Figlio & Rouse (2006); Ladd (1999); Pandey et al. (2009; 2010) Andersen (2008)
Inter-school relationships	-Increasing competition between schools -Increasing collaboration between schools	Jones & Egley (2004); Ohemeng & McCall-Thomas (2013); Keddie (2014)
Teachers' professional identity	-Erosion of teachers' professional identity and autonomy -De-professionalization, de-skilling, stress, work intensification, anxiety, de-motivation -Erosion of trust and collegial relationships -Introduction of the culture of performativity	Lundström (2015); Gewirtz (2003) Jeffrey & Woods (1996); Day (2002) Jeffrey (2002); Maxcy (2009); Troman (2000) Ball (2003); Jeffrey (2002); Jeffrey and Troman (2012); Gewirtz (2003)
Opportunistic behaviours	-Educational triage -Teaching to the test -Cream-skimming -Cheating	Booher-Jennings (2005); Reback (2008) Ohemeng & Thomas-McCall (2013) Jennings (2010) Jacob & Levitt (2003)
Curriculum and evaluation	-Increasing school spending in curriculum development -Curriculum alignment -Mixed effects on the curriculum -Narrowing the curriculum	Chiang, (2009) Mintrop & Trujillo (2007) Au (2007) Jones & Egley (2004); Klinger, et al. (2011);
Spending behaviours	-Increasing spending on teachers training, educational technologies and curriculum development -Better adequacy of the resources acquired by schools -Hiring expensive local consultancy to conduct test simulations.	Chiang (2009) Rubenstein, et al., (2008) Falabella (2014)
Inequalities	-Narrowing the racial achievement gaps -Mixed or no effects on equity -Increasing teachers' turnover in low-performing schools -Increasing class and race inequalities and achievement gaps	Carnoy & Loeb (2002) Lee & Wong (2004) Clotfelter, Ladd, Vigdor & Diaz (2004) Andersen (2008); Lipman (2002; 2004)

6. Conclusions: research and policy futures in education accountability

Accountability is by no means a new element in education systems. Political and bureaucratic approaches to accountability have a long tradition in many locations, particularly in the most advanced democracies. Nonetheless, there is currently a gradual shift of emphasis toward a form of accountability that focuses on the measurement of learning-outcomes through large-scale testing instruments. As we have seen, several different factors influence this policy trend: pressures on education systems from the global economy, normative emulation dynamics between countries and within international organizations, methodological changes in the measuring of learning outcomes and the spread of New Public Management ideas in the educational sector, among others.

The international dissemination of learning-based accountability policies in education is striking, because countries are adopting these types of accountability measures despite evidence about their effects being highly contradictory. For this reason, existing evidence should be considered with caution and mainly in relation to the contexts where it has been produced. Next, we reflect on the main tensions, absences and contradictions that prevail in existing research on accountability in education and accordingly, we point to issues for future research and debate.

Accountability policies can trigger both **collaboration and competition between schools**. The school context and the role of the principal appear to strategically mediate on whether a school's behaviour becomes more cooperative or competitive within test-based accountability regimes. However, this theme remains underexplored. This is not a minor issue, given the potential benefits that cooperation can have for educational systems. According to Sahlberg (2010, p.59), collaboration between teachers and schools "are essential conditions for system-wide innovation and change". Similarly, Keddie (2015) claims that collaboration can overcome a sense of isolation among teachers, contribute to the development of teachers' professionalism and improve the quality of teaching practices in schools. Thus, further research is needed to understand what types of accountability systems encourage collaboration attitudes within the educational system and under what conditions.

Regarding the **professional identity and autonomy of teachers**, existing research – generally of an ethnographic character – notes that performance-based accountability policies can have more drawbacks than benefits. These policies seem to constrain teachers' professional autonomy, erode social relationships and/or diminish levels of motivation. Nevertheless, there is also a body of literature that acknowledges the positive effects of accountability on teacher motivation, or shows that teachers and principals can negotiate the requirements coming from the accountability regime in a way that does not undermine their professional judgement. On the other hand, the effects of accountability policies on teachers' autonomy are not independent of its design and specific approach. While professional accountability approaches are based on teachers' professional knowledge (Darling-Hammond, 1989) and can empower teachers, providing them, for example, with "licensed autonomy" (Whitty, 2002), managerial, high-stakes or market approaches to accountability can diminish teachers' autonomy, constraining their scope of action in pedagogic and curricular decisions (see Crocco and Costigan,

2007). Again, this is an area requiring more research and that needs to conceive teachers' identities from a more institutional and collective (rather than an exclusively individual) perspective.

Existing literature has identified that performance-based accountability systems with an excessive focus on performance tend to trigger a range of so-called **undesired behaviours** at the school level, including student screening, inflation of results or so-called educational triage. These systems are also conducive to altering the educational process by, for example, teaching to the test and narrowing the scope of the curriculum. These behaviours are generally considered to be part of the unintended consequences of accountability systems; that is, certain accountability systems, unintentionally, create perverse incentives for principals and teachers to act instrumentally in order to improve school performance as rapidly as possible. These behaviours appear so frequently in the literature reviewed that we wonder whether they can be considered as "unintended" or "unexpected" any longer. Undesired or opportunistic behaviours are particularly recurrent in the context of high-stakes testing regimes, where professionals are under high levels of pressure to produce better learning outcomes. Further research is needed in order to understand how accountability systems can be designed in order to minimize these types of undesirable behaviours. In particular, future research should inquire the types of accountability policies that can prevent schools from placing too much emphasis on instrumental responses and instead contribute to promoting more expressive responses, including the development of innovation within the curriculum and in pedagogic terms.

