



HAL
open science

THINK PIECE ON ACCOUNTABILITY

Christian Maroy, Annelise Voisin

► **To cite this version:**

Christian Maroy, Annelise Voisin. THINK PIECE ON ACCOUNTABILITY: Background paper prepared for the 2017/8 Global Education Monitoring Report. [Research Report] UNESCO. 2017. halshs-01705982

HAL Id: halshs-01705982

<https://shs.hal.science/halshs-01705982>

Submitted on 10 Feb 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



United Nations
Educational, Scientific and
Cultural Organization



Global
Education
Monitoring
Report

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

Accountability in education: Meeting our commitments

UNESCO – Think piece on accountability

This paper was commissioned by the Global Education Monitoring Report as background information to assist in drafting the 2017/8 GEM Report, Accountability in education: Meeting our commitments. It has not been edited by the team. The views and opinions expressed in this paper are those of the author(s) and should not be attributed to the Global Education Monitoring Report or to UNESCO. The papers can be cited with the following reference: "Paper commissioned for the 2017/8 Global Education Monitoring Report, Accountability in education: Meeting our commitments". For further information, please contact gemreport@unesco.org.

Abstract

AUTHORS: CHRISTIAN MAROY¹ & ANNE LISE VOISIN²
CANADA RESEARCH CHAIR IN EDUCATION POLICIES, UNIVERSITY OF MONTREAL

Accountability and accountability policies have known international expansion for the last decades. In a context of a growing concern for the quality of education accountability has become a centerpiece of educational reforms, supposed to improve effectiveness and equity in education systems. Despite the increasing popularity of accountability policies in education the link between these policies and the improvement of pupils achievement remains unclear. Moreover accountability policies take several forms and rely on various tools depending on the contexts where they are implemented. This paper aims at providing some insights regarding the development of accountability policies in education. It first discusses the polysemic character of this notion to focus thereafter on the development of accountability policies in education systems. After proposing a typology of accountability policies and tools, we discuss the main research finding regarding the effects of accountability policies on effectiveness and equity in education systems to discuss then central issues for further research.

¹ christian.maroy@umontreal.ca, full Professor in education policy and sociology of education

² annelise.voisin@umontreal.ca, research assistant for the Canada research chair in education policies, PHD candidate in comparative education

Outline

Abstract.....	2
1. Introduction.....	4
2. Accountability and accountability policies in education	4
2.1 A multi-layered concept.....	4
2.2 ‘New’ accountability in education.....	5
3. Accountability policies in education, diversity of tools and rationales.....	8
3.1 Typology of accountability policies and tools: main dimensions	8
3.2 Four approaches to regulation and accountability.....	9
3.2.1 Regulation through strong accountability	10
3.2.2 Regulation through neo-bureaucratic accountability	11
3.2.3 Regulation approach through reflexive responsabilization and accountability	11
3.2.4 Regulation through soft accountability.....	12
4. Accountability in education: effects, side effects and normative implications –main research results.	13
4.1 Regulation through strong accountability and high-stakes accountability systems: effects.....	13
4.1.1 Effectiveness and high-stakes accountability systems	13
4.1.2 Equity and high-stakes accountability systems	15
4.1.3 Strategic behaviors, gaming practices and unintended effects	15
4.2 Normative implications and changes.....	16
5. Issues for further research.....	17
5.1 Goals and purpose of accountability policies	17
5.2 Means and tools of accountability policies: evidences, standards and benchmarks.....	17
5.3 Control and support: toward reflexive forms of accountability?	18
5.4 Cultural, institutional and socio-economic context: why does this matter?	19
Conclusion.....	20
Bibliography	22
APPENDIX 1: Accountability, central tools.....	28

1. Introduction

Accountability in education has become a centerpiece of educational reforms in the last three decades. A growing concern for effectiveness and quality in education systems has led a drive for reform and change. Many education systems 'have found themselves under pressure from a drive to increase the quality of education and improve outcomes for pupils in order to create a more skilled and educated workforce' (Osborn, 2006, p. 243). In this context, accountability has become both an 'instrument and a goal' (Bovens, 2006), supposed to improve performance of public schools and equity in education systems. Major education reforms such as the Education Reform Act in England (1988) or the well-known No Child Left Behind (2001) in the United-States, have introduced new policy tools which transform the way the State regulates the education system and held local organizations and actors accountable for their functioning and results.

In education, the concept of accountability increasingly refers to various dimensions and mechanisms intended to measure the outcomes of education systems and to hold individual or institutional actors accountable for reaching targets and goals fixed at the central level. Holding actors accountable for the use of public resources or the conformity of their actions to the rules and political demands is not new in public education institutions. However, the shift toward a 'new' accountability implies a move of central governance from inputs regulation to 'a model of steering by the results' (Linn, 2000).

This paper aims at providing and discussing some insights into the development of accountability policies (AP) in education. We first discuss the polysemic character of the concept of accountability, and emphasize the recurring features of accountability policies in education, as well as the social conditions that have fostered their emergence. Secondly, we present a typology of accountability policies and tools, reflecting the fact that accountability forms, tools and significance vary widely from one country to another. In a third part we synthesize the main findings of the core literature about the effects of AP on effectiveness and equity. Finally, we highlight a number of important issues in relation to AP as well as avenues for further research.

2. Accountability and accountability policies in education

2.1 A multi-layered concept

Accountability is an 'appealing but elusive concept' (Bovens, 2007) which takes many meanings and refers to various practices following the social field or the scientific discipline concerned (political science, financial accounting, public administration or education). Reviewing the literature, Lindberg (2013) has counted 'over 100 different "subtypes" and usages' of the term.

This Anglo-Saxon term is related historically and semantically to the term 'accounting' (Bovens, 2007) and is difficult to translate as such in other languages such as French (Broadfoot, 2000) or Slavic languages (Vesely, 2013). It has indeed a long tradition in political science, where it refers to the idea that:

'when a decision-making power is transferred from a principal (e.g. the citizens) to an agent (e. g. government), there must be a mechanism in place for holding the agent accountable for their decisions and tools for sanction' (Lindberg, 2013).

The political term is thus related to an obligation from the representatives to justify themselves to the constituencies who elect them (at least in democratic regimes). At the same time, higher administration officers (and their agents following the chain of command) should be held directly accountable by the government (or sometimes parliament). In these terms, accountability becomes a democratic tool for monitoring and controlling government and, furthermore, the administration's actions.

This meaning of accountability has broadened and also become blurrier, both in scientific literature and public discourse, as public administration reforms have been developing in various countries under the influence of New Public Management (NPM) (Hood, 1991) that situates accountability as a central device of good governance. In this context, the notion of accountability has become very popular, but also polysemic. It has either become a synonym for loosely defined political goals (such as 'good governance' or 'democracy') or has been related to various 'mechanisms for controlling and ensuring quality in public institutions' (Vesely, 2013, p. 5). However Bovens proposes a much more analytical definition of accountability:

'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 judgments, and the actor may face consequences' (Bovens, 2007, p. 450)

The advantage of this definition is that it can be used as an analytical tool structured around five operational questions to analyze and classify accountability features and tools:

1) Who is accountable? Who should render account? It could be organizations, such as political institutions (government and government administration), public sector organizations (schools), or individuals (politicians, higher officers, school principals, teachers...).

2) To whom (which forum) is the account to be rendered? According to this question, Bovens distinguishes various types of accountability: a) *political accountability* - when the forum comprises political actors or institutions, such as governments or constituencies; b) *managerial or administrative accountability* - when the forum is the administration or public service organization hierarchy; c) *professional accountability* - when the forum is a professional body, an audit office or 'chartered accountants' and accountability relates an actor to a professional peer ; d) *social (or market) accountability* - when the forum comprises actors from civil society, users of a service, clients, or interest groups (stakeholders) ; e) finally, *legal accountability* - when the forum comprises courts.

3) What is the relation (or the type of obligation) between the actor and the forum? A vertical relation refers 'to the situation where the forum formally wields power over the actor, perhaps due to the hierarchical relationship between actor and forum' (Bovens, 2007, p. 460). Horizontal accountability occurs when an account is given to stakeholders without formal obligation, as is the case in social accountability where relations are based on a social or moral obligation. Finally, a diagonal relation is an intermediary form (as is the case of accounts due to an audit or inspection body, without direct hierarchical power on the actor held to account).

4) & 5) About what does an account have to be rendered? On what basis is the judgment to be passed? It could be based on inputs, processes, outputs, or effects of the actions taken. These aspects can moreover be evaluated on the basis of different criteria (i.e. equity, effectiveness, efficiency, conformity, transparency, democracy) and different methods.

2.2 'New' accountability in education

The notion of accountability has been used in various ways in education as in other fields. If accountability often refers to a social practice - 'to be held to account' (Kogan, 1988; Leithwood, Edge & Jantzi, 1999; Broafoot, 2000; Jaafar & Anderson, 2007), various accountability approaches and tools are then defined or compared by raising the same questions as Bovens: Who is accountable? To whom is the account owed? What is being accounted for? What are the consequences of providing an account? For example, in an international 'state of the art', Leithwood (Leithwood, Edge & Jantzi, 1999; Leithwood & Earl, 2000) typifies a large panorama of various approaches and tools of accountability: *market based, decentralization of decision making, professional, and managerial* accountability. Accounts are to be given by teachers, principals or schools to various forums

(professional order, professional community, local community, district, State, parents) about either the processes or the outputs of education, with various consequences.

But following the recent policies conducted mostly in the United-States and England, accountability in education is increasingly understood in a narrower sense. Accountability is related to accounts to be given by schools and teachers about results and outputs to deliver, taking into account targets (in terms of qualification rate, performances at external exams) and standards (about curriculum or evaluation) fixed at the central level (the ministry, State level). These accounts are to be given to the chain of command (district or state) or to parents, on the basis of school and pupil performance assessment, related to various indicators of results³.

This 'new' accountability in education or 'performance-based accountability' systems (Carnoy & Loeb, 2002) imply a shift from input-based political regulation toward output-based regulation, where student results constitute the linchpin of accountability systems:

'In principle, focusing on student performance should move states away from input regulations - judging schools based on the number of books in the library and the proportion of certified staff, for example -toward a model of steering by results - using rewards, sanctions, and assistance to move schools toward higher levels of performance. In other words, the educational accountability should focus schools' attention less on compliance with rules and more on increasing learning for students' (Elmore & al. 1996, p. 65 in Linn, 2000, p. 12).

Performance-based accountability includes four elements (see Appendix 1): 1) standards (what the pupils should learn), 2) a testing system (or large-scale assessment system), usually by an external body, 3) public information about test results and an account explaining their sources or causes, 4) positive or negative consequences for schools (Harris & Herrington 2006). According to Figlio & Loeb (2011, p. 385) 'the most-developed accountability systems operate in the US, England, and Chile and they are also the systems on which the overwhelming majority of academic research has been based'. However, in other parts of the world, similar policies have been named with other labels as 'steuerung' policies (Altrichter, Heinrich & Soukup-Altrichter, 2011), 'steering policies' (de Landsheere, 1994), 'testing regimes' (Lingard, Martino & Rezai-Rashti, 2013) and 'evaluation policies' (Buisson-Fenet & Pons, 2014).

These performance-based accountability systems share common features in reference to the questions proposed by Bovens:

1. Accounts are to be given by individuals (teachers or principals, administrators) or organizations (schools, districts) to various organizations in the chain of command, such as districts, states, or to specialized agencies, such as inspection bodies.
2. The accountability relationship is vertical. Less frequently, the relationship can be horizontal, geared toward the local community.
3. Accountability is mostly based upon results or outputs of organizations (qualification and/or retention rate, pupils' performances in external assessments in key grades and subjects etc.). However, traditional objects of accountability (with respect to rules and procedures, use of budgets) do not disappear.

³ Harris and Herrington (2006), for example, distinguish two types of accountability characterizing some aspects of educational policies put in place in the United-States (US) in the last decades (1990-2005) – *government-based accountability* and *market-based accountability*. *Government-based accountability* corresponds to 'government efforts to measure the outcomes of students and schools, especially on the basis of student test scores, and to provide explicit rewards and punishments based on these measures' (Harris & Herrington, 2006, p. 217). Secondly, *market-based accountability* corresponds to policies providing parents with greater school choice. The basic assumption is that giving parents greater choice regarding the school attended by their children is the best way to develop competition between schools and to enhance the quality of public schooling. These types of accountability tend to be piled up over the traditional local public accountability, where school principals are held to account by district administrations, and elected school boards are accountable to their constituencies.

4. These outputs are evaluated according to some standards and measurable objectives (indicators, targets, benchmarks) and the actors held to account have to explain or justify potential gaps between their results and these standards to the forum.
5. Finally, the actors might have to face various consequences (symbolic or material) following this account. In particular, the literature has often opposed high-stake and low-stake accountability devices (Carnoy & Loeb, 2002).

Finally, accountability policies in education cannot be isolated from other public policies that have been implemented in many countries and education systems (such as decentralization, school autonomy and school choice) for two or three decades. These reforms of school governance and regulation should be related to major evolutions which have subjected governments to pressures or demands, in an at least partially converging sense, especially in the OECD industrialized countries (Ball, 1998):

- The development of economic globalization has accentuated business demands for a greater efficiency of public education systems, but also for greater attention given to the economy's needs for competencies. In this context, the discourse about the needs of the 'new knowledge economy' has led to emphasize the necessary improvement of the effectiveness and efficiency of the public education system (Brown, Lauder & Ashton, 2008).
- The Welfare State's crisis of legitimacy and funding, and the rise of neo-liberal political paradigms have led to call the bureaucratic modes of managing public action into question and to import managerial concerns heretofore characteristic of the private sector (preoccupations with efficiency and accountability) into the public sector. Indeed, the principles of the New Public Management (Hood, 1995) have had a major influence on public administration reforms and the development of decentralization and accountability policies in education. Nevertheless, it should be noted that these models have been implemented and recontextualized in various ways (see below).
- An increasing social demand on the part of middle classes has also emerged, favoring more quality, choice and the individualization of education pathways. Aside from the influence of the increasing individualization of social ties, this demand has its source in middle class anxiety in the face of the erosion of their social and professional positions (Ball, 2003a).
- There is also a phenomenon of globalization of education policies, at least in the form of the diffusion of reference models by various bodies, feeding the construction of new policy tools (in particular NPM and diverse post-bureaucratic models of governance [Maroy, 2012]). Such models sometimes serve to inspire, sometimes to legitimize the construction of national policies, notably through the circulation of ideas favored or initiated by transnational organizations (OECD, EU etc.) and policy networks (Ball, 2012; Rizvi & Lingard, 2000).

However, these external or internal evolutions and demands, pushing or pressing governments toward the development of performance-based accountability systems, do not lead to a single model of accountability policies. This is due to many factors, and we will only mention some of them, although they have been explored in more detail elsewhere (Maroy, Pons & Dupuy, 2017). First of all, the transnational models of accountability policies are recontextualized and translated in national policies in various ways. The international model could be subjected to idiosyncratic normative and cognitive 'bricolage' that adapt, hybridize the 'pure model' into a more legitimate model in the societal context (Maroy, 2009a; Steiner-Khamsi & Waldow, 2012). Moreover, there are political struggles around the model to put in place. Finally, the actual policy trajectory of accountability policies is often conditioned by the socio-economic context of the country and the path dependencies to societal or local institutions (Streeck & Thelen, 2005).

3. Accountability policies in education, diversity of tools and rationales

In our perspective, performance-based accountability policies or ‘steering by the results’ policies (Linn, 2000) intend to modify (to a lesser or greater extent) the set of coordination and control mechanisms put in place by the State and the educational authorities in order to orient or regulate the behaviors of local actors (for example teachers and school principals). In this sense, these policies bring forward new forms of political regulation⁴ of the education system that intend to improve the effectiveness or the equity of the system (Maroy, 2009b). As such, accountability policies are sharing four common traits (Maroy, 2013): 1) they are embedded in a new policy paradigm whereby the school is conceived no longer as a core institution within society, but as a performative system of production (Ball, 2003b). 2) Operational objectives of the school policy and system may be expressed in quantifiable data, which, in turn, become the standards and targets for the system. 3) Various testing tools of pupil achievements are central to assess the outputs of the system. 4) Individual or collective actors at different levels of the system are held accountable for these results, with various consequences for them. Moreover accountability policies stand on common policy tools (see Appendix 1) even if the features and range of tools used for accountability purpose can widely vary from an education system to another.

However, despite these similarities accountability policies take several forms depending on the context where they are implemented and rely on a variety of tools and rationales. This is related to diverse theory of change (Muller, 2000) and means by which the effectiveness and quality of the school system is supposed to be improved. In order to better discern the variations, significance and socio-political issues underlying these policies, we propose a theoretical typology of the diversity of rationales and policy tools used by accountability policies in various education systems⁵ (Maroy & Voisin, 2014; Maroy, 2015). We construct four types or approaches of accountability policies – *a regulation through strong accountability, a regulation through neo-bureaucratic accountability, a regulation through reflexive responsabilization and accountability, and a regulation through soft accountability* (see below). Consequently we highlight the fact that beyond some commonalities, these policies differ depending on their contexts, the instruments chosen by education systems to implement them, actor conceptions and the regulation theories underlying them.

3.1 Typology of accountability policies and tools: main dimensions

Our typology of accountability policies and tools is based on four dimensions. Two bear on the characteristics of policy tools deployed to implement policies (the degree to which measures are aligned and the implications of accountability for the actors), and two others on the theory of change embedded in policy tools (the conception of the actor targeted by the policy and the theory of change concerning his behavior).

Concerning the degree of alignment (strong or weak) between the tools and levels of action of education systems (central, intermediate, local), we can observe a tight and narrow coupling between tools and levels of action (when standards set criteria and provide guidelines that should orient local practices). In contrast, weak alignment involves instruments that are loosely coupled with one another and/or between levels of action. Another dimension to the tools, which is much discussed in the literature, is that of the nature and strength of the consequences faced by actors. In a high-stakes accountability system, repercussions for actors are high, considered in terms of incentives and sanctions as well as external reputation. In

⁴ Regulation is here to understand in a broad meaning as in the french term *régulation* which is broader than the formal regulation in english, translated by *règlementation* in french. We take the view that ‘social regulation’ denotes multiple, contradictory and sometimes conflicting processes for orienting the behaviours of actors and defining the rules of the game in a social system (Maroy, 2008 and 2012). ‘Political regulation’ by public authorities is not only institutionalized in legal mechanisms (in this case, political regulation essentially means formal or statutory regulation) but also, more recently, in incentive, evaluation, emulation, consultation and accountability mechanisms.

⁵ Our procedure aims to construct a typology of ‘ideal types’ in a Weberian perspective. It’s not a descriptive typology that brings together similar cases together, even if we will describe some cases that can illustrate our types. In this sense our typology is an intellectual construction that aims to provide a tool in order to better understand the complexity of accountability in education taking into account the tools they use and the rationales at the heart of these policies.

'soft' or 'reflexive' accountability systems (Dupriez & Mons, 2011), consequences for actors are lower and their nature could be quite different. Tools and mechanisms that help actors to develop reflexivity about their practices and to foster changes in practices, beliefs or professional identities are prioritized.

The two other dimensions of the typology bear on the theory of change and regulation underpinning policies. This theory is not necessarily made explicit in a developed discourse, although it may sometimes be the case⁶. Change theories are often embedded in the policy tools (PT) that are operationalizing the policy (Lascoumes & Le Galès, 2004). Change theory involves, on one hand, the conception of the actor that the policy intends to regulate and, on the other hand, the external or internal character of measures or dispositions by which an educational authority seeks to change or regulate the behavior of a local actor.

Thus the third dimension of our typology is that AP can consider the actor targeted by a policy (for example teachers) either as 'utilitarian', or 'reflexive'. In the first case, actors are considered as primarily moved by a rational logic and a calculating rationality. They are sensitive to constraints. In the second case, actors are conceived as 'reflexive' as well as culturally and socially situated. Actions respond not only to constraint but also to social obligation, they stand on culturally constructed and shared cognitive patterns. Improvement and change are conceived as the results of a reflection on one's own practices (Leithwood, Jantzi & Steinbach, 1995). Regulation therefore consists of institutional arrangements that might support this reflexivity and this learning process at an individual and/or collective level.