Regarding the impact of accountability reforms on **students' outcomes and educational inequalities**, the results we observed in the existing literature are clearly mixed. Such disparity is in part due to the fact that existing studies have largely been conducted in contexts with different regulatory frameworks and/or have evaluated experiences based on a range of accountability designs. These elements have undoubtedly affected the comparability of local or national case studies. However, studies using the same international databases - such as PISA - have also yielded non-converging results. Differences of a methodological nature are likely responsible for these different results (Dumay and Dupriez, 2014). According to Woessmann (2007, p.493):

> It is clear that this international evidence can only provide the "big picture" of results, revealing broad patterns but not specifics of implementation details. Surely, implementation is crucial with any of the institutional features discussed, and more detailed research is needed to learn how to implement competition, autonomy, and accountability in different circumstances.

Overall, studies focusing on the effects of accountability must make a more significant effort to open the 'black box' of schools in an attempt to contemplate **policy enactment dynamics** within their analytical frameworks. Considering what actually happens inside schools when accountability policies are in place – i.e. how teachers and principals interpret the accountability mandate and translate this mandate into specific educational practices - is a necessary condition for having a more comprehensive and nuanced understanding of the resulting impact of these policies in different contexts.

Likewise, future research should also pay more attention to **contextual variables** of a different nature, including the socio-economic environment and the particular configurations of the local education markets in which schools are embedded. In the literature reviewed, a clear predominance of a school effectiveness approach that gives priority to school-level factors as key mediating variables within accountability regimes is identified. Existing research tends to overlook the social and regulatory context in which accountability practices are embedded and

enacted. In short, more research is needed that attempts to understand the specific circumstances through which accountability can achieve expected results. Geographically speaking, this type of research is particularly needed in southern countries, since accountability in education literature research involving these countries is scarce.

Finally, it is important to note that dominant accountability approaches raise issues of **social justice** in relation to teachers, particularly in terms of defining factors that can be attributed to teachers' actions and behaviour. According to Leithwood and Earl (2000, p.5), accountability systems will be more legitimate if, rather than holding teachers accountable primarily for their students' learning outcomes (which depends on many factors), "the teacher or principal [is held] accountable for making the most productive uses of the resources available to them in an effort to move toward the goal [which] is very different from holding them accountable for actually achieving the goal".

Ultimately, the primary debate is not about whether accountability in education is necessary or not, but about the type of accountability that should be in place. To advance this debate, Lingard et al. (2016, p.15) suggest that we need to move away from a conception of accountability that narrowly focuses on learning outcomes, and re-conceptualize accountability "so that systems and schools are held accountable for their educative and social justice purposes, but in ways that are productive, democratic, and socially just". Similarly, Darling-Hammond et al. (2014) advocate a new accountability paradigm in education that is structured around the needs of students and educators. Among others, this new accountability approach should focus on:

> • creating stronger, more multidimensional ways of evaluating schools and more sophisticated strategies for helping them improve;
> • addressing the opportunity gap that has allowed inequalities in resources to deprive many students of needed opportunities to learn; and
> • developing an infrastructure for professional learning and accountability (...) that allows educators to acquire and share the knowledge and skills they need to enable students to learn. (Darling-Hammond, et al., 2014, p.31)

The results of our review also support the necessity for rethinking the current predominant paradigm of accountability in education. Advancing toward a new and more robust paradigm, however, does not only require debating specific policy instruments, but also, to paraphrase Peter Hall (1993), the main goals that accountability in education is meant to achieve, as well as the nature of the problems that this policy is meant to address. It also requires thinking about the policy processes behind the constitution of accountability systems, and about the subjects that need to be involved in such processes. In this respect, we conclude this paper by drafting a number of principles and notions that could contribute to structuring a more robust accountability approach in education. These are:

a) In terms of their *construction and architecture*, accountability systems should be participatory by involving key education stakeholders in the design and implementation of the system, and multi-directional by combining forms of political and administrative accountability and allowing schools to hold public administrations accountable for their educational commitments, not only the other way around.

b) In terms of the *problems* to be addressed, accountability systems must focus on both education quality and equity issues, and be holistic in the sense that they consider different educational dimensions that include, but move beyond learning outcomes and learning achievement.

c)   In terms of their *primary goals*, accountability systems must be formative rather than exclusively summative and, accordingly, focus further on educational processes. To this purpose, accountability systems should interact with and feed policies about professional development, in-service training and instructional improvement. In addition, challenging persisting educational inequalities and distributing additional educational opportunities should be among the main goals of more robust forms of accountability in education.

Debating each of these elements is a necessary step in addressing some of the most sensitive issues that the predominant model of learning-based accountability raises, and to advance toward new forms of accountability that are better aligned with democratic governance and social justice in education in the post-2015 period.

7. References

Alegre, M. À. 2015. *Què funciona en educació: la pregunta necessària*. Barcelona: Fundació Jaume Bofill & Ivàlua. Available at: 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*, Vol. 86, No. 2, pp.541-558. https://doi.org/10.1111/j.1467-9299.2008.00717.x

Anderson, J. A. 2005. *Accountability in education*. Paris: UNESCO-International Institute for Educational Planning.

Arksey, H., & O'Malley, L. 2005. Scoping studies: towards a methodological framework. *International journal of social research methodology*, Vol. 8, No. 1, pp.19-32. https://doi.org/10.1080/1364557032000119616

Apple, M. W.,. 1989. *Maestros y textos: una economía política de las relaciones de clase y de sexo en educación*. Barcelona: Paidós.

Aske, D. R., Connolly, L. S., & Corman, R. R. 2013. Accessibilty or accountability? The rhetoric and reality of No Child Left Behind. Journal of Economics & Economic Education Research, 14(3), pp.107-118.

Assaf, L. C. 2008. Professional identity of a reading teacher: responding to high-stakes testing pressures. *Teachers and Teaching: theory and practice*, *Vol. 14, No. 3*, pp.239-252. https://doi.org/10.1080/13540600802006137

Assaél, J., Acuña, F., Contreras, P., & Corbalán, F. 2014. Transformaciones en la cultura escolar en el marco de la implementación de políticas de accountability en Chile: Un estudio etnográfico en dos escuelas clasificadas en recuperación. *Estudios pedagógicos*, Vol. 40, No.2, pp. 07-26.

Au, W. 2007. High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher*, Vol. 36, No. 5, pp.258-267. https://doi.org/10.3102/0013189X07306523

Ball, S. J. 2003. The teacher's soul and the terrors of performativity. *Journal of education policy*, Vol. 18, No.2, pp.215-228. https://doi.org/10.1080/0268093022000043065

Ball, S. J., Maguire, M., & Braun, A. 2012. *How schools do policy. Policy Enactment in Secondary Schools.* New York: Routledge.

Barnhardt, S., Karlan, D., & Khemani, S. 2009. Participation in a School Incentive Programme in India. *The Journal of Development Studies*, *Vol. 45, No. 3*, 369-390. https://doi.org/10.1080/00220380802265058

Barroso, J. 2009. A utilização do conhecimento em política: o caso da gestão escolar em Portugal, Educaçao & Sociedade, Vol. 30, No. 109, pp.987-1007. https://doi.org/10.1590/S0101-73302009000400004

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*, Vol. 1, pp. 1-20.

Booher-Jennings, J. 2005. Below the bubble:"Educational triage" and the Texas accountability system. *American educational research journal*, Vol. 42, No. 2, pp. 231-268. https://doi.org/10.3102/00028312042002231

Bovens, M. 2007. Analysing and assessing accountability: a conceptual framework1. *European law journal*, Vol. 13, No.4, pp. 447-468. https://doi.org/10.1111/j.1468-0386.2007.00378.x

Braun, H., Chapman, L., & Vezzu, S. 2010. The black-white achievement gap revisited. *education policy analysis archives*, Vol. 18, No. 21, pp.1-95.

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, No. 71. *OECD Publishing.* DOI: http://dx.doi.org/10.1787/5k9fdfqffr28-en

Brooke, N. 2006. The future of educational accountability policies in Brazil. *Cadernos de Pesquisa*, Vol. *36*, No. 128, pp. 377-401. https://doi.org/10.1590/S0100-15742006000200006

Bruns, Barbara, Filmer Deon, and Harry A. Patrinos. 2011. Making Schools Work: New Evidence on Accountability Reforms. Washington, DC: World Bank Group. https://doi.org/10.1596/978-0-8213-8679-8

Buchanan, R. 2015. Teacher identity and agency in an era of accountability.*Teachers and Teaching*, Vol. 21, No. 6, pp. 700-719. https://doi.org/10.1080/13540602.2015.1044329

Carnoy, M., & Loeb, S. 2002. Does external accountability affect student outcomes? A cross-state analysis. *Educational evaluation and policy analysis*, Vol. 24, No. 4, pp. 305-331. https://doi.org/10.3102/01623737024004305

Carnoy, M., & Rhoten, D. 2002. What does globalization mean for educational change? A comparative approach. *Comparative education review*, Vol. 46, No. 1, pp.1-9. https://doi.org/10.1086/324053

Carnoy, M. 2016. Educational policies in the face of globalization: Whither the Nation State? In K. Mundy, A. Green, R. Lingard, & A. Verger (Eds.), *Handbook of global policy and policy-making in education* (pp. 27–42). West Sussex, UK: Wiley-Blackwell. https://doi.org/10.1002/9781118468005.ch1

Chiang, H. 2009. How accountability pressure on failing schools affects student achievement. *Journal of Public Economics*, *Vol. 93, No. 9*, pp.1045-1057. https://doi.org/10.1016/j.jpubeco.2009.06.002

Clotfelter, C. T., Ladd, H. F., Vigdor, J. L., & Diaz, R. A. 2004. Do school accountability systems make it more difficult for low-performing schools to attract and retain high-quality teachers?. *Journal of Policy Analysis and Management*, Vol. 23, No. 2, pp.251-271. https://doi.org/10.1002/pam.20003

Codd, J. 2005. Teachers as 'managed professionals' in the global education industry: The New Zealand experience. *Educational review*, Vol. 57, No.2, pp. 193-206. https://doi.org/10.1080/0013191042000308369

Crocco, M. S., & Costigan, A. T. 2007. The narrowing of curriculum and pedagogy in the age of accountability: Urban educators speak out. *Urban Education*, Vol. 42, No. 6, pp. 512–535. https://doi.org/10.1177/0042085907304964

Darling-Hammond, L. 1989. Accountability for professional practice. *Teachers College Record*, Vol. 91, No. 1, pp. 59–80.

Darling-Hammond, L. 2004. Standards, accountability, and school reform. *The Teachers College Record*, Vol. 106, No. 6, pp.1047-1085. https://doi.org/10.1111/j.1467-9620.2004.00372.x

Darling-Hammond, L., Wilhoit, G., & Pittenger, L. 2014. Accountability for college and career readiness: Developing a new paradigm. *Education Policy Analysis Archives*, Vol. 22, No. 86, pp.2-33.http://dx.doi.org/10.14507/epaa.v22n86.2014.