The fourth analytical dimension concerns the nature of the mediation on which the change is based. More specifically, mediations operate through the more or less intensive recourse to external devices, or they operate in acting on the internal dispositions of actors (Boltanski & Chiapello, 1999). In certain types of AP, policy ontology makes external measures key factors in the process of change and regulation – they are the pragmatic supports that tend to condition the orientation of individual or collective conduct from the outside. In contrast, other types of AP give greater importance to the interiority of actors, their ethos and internalized dispositions, as key vectors and mediations in the process of improving school performance (Mangez, 2001).

3.2 *Four approaches to regulation and accountability*

We construct four rationales (see Table 1) underlying accountability policies: **a regulation through strong accountability, a regulation through neo-bureaucratic accountability, a regulation through reflexive responsabilization and accountability, and a regulation through soft accountability**. We illustrate these 'ideal types' with the presentation of education systems that put in place national AP that are close to our models even if they cannot be reduced to the type they exemplify.

⁶ Besides the influence of NPM, networks of experts were able to play a key role in the formulation of these theories. Concerning 'soft accountability', Claude Thélot has for example theorized the 'mirror effect' in France (Mons, 2009). The role of the inspectorate in the conception of 'self-evaluation' has been important in the Scottish case (Ozga & Grek, 2012). In Canada and the US, economic theory has been very influential in the conception of 'high-stakes accountability', while the reflexive model has been influenced in Ontario by authors like Fullan, Rincon-Gallardo & Hargreaves, A. (2015)

Table 1: Four approaches to regulation by results

	Regulation through 'strong' accountability	Regulation through neo-bureaucratic accountability	Regulation through reflexive responsabilization and accountability	Regulation through 'soft' accountability
Stakes for actors	High	Moderate to low	Moderate to high	Low
Alignment of tools and levels of action	Strong	Strong	Strong	Weak
Conception of actor	Utilitarian	Utilitarian	Reflexive and socially situated	Reflexive and socially situated
Central mediation for the expected change	External devices (information, evaluation, control, support in case of a problem)	External devices (information, evaluation, control and support)	External devices (information, evaluation, control and support) and actors' dispositions	Actors' dispositions, evaluation and support measures
Examples	Texas, England	Québec	Ontario, Scotland	Belgium, France

3.2.1 Regulation through strong accountability

Regulation through strong accountability entails important consequences for actors, a tight coupling between policy tools and levels of action, as well as a conception of actors as primarily utilitarian, motivated by their own interest and sensitive to constraints. This 'strategic' conception can be summarized by Woessmann's quote:

'That is, if the actors in the education process are rewarded [extrinsically or intrinsically] for producing better student performance, if they are penalized for not producing high performance, they will improve performance' (2007, p. 473).

With *regulation through strong accountability*, change is supposed to occur through external measures. This accountability approach involves a highly developed information system about actors' performances and several features such as targets (performance standards for students and schools) and high-stakes testing (large-scale state-wide standardized testing) that lead to consequences for several actors (sanctions and rewards). The focus of this approach rests on making actors accountable for student results, as well as on setting a large range of repercussions (both material and symbolic) for key actors, such as teachers, head-teachers and schools at the central level.

Prime examples of this logic are Texas and England⁷ (Broadfoot, 2000; Ozga, 2009). As noted by Mc Neil, Coppola, Radigan & Vasquez Heilig (2008, p.3), the accountability system in Texas is 'an extreme form of centralized management with a strict hierarchy in which rules and sanctions are set at the top, with every level of the system accountable to the level above it for measurable performance'. This high-stakes accountability system combines statewide-standardized tests (for grades three, eight and ten), publication of aggregate student, school and district results, ranking of schools and districts, as well as major repercussions for actors (districts, school

⁷ England and the United-States are also characterised by well-developed mechanisms of school choice. In these contexts characterised by choice/markets, standards, accountability, privatization, and diversity (Hursh, 2005, p. 6), school choice mechanisms are overlapping with government-based accountability mechanisms (Harris & Herrington, 2006) and constitute other forms of regulation based on competition, market and choice favoured and encouraged at the federal and/or state levels.

administrators, schools and teachers) (Carnoy & Loeb, 2002). The alignment between levels of the education system with a focus on standardized objectives and actor accountability is organized via a tight hierarchical control and management, including the monitoring of test results and school performances. The Texas Assessment of Academic Skills (TAAS), aligned with curriculum and performance standards established by the Texas Education Agency, is the linchpin of an accountability system with serious ramifications for teachers, schools and districts, including financial incentives and career advancement measures (salary bonuses, takeover of staff management etc.). Despite the fact that the Texan system plans for both support measures and pressure on the actors, focus seems to be on the presence of external measures exercising pressure on actors.

3.2.2 *Regulation through neo-bureaucratic accountability*

Regulation through neo-bureaucratic accountability with a focus on test results also rests upon a utilitarian conception of the actor and a regulation based on a close alignment between tools and different levels of action. External procedures tighten higher administrative authorities' control over the results with a top-down approach. The emphasis is on formal accountability to higher authorities, using bureaucratic procedures and external tools with a stress on assessment and data collection tools but with low consequences attached. Regulation operates through actors' sensitivity to the constraints and rules that condition their action. Regulation is imposed through measures from the outside, without actors' moral assumption of responsibility being targeted by the policy. There is no calling for internal change of personal beliefs or values. Professional ethos is neither a basis nor a key driver for the change expected. In this sense, the regulatory action remains bureaucratic (Le Galès & Scott, 2008). The province of Quebec (Canada) can be categorized in this type of approach.

The 'Results-based management' education policy⁸ was introduced in Quebec by two major Laws: Law 124 in 2002 and Law 88 in 2008. This accountability system is based on a number of central features: strategic alignment, planning, contractual agreements and vertical accountability. Targets (including pupil achievements at national tests, dropout and graduation rates) are set at the central level (Ministry of education), and schools have to provide an account to their school boards about their performance and their contribution to pre-determined targets. School boards will, in turn, render an account to the Ministry of education. The main instruments of Quebec's accountability policy consist of external and vertical measures aimed at aligning the lower echelons with the targets determined by higher levels of decision. The reliance on external measure is a shared feature with the Texas approach. But these measures are different in nature, since in Quebec there is no emphasis on financial incentives or constraints. Instead, the approach is based on the deployment of tools based on bureaucratic accountability practices (for example in Quebec the account on results has to be given by principals to hierarchical authorities such as districts' administrator – see Appendix 1). Therefore, regulation theory rests upon a conception of utilitarian actors sensitive to the external rules and bureaucratic controls designed to encourage them to readjust their actions.

3.2.3 *Regulation approach through reflexive responsabilization and accountability*

Regulation through reflexive responsabilization and accountability entails high to moderate consequences for individuals and organizations, as well as instruments and levels of action that remain closely aligned. Policy regulation does not occur only through external measures but also by mobilizing actors' inner dispositions. This accountability approach is based on various external devices, sophisticated data collection and assessment tools focusing on both results and processes, as well as on some support and improvement measures. Cognitive and normative dispositions internal to the actors are heavily mobilized in self-evaluation processes which no longer focus solely on results, but also examine change and learning processes as well as the improvement of school and teacher practices. Examples of this approach include accountability measures established in Scotland (Ozga & Grek, 2012) and Ontario (Chang, Fisher & Rubenson, 2007; Jaafar & Anderson, 2007).

⁸ See Maroy & al. (2014) on the genesis and implementation of the 'Results-based management' (i.e Gestion Axée sur les Résultats) education policy

In Scotland, the School Self Evaluation model was introduced and promoted by the Scottish inspectorate in the 1990's. It was formalized through the publication of the report *How Good is Our School? Self-Evaluation Using Performance Indicators* by the Scottish inspection services (Scottish Office Education and Industry Department, SOEID, 1996 in Croxford, Grek & Jeelani Shaik, 2009). The managerial Scottish performance-based accountability logic involves a strict monitoring of school performance. It stands on a range of key instruments at the central level such as curriculum and performance standards, standardized testing, information tools (publication and monitoring of school results), and on the obligation for schools to use their results as the basis for a self-evaluation process that should lead to better performance. The Scottish education system includes a strict system of incentives and sanctions. However, support measures are central: they consist of interventions by external actors who assist schools in the process of self-evaluation in order to improve their processes (management, educational practices) and results. This process is intended to encourage the learning processes of professionals and administrators, leading to other ways of thinking. It relies on the internalization of the criteria on which these actors are evaluated (according to their performance). Therefore, the focus is, on the one hand, on actors assuming responsibility for the results and, on the other hand, on their capacity for reflexivity with respect to their professional practices. One might hypothesize that the action theory underlying the implementation of these instruments supposes a more reflexive actor, although sensitive to external incentives.

3.2.4 Regulation through soft accountability

In *regulation through soft accountability*, change theory primarily rests upon a conception of local actors as reflexive and supposedly socially inclined (by their ethos, training, and professional socialization) to improve their practices in the direction expected by educational authorities. Accountability mechanisms are poorly formalized and characterized by a loose coupling between both tools and standards, and between levels of action. Regulation operates through external measures – essentially assessment tools with low stakes associated, it relies mostly on an impact on the actors' perceptions through the 'mirror effect' (Thélot, 1998 in Mons, 2009) allegedly produced by test results. It counts on the dispositions of local school actors to engage in a reflection on their practices, stemming from better knowledge provided by the evaluation of their students' results. Furthermore, this reliance on local actors' internalized dispositions can be supported by the implementation of actions combining support and control to various degrees. This approach is favored in French-speaking Belgium and in France (Mons & Pons, 2006; Mons & Dupriez, 2010; Dutercq & Cuculou, 2013).

French-speaking Belgium has developed external evaluation and accountability policy for schools (Maroy & Mangez, 2011) for fifteen years. This encompasses the development of external end-of-year examinations and diagnostic evaluations, 'pedagogical standards' in terms of skills to be attained at different levels. Despite the development of these centralized tools (curriculum standards, evaluation frames of reference, and systematic and regular external evaluations), the autonomy of local educational authorities and schools remains significant in terms of management and pedagogy. As for external evaluations, they are subject to a two-fold monitoring by the system. On the one hand, the results of external evaluations may be used by the school's management as a tool for 'micro-steering'; on the other hand, the inspection services, whose role is to inspect schools, have the opportunity to assess pedagogical practices in relation to both the legal framework and the expected and observed 'performance' level, with regard to competencies defined as pedagogical objectives. In the case of a perceived deficiency, the inspector must notify the educational authorities concerned, who are then responsible for ensuring that improvements are made, without, nonetheless, planning sanctions or consequences (Maroy, 2009c). Finally, the alignment of policy tools remains relatively weak and the actual 'steering' from the top is mediated by the institutional autonomy of educational organizations and the actors' sense making of the policy. The external measures used to orient their behaviors are here weak.