Davis, D.R, Pool, J.E., Mits-Cash, M. 2000. Issues in implementing a new teacher assessment system in a large urban school district: Results of a qualitative field study. *Journal of Personnel Evaluation in Education*, Vol. 14, No. 4, pp. 285-306. https://doi.org/10.1023/A:1011139203370

Day, C. 2002. School reform and transitions in teacher professionalism and identity. *International journal of educational research*, Vol. 37, No. 8, pp. 677-692. https://doi.org/10.1016/S0883-0355(03)00065-X

Dee, T. S., & Jacob, B. 2011. The impact of No Child Left Behind on student achievement. *Journal of Policy Analysis and management*, *Vol. 30, No. 3*, pp. 418-446. https://doi.org/10.1002/pam.20586

Diamond, J., & Spillane, J. 2004. High-stakes accountability in urban elementary schools: challenging or reproducing inequality?. *The Teachers College Record*, Vol. 106, No. 6, pp. 1145-1176. https://doi.org/10.1111/j.1467-9620.2004.00375.x

Diamond, J. B. 2007. Where the rubber meets the road: Rethinking the connection between high-stakes testing policy and classroom instruction.*Sociology of Education*, Vol. 80, No. 4, pp. 285-313. https://doi.org/10.1177/003804070708000401

Douglas J. 2005. *Teacher professionalism and education transformation*. Braamfontein: Centre for Education Policy Development (CEPD).

Dumay, X., & Dupriez, V. 2014. Educational quasi-markets, school effectiveness and social inequalities. *Journal of Education Policy*, Vol. 29, No. 4, pp. 510-531. https://doi.org/10.1080/02680939.2013.850536

Elstad, E., Nortvedt, G.A., Turmo, A. 2009. The norwegian assessment system: An accountability perspective. *Cadmo*, Vol. 17, No. 2, pp. 89-103.

Eyben, R. 2008. Power, Mutual Accountability and Responsibility in the Practice of International Aid: A Relational Approach. IDS Working Paper 305

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* (ahead-of-print), pp.1-18.

Falabella, A.; de la Vega, F. 2016. Políticas de Responsabilización por Desempeño Escolar: Un Debate a Partir de la Literatura Internacional y el Caso Chileno. *Revista de Estudios Públicos*.

Figlio, D. N., & Rouse, C. E. 2006. Do accountability and voucher threats improve low-performing schools?. *Journal of Public Economics*, *Vol. 90, No. 1*, pp. 239-255. https://doi.org/10.1016/j.jpubeco.2005.08.005

Finnigan, K. S., & Gross, B. 2007. Do accountability policy sanctions influence teacher motivation? Lessons from Chicago's low-performing schools. *American Educational Research Journal*, *Vol. 44, No. 3*, pp. 594-630. https://doi.org/10.3102/0002831207306767

Finnigan, K.S. 2010. Principal Leadership and Teacher Motivation under High-Stakes Accountability Policies, *Leadership and Policy in Schools*, Vol. 9, No.2, pp. 161-189. https://doi.org/10.1080/15700760903216174

Finnigan, K.S. 2012. Principal Leadership in Low-Performing Schools: A Closer Look Through the Eyes of Teachers. *Education and Urban Society*, *Vol. 44, No. 2*, pp. 183–202. https://doi.org/10.1177/0013124511431570

Gaddis, S. M., & Lauen, D. L. 2014. School accountability and the black–white test score gap. *Social science research*, *44*, pp. 15-31. https://doi.org/10.1016/j.ssresearch.2013.10.008

Gewirtz, S. 2003. *The managerial school: Post-welfarism and social justice in education*. New York: Routledge.

Grek, S., Lawn, M., Lingard, B., Ozga, J., Rinne, R , Segerholm, C, & Simola, H. 2009. National policy brokering and the construction of the European Education Space in England, Sweden, Finland and Scotland. *Comparative Education*, *Vol. 45, No. 1*, pp. 5-21. https://doi.org/10.1080/03050060802661378

Gunter, H. M., Grimaldi, E., Hall, D., & Serpieri, R. 2016. *New Public Management and the Reform of Education: European Lessons for Policy and Practice.* New York: Routledge.

Hall, P. A. 1993. Policy paradigms, social learning, and the state: the case of economic policymaking in Britain. *Comparative politics*, Vol. 25, No. 3, 275-296. https://doi.org/10.2307/422246

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*. Rand Corporation.

Hangartner, J., & Svaton, C. J. 2013. From autonomy to quality management: NPM impacts on school governance in Switzerland. *Journal of Educational Administration and History*, Vol. 45, No. 4, pp. 354-369. https://doi.org/10.1080/00220620.2013.822352

Hanushek, E. A., Kain, J. F., Markman, J. M., & Rivkin, S. G. 2003. Does peer ability affect student achievement?. Journal of applied econometrics, Vol. 18, No. 5, pp. 527-544. https://doi.org/10.1002/jae.741

Hanushek, E. A., & Raymond, M. E. 2004. The effect of school accountability systems on the level and distribution of student achievement. *Journal of the European Economic Association*, Vol. 2, No. 2/3, pp. 406-415. https://doi.org/10.1162/154247604323068096

Hardy, I. 2014. A logic of appropriation: enacting national testing (NAPLAN) in Australia. *Journal of Education Policy*, Vol. 29, No. 1, pp. 1-18. https://doi.org/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,* Vol. 14, No. 2*,* pp. 113-138. https://doi.org/10.1007/s10833-012-9206-1

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.* New York: Routledge*. pp. 107-124*

Hursh, D. 2005. The growth of high-stakes testing in the USA: accountability, markets and the decline in educational equality. British Educational Research Journal, Vol. 31, No.5, pp.605-622. https://doi.org/10.1080/01411920500240767

Jacob, B. A., & Levitt, S. D. 2003. *Rotten apples: An investigation of the prevalence and predictors of teacher cheating* Working paper No. 9413. National Bureau of Economic Research. http://www.nber.org/papers/w9413

Jeffrey, B. 2002. Performativity and primary teacher relations. *Journal of education policy*, Vol. 17, No. 5, pp. 531-546. https://doi.org/10.1080/02680930210158302

Jeffrey, B. & Troman, G. 2012. *Performativity in UK Education. Ethnographic cases of its effects, agency and reconstructions*. Essex: E&E Publishing.

Jeffrey, B., & Woods, P. 1996. Feeling deprofessionalised: The social construction of emotions during an OFSTED inspection. *Cambridge Journal of Education*, Vol. *26, No.* 3, pp. 325-343. https://doi.org/10.1080/0305764960260303

Jennings, J. L. 2010. School choice or schools' choice? Managing in an era of accountability. *Sociology of Education*, Vol. *83*, No. 3, pp. 227-247. https://doi.org/10.1177/0038040710375688

Jennings, J., & Sohn, H. 2014. Measure for Measure How Proficiency-based Accountability Systems Affect Inequality in Academic Achievement. *Sociology of Education*, Vol. 87, No. 2, pp. 125-141. https://doi.org/10.1177/0038040714525787

Jones, B. D., & Egley, R. J. 2004. Voices from the frontlines: Teachers' perceptions of high-stakes testing. *Education policy analysis archives*, Vol. *12, No.* 39, 1-34.

Kamens, D. H., & Benavot, A. 2011. National, regional and international learning assessments: Trends among developing countries, 1960–2009.*Globalisation, societies and education*, Vol. *9*, No. 2, pp. 285-300. https://doi.org/10.1080/14767724.2011.577337

Katsuno, M. 2012. Teachers' professional identities in an era of testing accountability in Japan: the case of teachers in low-performing schools.*Education Research International*, Vol. *2012, pp. 1-8*. https://doi.org/10.1155/2012/930279

Keddie, A. 2014. 'It's like Spiderman… with great power comes great responsibility': school autonomy, school context and the audit culture. *School Leadership & Management*, Vol. 34, No. 5, pp. 502-517. https://doi.org/10.1080/13632434.2014.938040

Keddie, A. 2015. School autonomy, accountability and collaboration: a critical review. *Journal of educational administration and history*, Vol. *47, No.* 1, pp.1-17. https://doi.org/10.1080/00220620.2015.974146

Klinger, D. A., Maggi, S., & D'Angiulli, A. 2011. School Accountability and Assessment: Should We Put the Roof Up First?. *The Educational Forum,* Vol. 75, No. 2, pp.114-128. https://doi.org/10.1080/00131725.2011.552671

Kostogriz, A., & Doecke, B. 2011. Standards-based accountability: Reification, responsibility and the ethical subject. *Teaching Education*, Vol. 22, No. 4, pp.397-412. https://doi.org/10.1080/10476210.2011.587870

Koyama, J. 2013. Global scare tactics and the call for US schools to be held accountable. *American Journal of Education*, Vol. 120, No.1, pp. 77-99. https://doi.org/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*, Vol. *18*, No. 1, pp.1-16. https://doi.org/10.1016/S0272-7757(97)00044-7

Ladd, H. F., & Lauen, D. L. 2010. Status versus growth: The distributional effects of school accountability policies. *Journal of Policy Analysis and Management*, Vol. *29*, No. 3, pp.426-450. https://doi.org/10.1002/pam.20504

Lee, J., & Wong, K. K. 2004. The impact of accountability on racial and socioeconomic equity: Considering both school resources and achievement outcomes. *American educational research journal*, Vol. 41, No. 4, pp.797-832. https://doi.org/10.3102/00028312041004797

Leithwood, K., & Earl, L. 2000. Educational accountability effects: An international perspective. *Peabody Journal of Education*, Vol. 75, No. 4, pp. 1-18. https://doi.org/10.1207/S15327930PJE7504_1

Leicht, KT, Walter, T. Sainsaulieu, I., & Davies, S. 2009. New Public Management and New Professionalism across Nations and Contexts, *Current Sociology* Vol. 57, No. 4, pp.581-605. https://doi.org/10.1177/0011392109104355

Lingard, B. 2010. Policy borrowing, policy learning: Testing times in Australian schooling.Critical Studies in Education, Vol. 51, No. 2, pp. 129-147. https://doi.org/10.1080/17508481003731026

Lingard, B., Martino, W., Rezai-Rashti, G., & Sellar, S. 2016. *Globalizing Educational Accountabilities*. New York: Routledge.

Linn, R. 2000. Assessments and accountability. *Educational Researcher*, Vol. 29, No. 2, pp. 4-16. https://doi.org/10.3102/0013189X029002004

Lipman, P. 2002. Making the global city, making inequality: The political economy and cultural politics of Chicago school policy. *American Educational Research Journal*, Vol. 39, No. 2, pp.379-419. https://doi.org/10.3102/00028312039002379

Lipman, P. 2004. *High stakes education: Inequality, globalization, and urban school reform*. New York: RoutledgeFalmer. https://doi.org/10.4324/9780203465509

Locke, T., Vulliamy, G., Webb, R., & Hill, M. 2005. Being a 'professional'primary school teacher at the beginning of the 21st century: A comparative analysis of primary teacher professionalism in New Zealand and England. *Journal of education policy*, Vol. 20, No. 5, pp. 555-581. https://doi.org/10.1080/02680930500221784

Lundström, U. 2015. Teacher autonomy in the era of New Public Management. *Nordic Journal of Studies in Educational Policy*, Vol. 1, No.2, pp. 73-85. https://doi.org/10.3402/nstep.v1.28144

Maier, U. 2010. Effects of test-based reports on school and on teaching - Are international findings applicable to comparative tests in the German-speaking countries?. *Zeitschrift fur Padagogik*, Vol. 56, No. 1, pp.112-128.