In summary, different types and models of accountability can be identified in different contexts. They stand on various tools, features and theories of change. Whilst they share common purposes, mainly 'to shorten the feedback loops between policy makers, principals, and teacher (...) to set clear standards, measure students

performance, and use those measures to evaluate the effectiveness of school' (Cohen, 1996; O'Day & Smith, 1993 in Ladd & Lauen 2010, p. 426), they are prioritizing differentiated mechanisms in order to bring about change in education systems and practices.

4. Accountability in education: effects, side effects and normative implications –main research results

Identifying more effective approaches and tools in order to improve school results and student achievement is at the heart of politicians and social scientists concerns. If the rhetoric surrounding the development of accountability policies puts forward their presupposed positive impact on effectiveness and equity in education systems, finding strong evidence remains a challenge for many reasons: 1) empirical studies are conducted along various theoretical and epistemological perspectives; 2) most comparative evaluation studies focus on the effects of high-stakes accountability systems, mainly in the United-States (US). Moreover, because isolating the effects of system-wide reforms represents a methodological issue, studies often focus on key related tools such as high-stakes testing and incentives for actors; 3) fewer studies have looked at *soft accountability* or *reflexive accountability* systems.

In this section we will first discuss some of the main findings of evaluation studies in a *regulation through strong accountability* context, otherwise named high-stakes accountability systems, to provide some insights into the central questions: how far are high-stakes accountability policies reaching their objectives, in other words how effective are they for better performance? To what extent can they improve equity in the school system? In order to have a more comprehensive picture, we will also discuss some unintended effects of high-stakes accountability policies highlighted by this literature. As we will see, it is difficult to give a clear-cut answer to these questions. In this regard Lee's recommendation seems quite accurate:

'(...) educational policy makers and practitioners should be cautioned against relying exclusively on research that is consistent with their ideological positions to support or criticize the current high-stakes testing policy movement. They should become aware of potential biases arising from the uncertainty and variability of evidence in the literature' (2008, p. 629)

Secondly, we will discuss normative changes brought about by accountability policies, as highlighted by critical sociological studies.

4.1 *Regulation through strong accountability and high-stakes accountability systems: effects*

We present here the conclusions of extensive literature reviews, longitudinal or macro-statistics studies based on the Federal National Assessment of Educational Progress (NAEP) in the United-States, State driven tests, district driven tests (Chicago, New York), and meta-analyses that try to isolate and discuss the impact of high-stakes accountability systems on student performance and achievement gaps.

4.1.1 *Effectiveness and high-stakes accountability systems*

Slight positive effects on student performance... but variable and unstable over time

Main findings of US econometrical studies suggest that high-stakes accountability systems, particularly 'more stringent accountability systems' (Harris & Herrington, 2006) have a positive effect on student achievement

(Hanushek & Raymond, 2005; Jacob, 2005; Harris & Herrington, 2006; Lee, 2008; Figlio & Loeb, 2011). But this effect remains moderate (Lee, 2008)⁹.

However, there are important variations regarding school grades and disciplines (Carnoy & Loeb, 2002; Jacob, 2005; Lee, 2008; Chiang, 2009; Mons, 2009; Figlio & Loeb, 2011). Achievement gains are bigger for higher school levels, possibly because these grades are facing the largest incentives in order to improve student performance (Jacob, 2005, p. 772). They also tend to be larger in math than in reading or other disciplines tested (Jacob, 2005; Lee, 2008; Treisman & Fuller, 2001 in Mons, 2009; Dee & Jacob, 2009 in Figlio & Loeb, 2011).

Furthermore, achievement gains do not remain stable over time (Lee, 2008; Chiang, 2009; Mons, 2009). In this respect, as pointed out by Lee (2008, p. 619):

‘the volatility of gain scores requires that one look at changes in performance over the long run’

Differentiated results regarding specific mechanisms and tools

Impact on student achievement also differs regarding mechanisms and tools at the heart of accountability systems. Three main tools and mechanisms are usually studied: 1) the use of information for accountability purposes at the school level (Bruns, Filmer & Patrinos, 2011) – report cards, publication of test results and ranking of schools; 2) the introduction of high-stakes testing; 3) the system of incentives, sanctions and rewards targeted at different levels (teachers, schools, districts levels).

Considering the effect of report cards (publication of information regarding school results, with possibly pupil characteristics and breakdown regarding sub-groups), the assumption is that publishing student and school results allows the public to identify low-performing schools and districts, and possibly leads to better performance. The ‘scarlet letter’ effect ‘would suggest that educators wish to avoid, and will respond to stigmatization regardless of other incentives’ (Harris & Herrington, 2006, p. 220). But analysis of the core literature leads to mixed conclusions: as an example Bishop & al. (2006 in Harris & Herrington, 2006), Hanushek (2003 in Harris & Herrington, 2006), Carnoy & Loeb (2002) found that the gains in NAEP are greater for the states that use report cards, but Harris & Herrington (2004 in Harris & Herrington, 2006) found no effect. Finally, it is difficult to conclude on the positive effect of report cards as such on achievement gains.

Conversely, high-stakes testing, particularly Promotion and Graduation Exams (PGE) and State High School Exit Examination (HSEE), seems to have a positive impact on student achievement (Carnoy & Loeb, 2002; Harris & Herrington, 2006; Neill, 1998 in Grodsky, Warren & Kalogrides, 2009). For Harris & Herrington (2006) indeed, PGE is a ‘key player’ in the increase of student achievement.

Research consistently shows that incentives, sanctions and rewards for schools and individuals attached to high-stakes testing may have the potential to increase student achievement¹⁰. It is the conclusion of system-wide studies carried out by Chiang (2009) and Reback (2008) (as well as Hanushek & Raymond’ [2005] cross-States analysis and Harris & Herrington’ [2006] literature review).

Nevertheless, there is strong evidence that the introduction of sanctions linked to high-stakes testing has also led to the development of strategic behavior at the school or individual levels (see below). This has led many authors to raise some concerns about this mechanism.

⁹ Even if the evidence looks sometimes contradictory, or in Lee’s terms ‘mixed and inconclusive’ (2008, p. 628): see for example Amrein et Berliner (2003 in Grodsky & al., 2009, p. 5) who underline the lack of evidence regarding the link between high-stakes testing and achievement gains, or the discrepancies between study results highlighted by Lee (2008).

¹⁰ Bruns, Filmer & Patrinos (2011), in their report for the World Bank, put forward a possible positive impact of school-based-management and teacher incentives on learning outcomes.

4.1.2 *Equity and high-stakes accountability systems*

One of the key objectives of the introduction of high-stakes accountability systems in the US is also to reduce the achievement gap between students in a context of strong achievement inequalities (Carnoy & Loeb, 2002; Harris & Herrington, 2006), where underperforming groups are traditionally subgroups of students 'defined by their race, income and disability status' (Figlio & Loeb, 2011, p. 395).

A plausible negative impact with variations regarding social and ethnic sub-groups

Leaving aside some discrepancies between study conclusions, a consensus emerges regarding the lack of effect of high-stakes accountability policies on narrowing the achievement gap (Hanushek & Raymond, 2005; Harris & Herrington, 2006; Lee, 2008; Figlio & Loeb, 2011). On the contrary, it even tends to widen the performance gap between schools and students, mostly regarding the more disadvantaged ones. Hanushek & Raymond's (2005) conclusions show that the 'black-white' achievement gap has widened after the introduction of high-stakes accountability systems, while Dee's (2002 in Carnoy & Loeb, 2002) conclusions highlight 'reductions in educational attainment, particularly for black students'.

In a context of market and competition between schools, high-stakes accountability mechanisms also tend to increase social and ethnic segregation between schools (Hanushek & Raymond, 2005), and, as a consequence, the stigmatization of low-performing schools and students. Studies have shown that such mechanisms lower the capacity of low-performing schools to attract and retain high-qualified teachers (Wolf & Janssens, 2007; Finnigan & Gross, 2007; Figlio & Loeb, 2011) and consistently increase school turn over (Figlio & Loeb, 2011). Such evidence matters if we assume that teacher quality is one of the key factors for the quality of schooling and that school segregation is negatively correlated with equity in education systems (Demeuse & Baye, 2009; Dumay, Dupriez & Maroy, 2010). Consequently, it seems particularly important to have a closer look at the differentiated impact of high-stakes accountability on disadvantaged schools and students.

4.1.3 *Strategic behaviors, gaming practices and unintended effects*

Many authors have also highlighted the fact that high-stakes accountability systems, particularly sanctions and rewards, lead to the development of strategic behaviors and provide incentives for actors (at the state, district, school or individual level) 'to game the system' (Heilig & Darling-Hammond, 2008; Lee, 2010). Under accountability pressure, actors tend to put in place gaming strategies that aim at artificially rising test outcomes. These gaming strategies could range from the falsification of student and school results, the exclusion of low-performing students from testing and from school, to the increase of student retention in lower grades, as well as the placement of weaker students in special need education or 'limited English proficiency' that implies students' ineligibility to high-stakes testing (Linn, 2000; Webb, 2005; Mons, 2009; Ladd & Lauen, 2010; Figlio & Loeb, 2011).

In a context of high-stakes testing, 'given the consequences attached to test performance in certain subjects, one might expect teachers and students to shift resources and attention toward subjects included in the accountability program' (Jacob, 2005, p. 786). The very well documented phenomenon of 'teaching to the test' consists in focusing learning activities and taught contents on test preparation. It leads to a focus on short-term learning outcomes and a reduction of the taught curriculum, particularly for low-performing students. The shift could also translate into increased attention paid to students who are more likely to improve their performance – the ones just below proficiency level (Ladd & Lauen, 2010). Moreover, the threat of sanctions also affects teacher motivation and increases teacher and student stress (Finnigan & Gross, 2007; Mons, 2009). These collateral effects of accountability policies could lead to a downward cycle. Further investigations are thus needed in order to bring some insights regarding the relationship between accountability policies, particularly high-stakes accountability policies, and equity.

4.2 Normative implications and changes

Critical sociological studies, among others, have highlighted normative implications and changes that new accountability policies and mechanisms bring about in education systems. These critical studies introduce no distinction between various accountability policies types or rationales. However, we can hypothesize that most of their conclusions are plausible for high-stakes accountability policies.

At the school level, the pressure put by high-stakes testing and the emphasis on test scores involve a narrowing of the attention of pupils and teachers to the subjects and disciplines assessed by external tests, most particularly for most disadvantaged pupils. With it comes the risk, highlighted by Darling-Hammond (2004, p. 18 in Hong & Young, 2008, p. 6) that overemphasis on test scores will lead 'to test-based instruction that ignores critical real world skills, especially for lower-income and lower performing students; and to less useful and engaging education'. It goes hand in hand with the reinforcement of a strategic attitude toward schooling. This focus on short-term, cognitive skills, also questions the role of schooling and education in the development of critical ways of thinking. In the classroom, Hargreaves (2003) underlines the loss of creativity and open collective work when there is less room for individual or local curriculum choices and less time for cooperation, due to a unilateral focus on achievement and tests to pass. The pressure for results and performance could also place teachers in low-performing schools in front of a dilemma:

'Teaching is a profession that is typically driven by ethical motive or intrinsic desire (...) Most teachers, therefore, expect to teach in congruence with their moral purpose, i.e. so that students would understand and learn to promote their personal development and growth, not only for favourable exam scores or other externally set conditions of progress' (Sahlberg, 2010, p. 50).

As highlighted by Pollard and Osborn (2001), the research of effectiveness *per se* can also challenge the affective and humanistic dimension of teaching and learning. Performance standards can also be viewed as reducing the diversity, complexity and the constitutive uncertainty of the teaching and learning relationship (Ball, 2003b; Maroy, 2006). For Ball (2003b, p. 217), the research of effectiveness leads to a performativity regime that translates 'complex social processes and events into simple figures or categories of judgments'. It deeply modifies teachers' identities and conceptions of professionalism, with an emphasis on competition and performance rather than on professional judgment and cooperation: 'a new basis for ethical decision-making and moral judgment is erected by the "incentives" of performance' (Ball, 2003b, p. 218).

Furthermore, Mons (2008) shows that in many European educational contexts, we see the development of control on teachers and schools in various forms, from the traditional inspection of individual teaching processes, self-evaluation, interviews with the principal, to schools being audited. This development of teacher's work control is often linked to the much-discussed risk of teacher deprofessionalization. Professionalization is not just a matter of expertise over complex and uncertain tasks, it is also the ability of a group to set/negotiate their employment conditions, the content of their work, partially insulated from the pressure of customers/users, or managers (Freidson, 2001). There is an on-going loss of power and de-professionalization of teachers, to the benefit of either parents or school managers (Gewirtz, Mahony, Hextall & Cribb, 2008). There is also a risk of redefinition of professionalism. Professionalism in the 'full sense' involves autonomy in decision-making and reflection, as well as ethical autonomy, which may lead to questioning the orientation of policies. At the opposite, a managerial professionalism is currently developing: expertise and autonomy are expected on processes, while goals and normative frames are supposed to be a political choice, which have been made by policy makers in a democratic way (Ball, 2003b; Gerwirtz, 2002; Maroy, 2006).

To summarize this section, it is difficult to give a clear-cut answer about the effects of high-stakes accountability policies on effectiveness and equity in education systems. Firstly, discrepancies in study conclusions, methodological challenges, as well as the variety of contexts and policy tools could explain the difficulty of giving

an assertive conclusion. Secondly, if we can cautiously acknowledge the moderate but unstable positive impact of these policies on student achievement, their potentially negative impact on equity and other unintended effects need further investigation. In addition, the processes that mediate the relationship between new forms of accountability, effectiveness and equity remain, from our point of view, central to understanding the implementation of education policies, and require further research. Finally, normative changes matter and deserve to be questioned.

5. Issues for further research

In this section, we will emphasize some issues that need to be taken into account when thinking about the general relevance and choices to be made concerning the institutional design of accountability. Meanwhile, we will point out some questions that deserve further research. We will develop these points successively in relation with the goals of accountability policies, their means, and finally the general context (cultural, economic or institutional) within which they are developed.

5.1 *Goals and purpose of accountability policies*

The goals of accountability policies are often oriented toward the improvement of the effectiveness and equity of the schooling system, on the basis of some major indicators, such as the qualification rate, the retention rate, or the pupil's average performance at external exams in certain key topics and stages of the curriculum. Promoting success for all, reducing the qualification or performance gaps between different sub-groups of pupils (defined in socio-economic or ethnic categories) are highly consensual goals that could hardly be contested by school professionals, stakeholders or users.

However, a fundamental critique is related to the reduction of the education goals enacted by accountability policies, but also to short-term acquisitions and the 'shallow' cognitive skills that high-stakes testing could involve (as has been discussed earlier). Due to the fact that monitoring of results and progress is focused on key subjects (and improvement in related practices), accountability policies tools tend to reinforce a curricular reduction and a symbolic hierarchy between the different subjects to learn, for pupils as well as for teachers and principals. This means that the general goals defined by democratic societies and governments (i.e. defined by some authors as qualification, socialisation, subjectification (Dubet & Martuccelli, 1996; Biesta, 2009) tend to be reduced in practice to certain learning outcomes in specific fields (such as math, mother tongue and sciences) at the expense of complex cognitive learning or other, broader 'non-cognitive' learning, related to socialization or education in a broad meaning – such as social skills and moral values. This trend or drift raises the question of the selective character of the indicators of outcomes (in relation to the missions and goals of the education system) and the way monitoring and support tools are selectively oriented towards certain selective results.

Therefore, in terms of further research, we should encourage the development of research on accountability policies' effects on 'non-cognitive' learning and skills: is such learning really impeded or weakened? Moreover, considering the development of 'non-cognitive' learning indicators, the question of their relevance, validity and fidelity should also be put forward and questioned. Some research has started on these topics (see for example Van Landeghem & al., 2002), but should be further developed, given the present day's context, where in the general role of schooling in terms of social cohesion tends to be more and more crucial. In other words, accountability policies have certainly some economic rationale (for the qualification goal of schooling), but its societal or social relevance for the socialization goals of education have to be questioned.

5.2 *Means and tools of accountability policies: evidences, standards and benchmarks*

Concerning the means of accountability policies, one key open question is the relative effectiveness of the various policy tools (or the theory of change embedded in those tools) used to implement it. We have seen

(chapter 3) important limits or drifts when it comes to its actual implementation. The curricular narrowing strategic adaptation of professionals (i.e. ‘teaching to the test’, cheating practices, filtering and tracking students and all other gaming practices developed to reach targets) are drifts that cast some doubt about the ‘progress’, if any, made in terms of effectiveness and equity. Moreover, it shows that accountability policies – in particular high-stakes accountability policies - are conditioned by the way in which they are implemented, and by the strategic ways in which actors use the tools or rules linked to these policies. In other words, accountability policies are transformed and translated by local or meso level actors who make sense of these policies, with their own cognitive or normative frames (Spillane, Reiser & Gomez, 2006).

These cognitive and strategic processes of implementation and use of accountability policies should also be an important research avenue to be developed. These questions matter particularly for low-performing schools and students that are also often the most disadvantaged ones. Indeed, they seem to struggle and suffer the most from the unintended, adverse effects high-stakes accountability tends to create.

We have already discussed the question of results indicators in relation with the spectrum of skills and knowledge they embrace. About these indicators, another key research (and political) question should also be to define the adequate level of proficiency expected in the various domains evaluated. These questions are of course to be discussed democratically in each society. Although we cannot develop this point here, it is useful to briefly recall that the development of international testing initiatives (such as the PISA project led by OECD or the TIMSS initiative by LEA) constitute international benchmarks (or regional ones) that have varied but serious impacts on national education policy and debate, through the dynamics of comparison and competition that they tend to develop (Carvalho, 2009; Grek, 2009).

This leads us to raise the following question: what should be the relevant level of proficiency considered as benchmarks in various subjects and various national and regional contexts? This question should also be an avenue for future research. In particular, what is the good balance of various rationales and education goals to be taken into account to define these levels? To what extent do international standards and benchmarks have a positive influence? Do they change the balance between the various and complex education goals and rationales taken into account in a given national context (i.e. prioritizing economic goals over societal ones) (Lingard, Martino & Rezai-Rashti, 2013)?

Finally, the question of the knowledge base and technical tools and indicators to measure schools’ and teachers’ net effect should be looked into. Accountability policies, and new forms of accountability in education imply that teachers and school actors bear the responsibility (at least partly) for students’ results and improvement. This raises the question of who has to be held accountable for students’ learning and, in the case of sanctions, who has to take the blame (individually or collectively). This judgment should be fair and technically requires that the indicators and measurements of results be very accurate. The data available should also be able to actually disentangle the effect of the context and the effect of the teacher, the class or the school. This classical problem of effectiveness research could be theoretically and methodologically solved in controlled research environments (Harker & Nash, 1996; De Fraine, Van Damme & Onghena, 2002). However, it is not guaranteed that the same methodological caution can and is developed in all local schooling contexts by inspection bodies or local authorities in charge of holding schools accountable. Indeed, the quality of the knowledge base and the capacity of local actors to use it in a fair manner are important conditions for the legitimacy of the accountability process. Finally, the knowledge basis for accountability and the technical capacity to control context effects on student’s achievement matters, as seen above.

5.3 Control and support: toward reflexive forms of accountability?

Another key question related to the means of accountability policies is the ‘good’ balance between support and control in the monitoring process and the search for improvement within the schooling system. In all

accountability policies and systems quantitative targets and the evaluation of results are central tools to diagnose the relative effectiveness of educational organizations and professionals. Nevertheless, there are at least four distinct types of accountability tools and rationales that organize the follow up and the monitoring of actions in very different ways (see the typology, above) and that give very different weight to support and control.

Furthermore, if most research on accountability policies has been carried out in the context of Anglo-Saxon countries and systems, mostly oriented by *regulation through strong accountability*, further research should focus on the other rationales, in particular the 'reflexive' and 'soft' accountability policies and tools. Some researchers consider reflexive accountability as more effective than *regulation through strong accountability* (Normand & Derouet, 2011; Fullan, Ricon-Gallardo & Hargreaves, 2015). This should be validated by further research, looking not only at classical indicators of results (effectiveness and equity) but also collateral and unexpected results. The process of improvement (in particular the balance between capacity building and control, the use and limits of data-set analysis) should also be the object of more attention.