Maroy, C.; Voisin, A. 2013. As transformações recentes das políticas de accountability na educação: desafios e incidências das ferramentas de ação pública. *Educação & Sociedade*, Vol. 34, No. 124, pp. 1-24. https://doi.org/10.1590/S0101-73302013000300012

Martinsson, J. 2011. Global Norms: Creation, Diffusion, and Limits. Washington, DC: World Bank/Communication for Governance and Accountability,
http://siteresources.worldbank.org/EXTGOVACC/Resources/FinalGlobalNormsv1.Pdf (accessed, December 15, 2016). https://doi.org/10.1596/26891

Maxcy, B. 2009. New public management and district reform: Managerialism and deflection of local leadership in a Texas school district. *Urban Education*, Vol. 44, No. 5, pp. 489-521. https://doi.org/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, March 2007,* Vol. 44, No. 1, pp. 77–111. https://doi.org/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,* Vol. 30, No. 2, pp. 521–46. https://doi.org/10.1080/15235882.2006.10162888

Meyer, H. D., & Benavot, A. (Eds.). 2013. *PISA, power, and policy: The emergence of global educational governance*. Oxford: Symposium Books. https://doi.org/10.15730/books.85

Mintrop, H. 2003. The limits of sanctions in low-performing schools: A study of Maryland and Kentucky schools on probation. *Education policy analysis archives*, Vol.10, No. 3, pp. 1-30. https://doi.org/10.14507/epaa.v11n3.2003

Mintrop, H., & Trujillo, T. 2007. The practical relevance of accountability systems for school improvement: A descriptive analysis of California schools.*Educational Evaluation and Policy Analysis*, Vol. 29, No. 4, pp.319-352. https://doi.org/10.3102/0162373707309219

Møller, J., Skedsmo, G. 2013. Modernising education: New Public Management reform in the Norwegian education system. Journal of Educational Administration and History, Vol. 45, No. 4, pp. 336-353. https://doi.org/10.1080/00220620.2013.822353

Montecinos, C., Madrid, R., Fernández, M. B., & Ahumada, L. 2014. A goal orientation analysis of teachers' motivations to participate in the school self-assessment processes of a quality assurance system in Chile. *Educational Assessment, Evaluation and Accountability*, Vol. 26, No. 3, pp.241-261. https://doi.org/10.1007/s11092-014-9190-5

Montecinos, C., Ahumada, L., Galdames, S., Campos, F., & Leiva, M. V. 2015. Targets, threats and (dis)trust: The managerial troika for public school principals in Chile. *Education policy analysis archives*, Vol. 23, No. 87, pp. 1-23. https://doi.org/10.14507/epaa.v23.2083

Moos, L. 2013. School leadership in a contradictory world. Revista de Investigacion Educativa, Vol. 31, No. 1, pp. 15-29. https://doi.org/10.6018/rie.31.1.162511

Mosoge, M. J., & Pilane, M. W. 2014. Performance management: the neglected imperative of accountability systems in education. *South African Journal of Education*, Vol. 34, No. 1, pp.1-18. https://doi.org/10.15700/201412120947

Müller, J. & F. Hernández. 2010. On the geography of accountability Comparative analysis of teachers' experiences across seven European Countries. *Journal of Educational Change*, Vol. 11, pp.307-322. https://doi.org/10.1007/s10833-009-9126-x

OECD. 2011. "School autonomy and accountability: Are they related to student performance?" *PISA in Focus, No, 9,* OECD Publishing, Paris.

OECD. 2012. *Equity and Quality in Education - Supporting Disadvantaged Students and School*. Paris: OECD

OECD. 2013. *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*. Paris: OECD

Ohemeng, F., & McCall-Thomas, E. 2013. Performance management and "undesirable" organizational behaviour: Standardized testing in Ontario schools. *Canadian Public Administration*, Vol. *56*, No. 3, pp. 456-477. https://doi.org/10.1111/capa.12030

Osborn, M. 2006. Changing the context of teachers' work and professional development: A European perspective. *International journal of educational research*, Vol. *45*, No. 4, pp. 242-253. https://doi.org/10.1016/j.ijer.2007.02.008

Osses, A., Bellei, C., & Valenzuela, J. P. 2015. External technical support for school improvement: critical issues from the Chilean experience. *Journal of Educational Administration and History*, Vol. 47, No. 3, pp. 272-293. https://doi.org/10.1080/00220620.2015.1038699

Ozga, J. 2013. Accountability as a policy technology: Accounting for education performance in Europe. International Review of Administrative Sciences, Vol. 79, No. 2, pp. 292-309. https://doi.org/10.1177/0020852313477763

Palmer, D. and V.S. Rangel. 2011. High Stakes Accountability and Policy Implementation: Teacher Decision Making in Bilingual Classrooms in Texas. *Educational Policy*, Vol. 25, No. 4, pp. 614–647. https://doi.org/10.1177/0895904810374848

Patrinos, H. A., & Fasih, T. 2009. *Decentralized decision-making in schools: The theory and evidence on school-based management*. Washington DC: World Bank Publications. https://doi.org/10.1596/978-0-8213-7969-1

Pandey, P., Goyal, S., & Sundararaman, V. 2009. Community participation in public schools: impact of information campaigns in three Indian states. *Education Economics*, Vol. 17, No. 3, pp.355-375. https://doi.org/10.1080/09645290903157484

Pandey, P., Goyal, S., & Sundararaman, V. 2010. Public participation, teacher accountability and school outcomes in three states. *Economic and Political Weekly*, Vol. 45, No. 24, pp. 75-83.

Pavon, F. Y. 2008. Improving Educational Quality in Honduras: Building a demanddriven education market. *Journal of Public and International Affairs*, Vol. 19, No. 1, pp. 194-213.

Perryman, J., S. Ball , M. Maguire & A. Braun. 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, Vol. 59, No.2, pp.179-195. https://doi.org/10.1080/00071005.2011.578568

Pollitt, C., & Bouckaert, G. 2011. *Public Management Reform: A comparative analysis-new public management, governance, and the Neo-Weberian state*. Oxford: Oxford University Press

Powers, J. M. 2003. An analysis of performance-based accountability: Factors shaping school performance in two urban school districts. *Educational Policy*, Vol. 17, No. 5, pp.558-585. https://doi.org/10.1177/0895904803256789

Ranson, S. 2003. Public accountability in the age of neo-liberal governance, *Journal of Education Policy*, Vol. 18, No. 5, pp.459-480. https://doi.org/10.1080/0268093032000124848

Reback, R. 2008. Teaching to the rating: School accountability and the distribution of student achievement. *Journal of Public Economics*, Vol. 92, No. 5, pp.1394-1415. https://doi.org/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*, Vol. *6*, No. 3, pp.207-241. https://doi.org/10.1257/pol.6.3.207

Resnik, J. 2011. The construction of a managerial education discourse and the involvement of philanthropic entrepreneurs: The case of Israel, *Critical Studies in Education*, Vol. 52, No. 3, pp.251-266. https://doi.org/10.1080/17508487.2011.604075

Rex, L., & Nelson, M. 2004. How teachers' professional identities position high-stakes test preparation in their classrooms. *The Teachers College Record*, Vol. 106, No.6, pp.1288-1331. https://doi.org/10.1111/j.1467-9620.2004.00380.x

Robinson, S. 2012. Constructing teacher agency in response to the constraints of education policy: adoption and adaptation. *Curriculum Journal*, Vol. 23, No. 2, pp.231-245. https://doi.org/10.1080/09585176.2012.678702

Roderick, M., Jacob, B. A., & Bryk, A. S. 2002. The impact of high-stakes testing in Chicago on student achievement in promotional gate grades. *Educational evaluation and policy analysis*, Vol. 24, No. 4, pp.333-357. https://doi.org/10.3102/01623737024004333

Ryan, J. E. 2004. Perverse Incentives of the No Child Left behind Act, The. *NYUL Review.*, Vol. 79, pp.932-989.

Rubenstein, R., Ballal, S., Stiefel, L., & Schwartz, A. E. 2008. Equity and accountability: The impact of state accountability systems on school finance. *Public budgeting & finance*, Vol. 28, No. 3, pp. 1-22. https://doi.org/10.1111/j.1540-5850.2008.00908.x

Rustique-Forrester, E. 2005. Accountability and the pressures to exclude: A cautionary tale from England. *education policy analysis archives*, Vol. 13, No. 26, pp.1-41.

SADTU. 2011. *SADTU rejects Teacher Performance Appraisal in current form*. Declaration of the National General Council of the South African Democratic teachers Union (SADTU). Available at http//:www.politicsweb.co.za.

Sahlberg, P. 2010. Rethinking accountability in a knowledge society. *Journal of Educational Change*, Vol. 11, No. 1, pp. 45-61. https://doi.org/10.1007/s10833-008-9098-2

Santos, L,L,C. 2004. Formação de professores na cultura do desempenho. *Educ. Soc*, Vol. 25, No. 89, pp.1145-1157. https://doi.org/10.1590/S0101-73302004000400004

Sayed, Y. 2011. After 2015: time for an education quality goal?. *Compare*, Vol. 41, No. 1, pp.129-130. https://doi.org/10.1080/03057925.2011.534847

Skrla, L., Scheurich, J. J., Johnson Jr, J. F., & Koschoreck, J. W. 2001. Accountability for equity: can state policy leverage social justice?. *International Journal of Leadership in Education*, Vol. 4, No. 3, pp.237-260. https://doi.org/10.1080/13603120120176

Sloan, K. 2006. Teacher identity and agency in school worlds: beyond the all-good/all-bad discourse on accountability-explicit curriculum policies. *Curriculum Inquiry*, Vol. *36*, No. 2, pp.119-152. https://doi.org/10.1111/j.1467-873X.2006.00350.x

Smith, M.L, L. Miller-Kahn, W. Heinecke, and P.F. Jarvis. 2004. *Political Spectacle and the Fate of American Schools*. New York: Routledge. https://doi.org/10.4324/9780203465554

Spillane, J., & A.W. Kenney. 2012. School Administration in a Changing Education Sector: The U.S. Experience. *Journal of Educational Administration,* Vol. 50, No.5,pp.541–61. https://doi.org/10.1108/09578231211249817

Stecher, B., Hamilton, L.S., & Gonzalez, G. 2003. *Working smarter to leave no child left behind.* Santa Monica, CA: Rand Corporation.

Springer, M. G. 2008. The influence of an NCLB accountability plan on the distribution of student test score gains. *Economics of Education Review*, Vol. *27*, No. 5, pp.556-563. https://doi.org/10.1016/j.econedurev.2007.06.004

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).

Tolofari, S. 2005. New public management and education. Policy Futures in Education, Vol. 3, No. 1, pp. 75–89. https://doi.org/10.2304/pfie.2005.3.1.11

Troman, G. 2000. Teacher stress in the low-trust society. *British journal of sociology of education*, Vol. 21, No. 3, pp.331-353. https://doi.org/10.1080/713655357

UNESCO. 2009. *EFA Global Monitoring Report 2009. Overcoming inequality: why governance matters*. Paris: UNESCO

UNESCO. 2015. Framework for Action Education 2030: Towards inclusive and equitable quality education and lifelong learning for all. Available at: http://www.uis.unesco.org/Education/Documents/wef-framework-for-action.pdf

Van Zanten, A. 2002. Educational change and new cleavages between head teachers, teachers and parents: global and local perspectives on the French case, *Journal of Education Policy*, Vol. 17, No. 3, pp.289-304. https://doi.org/10.1080/02680930210127568