The variety of means and approaches of accountability policies should also be related to the question of teacher professionalism and autonomy. The evolution and the variation of teacher professionalism in relation with the variety of the models of accountability could also be analyzed.

Finally, considering the side effects of high-stakes accountability and the resistance it arouse at the local level (but also State level in the case of the US [Hursh, 2005]), one can argue that more intelligent forms of accountability should stress the principle of mutual responsibility that can be seen as a two-way process (Sahlberg, 2010). On the one hand, schools should be held accountable for the outcomes of schooling, which not only encompass cognitive learning but also larger outcomes. On the other hand, authorities (including policy makers) should provide resources and support needed in order to achieve these 'jointly agreed goals' (Op. cit., p. 54).

5.4 *Cultural, institutional and socio-economic context: why does this matter?*

Last but not least is the issue of the cultural, institutional or socio-economic contexts. We have seen that accountability policies have initially been developed in Anglo-Saxon countries (USA, UK, Canada, Australia and New-Zealand) but also in Chili. In European, Asian and Latin-American contexts, other policies are underway, under various labels, centered on the external testing of pupils, school (self) evaluation and accountability. Due to the general context of globalization of education policies, this development of accountability policies in education will also probably concern the African continent and, more generally, developing countries. In this context, it is worth putting forward the questions of the cultural, institutional or socio-economic conditions that could influence or impede on the effectiveness and relevance of educational policies. Taking advantage of the (important and various) research that has emphasized the key problems with the 'borrowing-lending' of policies from the North to the South (Steiner-Khamsi & Waldow, 2012), especially concerning various regulation and governance features (such as decentralization, privatization, or partnership) (De Grauwe & Lugaz, 2007), further research or syntheses of research should be done on the various conditions that could impact the relevance or the feasibility of the diffusion of accountability policies.

Further research should be focused on the aspects of the schooling system (and other education policies) that seem to be pre-conditions for accountability policies. Indeed, accountability systems seem to be conditioned by the existence of an infrastructure of relevant and valid data (on educational resources, processes, outcomes), and the distribution of knowledge and skills necessary among professionals in order to use and interpret this data. Secondly, a prerequisite should also be the development of education and training systems for teachers and principals, or the existence of a staff of pedagogic counselors able to play their role in the monitoring of schools. More crucially, the actual use and effectiveness of these human resources and knowledge-based infrastructures are themselves related and conditioned by other economic and material resources that are basic

conditions for access or success in education. An obvious example is the importance of schooling infrastructure and pedagogic material, or the level of funding resources available, that are the conditions for ‘the capacity building’, the professional development of teachers and the educational workforce (Mingat & Suchaut, 2000). These conditions matter and condition the equity of the system. As the low-performing schools fail to answer to accountability pressure and have a lack of capacity (Harris & Herrington, 2006; Figlio & Loeb, 2011), an important emphasis has to be put on capacity building, as an important pre-condition for accountability systems:

‘If school resources must be at a certain level to bring about positive performance improvements, or if principals and teachers have sufficient resources but lack the specific policy and practice knowledge necessary to implement highly successful instructional policy and practice changes, then accountability might not lead to meaningful improvements in student outcomes’ (Figlio & Loeb, 2011, p. 401).

In other terms, the reflexive, critical thinking and research about the institutional design of accountability policies in any context should not be isolated from other education policies (as funding policies, development of schools infrastructure, teacher’s education and professional development, external evaluation and data sets capacities etc.). Further research on these various pre-conditions are thus crucial for an *ex-ante* evaluation of the relevance of accountability policies in each national context, and for the institutional design of these policies adapted to this context.

Conclusion

This think piece on accountability in education aimed at providing some insights into the central issues regarding the development of accountability in education. To conclude, we would like to highlight some key elements. First it appears that accountability and accountability policies in education rest upon certain key features, marking a shift toward ‘new’ accountability in education. This shift implies a move of central governments from input regulation to output-regulation, and it cannot be isolated from other public and education policies that currently take place in many education systems. If some major evolutions (both internal and external to education systems) can be linked to the development of performance-based accountability, we saw that these policies can take various forms depending on the context, even if they share common features.

In order to better understand the rationales and socio-political issues underlying these policies, we proposed a typology of accountability policy and tools in education. Our four types of accountability stand on the one hand on the tools used for accountability purposes, on the other hand on the theory of change embedded in the policy ontology. We highlighted the fact that different logics and different conceptions of what should bring about change underpin these policies, even if the model that is most frequently analyzed remains the one in place in England and in the US (with variations between these education systems), *i.e.* the high-stakes accountability system. Regarding this type of *regulation through strong accountability*, empirical literature and factual findings put forward unintended consequences and uncertainties, rather than a clear consensus regarding the positive effects of this accountability model. We raised the question of the potential link between the side effects of high-stakes AP and a move toward more reflexive forms of accountability in a context of a growing need (or call) for greater accountability.

In our perspective, fundamental issues in relation to accountability policies are: 1) the goal and purpose of these policies 2) the appropriate means for accountability in education 3) the cultural, institutional and socio-economic context of countries and education systems.

Regarding the goals of accountability policies and the tendency to focus on key cognitive learning outcomes as targets that students have to reach, it appears that broader non-cognitive learning could be overlooked in the pursuit of effectiveness. In this regard, setting standards in terms of content, and more importantly, performance standards, is highly political. The definition of levels of performance (from minimum competencies to high

proficiency level) as well as their implications for all the students should be carefully analyzed. Moreover, the definition of cognitive and non-cognitive learning standards should be put in perspective with the missions and goals of the education system at large if we assume that the general role of schooling in terms of social cohesion is crucial. Furthermore, some technical issues also require the development of further research: the validity and reliability of standards and assessment tools that evaluate the progress of students and schools that is the linchpin of performance-based accountability systems.

Given the lack of empirical consensus regarding the effect of accountability policies on the effectiveness and equity of the education system, it is worth highlighting that empirical findings should be used with caution. Further research is clearly needed to further analyze the contextualized processes that affect the link between these policies and their tools, effectiveness and equity in education systems. In addition, further research on other types of accountability policies (*i.e. regulation through neo-bureaucratic accountability, regulation through reflexive responsabilization and accountability, through soft accountability*) is also necessary in order to better understand their implications and effects. Here again, because those who seem to be adversely affected by potential unintended and negative effects of high-stakes accountability systems are the most disadvantaged ones (schools and students), the question of the support, and the balance between control and support provided appears central.

This leads us to conclude on the importance of the cultural, institutional and socio-economic context of countries and education systems. A large body of research has shown that many factors have an impact on school effectiveness, equity in education systems and the quality of schooling. If teacher training and the capacity of schools to attract and retain high-qualified teachers, time and content standards (the increase in the number of school days, the length of the school day, and course requirements [Harris & Herrington, 2006, p. 214]) as well as school composition have an impact on the effectiveness and equity of education systems, 'It is almost a tautology to say that schools are better able to respond to outside pressure when capacity is high' (Harris & Herrington 2006, p. 223). The central question of school capacity building in a context of accountability remains central and needs further investigation.

Bibliography

- Altrichter, H., Heinrich, M., and Soukup-Altrichter, K. (Dir.). 2011. *Shulentwicklung durch Schulprofilierung? Zur Veränderung von Koordinations-mechanismen im Schulsystem*. Wiesbaden: VS Verlag.
- Ball, S. J. 1998. Big Policies/Small World: An Introduction to International Perspectives in Education Policy. *Comparative Education*, Vol. 34, No. 2, pp.119-130.
- Ball, S. J. 2003a. *Class Strategies and the Education Market*. London: Routledge Falmer.
- Ball, S. J. 2003b. The teacher's soul and the terrors of performativity. *Journal of Education Policy*, Vol. 18, No. 2, pp. 215-228.
- Ball, S. J. 2012. *Global Education Inc. New policy networks and the neo-liberal Imaginary*. London & New York: Routledge
- Biesta, G. 2009. Good education in an age of measurement : on the need to reconnect with the question of purpose in education. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)*, Vol. 21, No. 1, pp. 33-46.
- Boltanski, L. and Chiapello, E. 1999. *Le nouvel esprit du capitalisme*. Paris Gallimard
- Bovens, M. 2006. Analysing and assessing public accountability. A conceptual framework. *European Governance Papers (EUROGOV)*. N. C-06-01
- Bovens, M. 2007. Analysing and Assessing Accountability: A Conceptual Framework. *European Law Journal*, Vol. 13, No. 4, pp. 447-468.
- Broadfoot, P. 2000. Un nouveau mode de régulation dans un système décentralisé: l'État évaluateur. *Revue française de pédagogie*, pp. 43-55.
- Brown, P., Lauder, H. and Ashton, D. 2008. Education, Globalisation and the Future of the Knowledge Economy. *European Educational Research Journal*, Vol. 7, No. 2, pp. 131-156.
- Bruns, B., Filmer, D., and Patrinos, H. A. 2011. *Making schools work: New evidence on accountability reforms*. World Bank Publications.
- Buisson-Fenet, H. and Pons, X. 2014. *School Evaluation Policies and Education States: Trends in four European Countries*. Berne: Peter Lang.
- Carnoy, M. and 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.
- Carvalho, L. M. 2009. The N-Moves of Pisa. An introduction to 'Pisa and educational Public Policies' *Sisifo/Educational Sciences Journal*, Vol. 10, pp. 3-13.
- Chang, A., Fisher, D. and Rubenson, K. 2007. *The evolution of professionalism: Educational policy in Provinces and Territories of Canada*. Vancouver: University of British Columbia Press.