Valli, L., & Buese, D. 2007. The changing roles of teachers in an era of high-stakes accountability. *American Educational Research Journal*, Vol.44, No.3, pp. 519-558. https://doi.org/10.3102/0002831207306859

Verger, A., & Curran, M. 2014. New public management as a global education policy: its adoption and re-contextualization in a Southern European setting. *Critical studies in education*, Vol. 55, No. 3, 253-271. https://doi.org/10.1080/17508487.2014.913531

Verger, A., & Normand, R. 2015. New public management and education: theoretical and conceptual elements for the study of a global education reform model. *Educação & Sociedade*, Vol. *36*, No. 132, pp.599-622. https://doi.org/10.1590/ES0101-73302015152799

Vesely, A. 2012. The Institutionalisation of Non-responsibility, Efficiency or Conformity? Organisational Reform of Public Services Based on Accountability Theory. *Sociologicky Casopis-Czech Sociological Review*, Vol. 48, No. 4, pp.757-784

Volante, L. 2004. Teaching to the Test: What Every Educator and Policy-Maker Should Know. *Canadian Journal of Educational Administration and Policy*.

Webb, P. T. 2002. Teacher power: The exercise of professional autonomy in an era of strict accountability. *Teacher Development*, Vol. 6, No.1, pp. 47-62. https://doi.org/10.1080/13664530200200156

West, A. P. Mattei & J. Roberts. 2011. Accountability and Sanctions in English Schools, *British Journal of Educational Studies*, No. 59, Vol. 1, pp.41-62. https://doi.org/10.1080/00071005.2010.529416

Whitty, G. 2002. *Making sense of education policy: Studies in the sociology and politics of education*. London: Sage Publications Ltd.

Woods, P., & Jeffrey, B. 2002. The reconstruction of primary teachers' identities. *British journal of sociology of education*, Vol. 23, No. 1, pp. 89-106. https://doi.org/10.1080/01425690120102872

Woessmann, L. 2007. International evidence on school competition, autonomy, and accountability: A review. *Peabody journal of education*, No. 82, Vol. 2-3, pp.473-497. https://doi.org/10.1080/01619560701313176

World Bank. 2006. *Social Accountability Sourcebook*. Washington DC: World Bank

World Bank. 2011. *Learning for All: Investing in People's Knowledge and Skills to Promote Development*. Washington DC: World Bank.

World Bank. 2015. What Matters Most for School Autonomy and Accountability: A Framework Paper. *SABER Working Paper Series, 9.*

World Education Forum. 2015. *Incheon Declaration: Education 2030: Towards inclusive and equitable quality education and lifelong learning for all.* Available at: http://unesdoc.unesco.org/images/0023/002338/233813M.pdf

8. Appendix

Box A1. Keywords of the scoping review search

Effects

1)   TITLE-ABS-KEY ("education"OR "education reform"OR "School")AND TITLE-ABS-KEY("Effective school research"OR "Efficiency"OR "Effectiveness"OR "Effects" OR "Outcomes")AND TITLE-ABS-KEY("New Public Management"OR "NPM"OR "School autonomy"OR "Institutional autonomy"OR "Accountability"OR "Merit-based pay" OR "School-based management"OR "Public management reform") AND NOT ("Higher education") AND (LIMIT-TO(SUBJAREA,"SOCI") OR LIMIT-TO(SUBJAREA,"BUSI") OR LIMIT-TO(SUBJAREA,"ECON") OR LIMIT-TO(SUBJAREA,"ARTS")) AND (LIMIT-TO(SUBJAREA,"SOCI") OR LIMIT-TO(SUBJAREA,"BUSI") OR LIMIT-TO(SUBJAREA,"ECON"))

Enactment

2)   TITLE-ABS-KEY ("education"OR "education reform"OR "School")AND TITLE-ABS-KEY("Enactment"OR "Implementation"OR "Public Service Motivation"OR "Teachers' motivation")AND TITLE-ABS-KEY("New Public Management"OR "NPM"OR "School autonomy"OR "Institutional autonomy"OR "Accountability"OR "Merit-based pay" OR "School-based management"OR "Public management reform") AND NOT ("Higher education") AND (LIMIT-TO(SUBJAREA,"SOCI") OR LIMIT-TO(SUBJAREA,"BUSI") OR LIMIT-TO(SUBJAREA,"ECON") OR LIMIT-TO(SUBJAREA,"ARTS")) AND (LIMIT-TO(SUBJAREA,"SOCI") OR LIMIT-TO(SUBJAREA,"BUSI") OR LIMIT-TO(SUBJAREA,"ECON"))

Diffusion

3)   TITLE-ABS-KEY ("education"OR "education reform"OR "School")AND TITLE-ABS-KEY("International organi*ations"OR "OECD"OR "PISA"OR "PUMA"OR "world bank"OR "European Commission"OR "policy difussion"OR "policy adoption"OR "agenda-setting"OR "global education agenda"OR "policy entepreneur"OR "policy learning"OR "emulation"OR "dissemination"OR "globali*ation")AND TITLE-ABS-KEY("New Public Management"OR "NPM"OR "School autonomy"OR "Institutional autonomy"OR "Accountability"OR "Merit-based pay" OR "School-based management"OR "Public management reform") AND NOT ("Higher education") AND (LIMIT-TO(SUBJAREA,"SOCI") OR LIMIT-TO(SUBJAREA,"BUSI") OR LIMIT-TO(SUBJAREA,"ECON") OR LIMIT-TO(SUBJAREA,"ARTS")) AND (LIMIT-TO(SUBJAREA,"SOCI") OR LIMIT-TO(SUBJAREA,"BUSI") OR LIMIT-TO(SUBJAREA,"ECON"))