- Chiang, H. 2009. How accountability pressure on failing schools affects student achievement. *Journal of Public Economics*, Vol. 93, No. 9-10, pp. 1045-1057.
- Croxford, L., Grek, S. and Shaik, F. J. 2009. Quality assurance and evaluation (QAE) in Scotland : promoting self-evaluation within and beyond the country. *Journal of Education Policy*, Vol. 24, No. 2, pp. 179-193.
- Darling-Hammond, L. 2006. *Standards, assessments, and educational policy: In pursuit of genuine accountability*. Princeton, NJ: Educational Testing Service.
- De Fraine, B., Van Damme, J. and Onghena, P. (2002). Accountability of Schools and Teachers: what should be taken into account? *European Educational Research Journal*, Vol. 1, No. 3, pp. 403-428.
- De Grauwe, A. and Lugaz, C. 2007. Décentralisation de l'éducation en Afrique francophone de l'Ouest: Réalités et défis au niveau local. *International Review of Education*, Vol. 53, No. 5-6, pp. 613-638.
- De Landsheere, G. 1994. *Le pilotage des systèmes d'éducation*. Bruxelles: De Boeck.
- Demeuse, M. and Baye, A. 2009. Indicateurs d'équité éducative. Une analyse de la ségrégation académique et sociale dans quelques pays européens. *Revue Française de Pédagogie*, Vol. 165, pp. 91-103.
- Dubet, F. and Martuccelli, D. 1996. Théories de la socialisation et définitions sociologiques de l'école. *Revue française de sociologie*, Vol. XXXVII, pp. 511-535.
- Dumay, X., Dupriez, V. and Maroy, C. 2010. Ségrégation entre écoles, effets de la composition scolaire et inégalités de résultats, *Revue française de sociologie*, Vol. 3, No. 51, pp. 461-480.
- Dupriez, V. and Mons, N. 2011. Les politiques d'accountability. Du changement institutionnel aux transformations locales. *Education comparée*, Vol. 5, pp. 7-16.
- Dutercq, Y. and Cuculou, S. 2013. La performance comme outil de gouvernance : quelles conséquences sur le travail des enseignants de l'école primaire? Dans C. Maroy (dir.), *L'école à l'épreuve de la performance* (p. 195-210). Bruxelles: De Boeck.
- Figlio, D. and Loeb, S. 2011. School Accountability. Dans E. A. Hanushek, S. Machin & L. Woessmann (dir.), *Handbooks in Economics*, Vol. 3 (p. 383-421). The Netherlands: North-Holland.
- Finnigan, K. S. and Gross, B. 2007. Do Accountability Policy Sanctions Influence Teacher Motivation? Lesson from Chicago's low-performing schools, *American Educational Research Journal*, Vol. 44, No. 3, pp. 594 –629.
- Freidson, E. 2001. *Professionalism. The third logic*. Cambridge (UK): Polity.
- Fullan, M., Rincon-Gallardo, S. and Hargreaves, A. 2015. Professional Capital as Accountability, *Education Policy Analysis Archives*, Vol. 23, No. 15, pp. 1-17.
- Gewirtz, S. 2002. *The managerial school: post-welfarism and social justice in education*. London: Routledge.
- Gewirtz, S., Mahony, P., Hextall, I. and Cribb, A. 2008. Changing teacher professionalism. *International trends, challenges and ways forward*. New York: Taylor & Francis.

- Grek, S. 2009. Governing by numbers: the PISA 'effect' in Europe. *Journal of Education Policy*, Vol. 24, No. 1, pp. 23-37.
- Grodsky, E., Warren, J. R. and Kalogrides, D. 2009. State High School Exit Examinations and NAEP Long-Term Trends in Reading and Mathematics, 1971-2004. *Educational Policy*, Vol. 23, No. 4, pp. 589-614.
- Hanushek, E. A. and Raymond, M. E. 2005. Does School Accountability Lead to Improved Student Performance? *Journal of Policy Analysis and Management*, Vol. 24, No. 2, 297-327.
- Hargreaves, A. 2003. *Teaching in the knowledge society. Education in the age of insecurity*. New York: Teachers College Press.
- Harker, R. and Nash, R. 1996. Academic Outcomes and School Effectiveness: Type 'A' and Type 'B' Effects. *New Zealand Journal of Educational Studies*, Vol. 32, No. 2, pp.143-170.
- Harris, D. N. and Herrington, C. D. 2006. Accountability, Standards, and the Growing Achievement Gap: Lessons from the Past Half-Century. *American Journal of Education*, Vol. 112, No. 2, pp. 209-238.
- Heilig, J. V. and Darling-Hammond, L. 2008. Accountability Texas-style: The progress and learning of urban minority students in a high-stakes testing context. *Educational Evaluation and Policy Analysis*, Vol. 30, No. 2, pp. 75-110.
- Hong, WP. and Young, P. 2008. Does high-stakes testing increase cultural capital among low-income and racial minority student? *Education Policy Analysis Archives*, Vol. 16, No. 6, pp. 2-18.
- Hood, C. 1991. A Public Management for All Seasons? *Public Administration*, Vol. 69, pp. 3-19.
- Hood, C. 1995. Contemporary public management: a new global paradigm? *Public policy and administration*, Vol. 10, No. 2, pp. 104-117.
- Hursh, D. 2005. Neo-liberalism, Markets and Accountability: transforming education and undermining democracy in the United States and England. *Policy Futures in Education*, Vol. 3, No. 1, pp. 3-15.
- Jaafar, S. B. and Anderson, S. 2007. Policy Trends and Tensions in Accountability for Educational Management and Services in Canada. *Alberta Journal of Educational Research*, Vol. 53, No. 2, pp. 207-227.
- Jacob, B. 2005. Accountability, incentives and behavior: the impact of high-stakes testing in the Chicago Public Schools. *Journal of Public Economics*, Vol. 89, pp. 761-796.
- Kogan, M. 1988. *Education accountability: an analytic overview*. London: Dover, NH, USA Hutchinson Education.
- Ladd, H. F and Lauen, F. 2010. Status versus Growth: The Distributional Effects of School Accountability. *Policies Journal of Policy Analysis and Management*, Vol. 29, No. 3, pp. 426-450.
- Lascombes, P. and Le Galès, P. (Dir.). 2004. *Gouverner par les instruments*. Paris: SciencesPo.Les Presses.
- Lee, J. (2008). Is test driven external accountability effective? Synthesizing the evidence from cross-state Causal comparative and correlational studies. *Review of Educational Research*, 78(3), p. 608-644.

- Lee, J. 2010. Trick or treat: New ecology of education accountability system in the USA. *Journal of Education Policy*, Vol. 25, No. 1, pp. 73-93.
- Le Galès, P. and Scott, A. J. 2008. Une révolution bureaucratique britannique? *Revue française de sociologie*, Vol. 49, No. 2, pp. 301-330.
- Leithwood, K., Edge, K. and Jantzi, D. 1999. *Educational accountability: The state of the art*. Bertelsmann Foundation Publishers: Gutersloh/Germany.
- Leithwood, K., Jantzi, D. and Steinbach, R. 1995. An organizational learning perspective on school responses to central policy initiatives. *School Organisation*, Vol. 15, No. 3, pp. 229-244.
- Leithwood, K. and Earl, L. 2000. Educational Accountability Effects: An International Perspective. *Peabody Journal of Education*, Vol. 75, No. 4, pp. 1-18.
- Lindberg, S. I. 2013. Mapping accountability: core concept and subtypes. *International Review of Administrative Sciences*, Vol. 79, No. 2, pp. 202-226.
- Lindle, J.C. 2009. Assessment Policy and Politics of information in G. Sykes, Scheinder, B., D.N. Plank & T.G. Ford (Dir.). *Handbook of education policy research* (p. 319-329). New York: Routledge.
- Linn, R. L. 2000. Assessments and accountability. *Educational Researcher*, Vol. 29, No. 2, pp. 4-16.
- Lingard, B., Martino, W. and Rezai-Rashti, G. 2013. Testing regimes, accountabilities and education policy: commensurate global and national developments. *Journal of Education Policy*, Vol. 28, No. 5, pp. 539-556.
- Mangez, E. 2001. Régulation de l'action éducative dans les années quatre-vingt-dix. *Education et sociétés*, Vol. 8, pp. 81-96.
- Maroy, C. 2006. Les évolutions du travail enseignant en France et Europe. Facteurs de changement, incidences et résistances dans l'enseignement secondaire. *Revue Française de Pédagogie*, vol. 155, pp. 111-142.
- Maroy, C. 2008. Régulation des systèmes éducatifs. Dans A. Van Zanten (dir.). *Dictionnaire de l'éducation* (p. 574-578). Paris: Presses universitaires de France.
- Maroy, C. 2009a. Convergences and Hybridization of Educational Policies Around 'Post-Bureaucratic' Models of Regulation. *Compare. A journal of comparative and international education*, Vol. 39, No. 1, pp. 71-84.
- Maroy, C. 2009b. Introduction to the sub issue 'New modes of regulation of education systems'. *Compare: A Journal of Comparative and International Education*, Vol. 39, No. 1, pp. 67-70.
- Maroy, C. 2009c. Réforme de l'inspection et montée de la régulation par les résultats en Belgique. Incidences des institutions, des intérêts et des idées. Dans G. Pelletier (dir.), *La gouvernance en éducation. Régulation et encadrement dans les politiques éducatives* (p. 53-71). Bruxelles: De Boeck.
- Maroy, C. and Mangez, C. 2011. La construction des politiques d'évaluation et de pilotage du système scolaire en Belgique francophone: nouveau paradigme politique et médiation des experts. *Gouverner l'éducation par les nombres*, pp. 53-76.

Maroy, C. 2012. Towards Post-Bureaucratic Modes of Governance: A European Perspective. In G. Steiner-Khamsi & F. Waldow (dir.), *World Yearbook of Education 2012. Policy Borrowing and Lending in Education* (p. 62-79). London & New York: Routledge.

Maroy, C. (Dir.). 2013. *L'école à l'épreuve de la performance: Les politiques de régulation par les résultats : trajectoires nationales, usages locaux*. Bruxelles: DeBoeck.

Maroy, C. and Voisin, A. 2014. Une typologie des politiques d'accountability en éducation: l'incidence de l'instrumentation et des théories de la régulation. *Education comparée/nouvelle série*, Vol. 11, pp. 31-58.

Maroy, C., Mathou, C., Vaillancourt, S. and Voisin, A. 2014. La trajectoire de la « Gestion axée sur les résultats » au Québec : récits d'action publique, intérêts des acteurs et médiations institutionnelles dans la fabrication d'une politique éducative. Rapport de l'axe 1 du projet de recherche New-Age Québec.

Maroy, C. 2015. Comparing Accountability Policy Tools and Rationales : Various Ways, Various Effects? Dans H.-G. Kotthoff and E. Klerides (dir.), *Governing Educational Spaces : Knowledge, Teaching and Learning in Transition* (p. 35-58). Boston, Rotterdam, Taipei: Sense Publishers.

Maroy, C., Pons, X., and Dupuy, C. (2017). Vernacular Globalisations. *Accountability in Education Policy in France and Quebec*. *Journal of Education Policy*, Vol. 32, No. 1, pp. 100-122.

Mc Neil, L., Coppola, E. Radigan and Vasquez Heilig, J. 2008. Avoidable Losses: High-Stake Accountability and the Dropout crisis. *Education Policy Analysis Archive*, Vol. 16, No. 3.

Mingat, A. and Suchaut, B. (2000). *Les systèmes éducatifs africains : une analyse économique comparative*. De Boeck.

Mons, N. 2008. Évaluation des politiques éducatives et comparaisons internationales. *Revue française de Pédagogie*, Vol. 164, pp. 5-13.

Mons, N. 2009. Les effets théoriques et réels de l'évaluation standardisée. *Revue française de pédagogie*, Vol. 169, pp. 99-140.

Mons, N. and Dupriez, V. 2010. Les politiques d'accountability. *Recherche et formation*, Vol. 65.

Mons, N. and Pons, X. 2006. *Les standards en éducation dans le monde francophone. Une analyse comparative*. Neuchâtel: Institut de recherche et de documentation pédagogique.

Muller, P. 2000. L'analyse cognitive des politiques publiques: vers une sociologie politique de l'action publique. *Revue française de science politique*, Vol. 50, No. 2, pp.189-208.

Normand, R. and Derouet, J.-L. 2011. Evaluation, développement professionnel et organisation scolaire. *Revue française de Pédagogie*, Vol. 174, pp. 5-20.

Osborn, M. 2006. Changing the context of teachers' work and professional development: A European perspective. *International Journal of Educational Research*, Vol. 45, pp. 242-253.

Ozga, J. 2009. Governing education through data in England: From regulation to self - evaluation. *Journal of Education Policy*, Vol. 24, No. 2, pp.149-162.

Ozga, J. and Grek, S. 2012. Governing through Learning. School Self-Evaluation as a Knowledge-based Regulatory Tool. *Recherches sociologiques et anthropologiques*, Vol. 43, No. 2, pp. 35-52.

Pollard, A., Broadfoot, P., McNess, E., Osborn, M. and Triggs, P. 2001. *What Teachers Do: changing policy and practice in primary education*. A&C Black.

Reback, R. 2008. Teaching to the rating: School accountability and the distribution of student achievement. *Journal of Public Economics*, Vol. 92, No. 5-6, pp. 1394-1415.

Rizvi, F. and Lingard, B. (2000). Globalization and Education: Complexities and Contingencies. *Educational Theory*, Vol. 50, No. 4, pp. 419-426.

Sahlberg, P. 2010. Rethinking accountability in a knowledge society. *Journal of Educational Change*, Vol. 11, pp. 45-61.

Spillane, J. P., Reiser, B. J. and Gomez, L. M. 2006. Policy implementation and Cognition. The role of human, social, and distributed cognition in framing policy implementation. Dans M. I. Honig (dir.), *New directions in education policy implementation. Confronting complexity* (p. 47-64). Albany: State University of New York Press.

Steiner-Khamsi, G. and Waldow, F. (Dir.). 2012. *World Yearbook of Education 2012. Policy Borrowing and Lending in Education*. London & New York: Routledge.

Streeck, W. and Thelen, K. 2005. *Beyond continuity: institutional change in advanced political economies*. Oxford: Oxford University Press.

Van Landeghem, G., Van Damme, J., Opdenakker, M. C., De Fraine, B. and Onghena, P. 2002. The Effects of Schools and Classes on Noncognitive Outcomes. *School Effectiveness and School Improvement*, Vol. 13, No. 4, pp. 429-451.

Veselý, A. 2013. Accountability in Central and Eastern Europe: concept and reality. *International Review of Administrative Sciences*, Vol. 79, No. 2, pp. 310-330.

Webb, P. T. 2005. The anatomy of accountability, *Journal of Education Policy*, Vol. 20, No. 2, pp.189-208.

Wolf, I. and Janssens, F. J. G. 2007. Effects and side effects of inspections and accountability in education: an overview of empirical studies, *Oxford Review of Education*, Vol. 33, No. 3, pp. 379-396.

Woessmann, L. 2007. International Evidence on School Competition, Autonomy and Accountability: a Review. *Peabody Journal of Education*, Vol. 82, No. 2-3, pp. 473-497.

APPENDIX 1: Accountability, central tools

Every accountability system stands on common policy tools, more particularly on four components: 1) standards (content and performance standards), 2) large-scale assessment and standardized testing, 3) accountability mechanisms that formalized the action of providing an account, 4) consequences for several actors (individuals or institutions). If these policy instruments and tools remain central for accountability purpose, the nature and range of tools can vary from one education system to another.

1. Standards

If the type and object of standard can widely vary (from teacher competencies and level of qualification, to student behavior etc.) curriculum and performance standards are central to accountability policies. Curriculum standards (which may also be called content or academic standards, core content frameworks) refer to knowledge that students should acquire (Lindle, 2009); they define what students should know, what has to be taught and what matters. In the United-States, the standard-based reforms of the 1990s aimed at establishing common standards within the curriculum of all States, to align several instruments on this basis (such as teacher education and professional development mechanisms based on these standards). The question of the tight alignment between various central instruments of the education system is at the heart of these large-scale reforms, as pointed out by Darling-Hammond (2006, p. 5):

if governments can clearly specify what students should know and be able to do, then these standards can shape curriculum, assessment, instruction, teacher education, professional development, the allocation of resources, and all of the other elements of the educational system

Content standards provide guidance on learning activities, while performance standards deal with the level of performance that has to be reached by students, schools and teachers. Performance standards introduce and fix some guidelines on how well students and schools should and might perform (Lindle, 2009). With it, comes the question of ‘how good is good enough’ (Linn, 2000, p. 9)? In England, the 1988 Education Reform Act established ‘the imposition of a National Curriculum, national testing at Key Stages (7, 11 and 14 years) and explicit targets and expectations of achievement’ (Osborn, 2006, p. 245). In the United-States, No Child Left Behind, following the same standards-based reform path, introduced the obligation for States to determine performance standards and ‘what it means to be proficient on the state assessments and to evaluate schools based on whether their students, in aggregate and by subgroup, are progressing adequately toward an ultimate goal of 100% proficiency by 2014’ (Figlio & Loeb, 2011, p. 385). Content and performance standards then became the basis of a high stakes accountability system run at the federal level.

2. Assessment and testing

Assessment is the central piece of accountability policies (Linn, 2000; Carnoy & Loeb, 2002; Lindle, 2009). If assessment and testing are not new in many education systems, the use of large-scale assessment as a policy instrument (e.g. in order to steer the system) and of student test results for accountability purposes in central to performance-based accountability systems are fairly recent. What is also relatively new is the use of student test scores in order to measure teacher, school and district performance. Standards, assessment tools and accountability mechanisms are becoming increasingly connected. Indeed, ‘during the past 20 years test-based accountability has held school, teachers, and students increasingly accountable for learning as measured by knowledge tests’ (Carnoy & al. 2003; Hamilton & al. 2002; Jaafar and Anderson 2007; Ladd and Fiske 2003; Møller 2009; Popham 2007 in Sahlberg, 2010, p. 50)

As highlighted by Webb, ‘this so-called “new” form of accountability intended to make educators’ work more visible through inspections, observations, performances, and public reporting of test scores’ (Webb, 2005, p.

192). Because test scores became the basis on which performance is evaluated, policy makers need data and results in order to monitor system progress:

Testing can be used in several ways. It can be an indicator to tell administrators and teachers whether they are reaching the organization's goals and to provide information on which elements of the curriculum are reaching students and which are not. It can be used as a measure of success or failure in an incentive system. It can be used as a gauge to increase standards, to assess curricula, or to provide technical assistance. It can be used as a mechanism to allocate additional resources in order to improve outcomes for groups having difficulty reaching the standards. (Carnoy & Loeb, 2002, p. 307)

3. Accountability mechanisms

Mechanisms that organize the action of rendering an account are also central to accountability policies and systems. These mechanisms formalize the 'obligation to explain and justify conduct' (Bovens, 2006, p. 9) that is constitutive of accountability relationships. As pointed out by Bovens (Idem.) explanations and justifications:

'involve not just the provision of information about results, but also the possibility of debate, of questions by the forum [e.g individuals, superiors, agencies to whom the account has to be given] and answer by the actor, and eventually of judgment of the actor by the forum'

Even if the nature of obligation and the mechanisms that organize the action of providing an account are not always very detailed and clear in the literature, for Wagner (1987 in Leithwood & Earl, 2000, p. 8), the obligation or requirement for an account to be given is central to accountability.

These mechanisms can be, in the case of hierarchical accountability, the obligation for the teacher to justify his student's results to the head-teacher, or for the head-teacher to provide information to a district administrator regarding actions taken at the school level in order to improve student test scores. For example, in the case of Quebec (Canada), accountability mechanisms take the form of formal accountability documents linked to contract agreements. These accountability documents include several data regarding schools and district results (in terms of achievement rates, dropout rates etc.) linked to the targets and goals included in contract agreements. They constitute the basis of an annual meeting (at least) between school principal and district administrator, during which the school principal has to justify the school's results and the measures taken in order to reach the target. On the same basis, the district administrator has to justify his district's results and actions taken in order to improve the results in line with the target fixed in the contract agreement between the ministry and the districts (for further information see Maroy & al., 2015, forthcoming).

4. Incentives, sanctions and rewards

Last but not least is the very well documented system of incentives for several actors. Many authors have pointed out that repercussions for actors (individuals or institutions – schools and districts) linked to test results are key instruments of accountability systems. They raise the following questions: which actors should be targeted at which level, and who has to be held accountable for student achievement and face the consequences in the case set requirements are not met?

Incentives and consequences for actors can be more or less formalized (Bovens, 2006). In the case of less formalized consequences, for example, 'it seems likely that some response will occur, but this response often will be muted and almost by definition unpredictable' (Leithwood & Earl, 2000, p. 8). Incentives and consequences can take several forms, ranging from teacher's intrinsic motivation to help students to learn and public recognition, to material and financial sanctions and rewards, such as additional resources or funding cuts for schools, salaries and financial bonuses for individuals - career advancements, transfers, firing. They often constitute the basis on which accountability systems can be differentiated or characterized, as highlighted in our typology.