Global Perspectives on High-Stakes Teacher Accountability Policies: An Introduction

Jessica Holloway
Research for Educational Impact (REDI) Centre, Deakin University
Australia

Tore Bernt Sørensen
University of Bristol
United Kingdom

&

Antoni Verger
Universitat Autònoma de Barcelona
Spain

Abstract: The aim of this special issue, "Global Perspectives on High-Stakes Teacher Accountability Policies", is to provide insights into a diverse set of policies focusing on teachers' accountability, including the underpinning ideas and cultural and socio-economic contexts of these policies, as well as their effects on teachers' work, the teaching profession and the broader educational environment. While these articles highlight the influence of the "global testing culture" on education systems world-wide, they also demonstrate the need for understanding accountability systems as context-specific. As such, we urge scholars to consider the social, historical, political and geographical contexts within which their research is situated and to promote a research agenda that looks at the specific responses and effects that accountability policies produce in different regulatory settings. This introductory article, first, clarifies the main focus and conceptual framework of the special issue and, second, presents an overview of the papers included in the issue and their main contents.

Keywords: teacher accountability; teacher evaluation; high-stakes testing; global testing culture

Perspectivas globales sobre políticas de rendición de cuentas de los profesores: Una introducción

Resumen: En esta edición especial, "Perspectivas globales sobre las políticas de responsabilidad del profesor de High Stakes" examinar un conjunto diversificado de políticas centrado en la rendición de cuentas de los profesores, incluidas las ideas subyacentes, y los contextos culturales y socioeconómicos de estas políticas, así como los Sus efectos sobre los profesores 'trabajo, la profesión docente y el ambiente educativo más amplio. Mientras estos artículos destacan la influencia de la "cultura de pruebas globales" en los sistemas educativos en todo el mundo, también demuestran la necesidad de entender los sistemas de responsabilización como específicos del contexto. Como tal, pedimos a los estudiosos que consideren los contextos sociales, históricos, políticos y geográficos dentro de los cuales su investigación se sitúa y promueva la agenda de investigación que analice las respuestas y los efectos específicos que las políticas de responsabilidad producen en diferentes configuraciones regulatorias. Este artículo introductorio, en primer lugar, aclara el enfoque principal y el marco conceptual de la cuestión especial y, en segundo lugar, presenta una visión general de los trabajos incluidos en la cuestión y sus principales contenidos.

Palabras-clave: rendición de cuentas de los profesores; Evaluación docente; Pruebas de alto riesgo; Cultura de pruebas global

Perspectivas Globais sobre políticas de prestação de contas dos professores: Uma introdução

Resumo: Esta edição especial, "Perspectivas Globais sobre políticas de responsabilização professor da High Stakes" examinar um conjunto diversificado de políticas com foco na prestação de contas dos professores, incluindo as ideias subjacentes, e os contextos culturais e sócio-econômico destas políticas, bem como os seus efeitos sobre os professores 'trabalho, a profissão docente eo ambiente educacional mais ampla. Enquanto esses artigos destacam a influência da "cultura de testes globais" nos sistemas educacionais em todo o mundo, eles também demonstram a necessidade de entender os sistemas de responsabilização como específicos do contexto. Como tal, pedimos aos estudiosos que considerem os contextos sociais, históricos, políticos e geográficos dentro dos quais sua pesquisa se situa e promova a agenda de pesquisa que analise as respostas e os efeitos específicos que as políticas de responsabilidade produzem em diferentes configurações

regulatórias. Este artigo introdutório, em primeiro lugar, esclarece o foco principal e o quadro conceitual da questão especial e, em segundo lugar, apresenta uma visão geral dos trabalhos incluídos na questão e seus principais conteúdos.
Palavras-chave: prestação de contas dos professores; Avaliação de professores; Testes de alto risco; Cultura de testes globais

Introduction

The aim of this special issue, "Global Perspectives on High-Stakes Teacher Accountability Policies", is to provide insights into a diverse set of policies focusing on teachers' accountability, including the underpinning ideas and cultural and socio-economic contexts of these policies, as well as their effects on teachers' work, the teaching profession and the broader educational environment. This introductory article, first, clarifies the main focus and conceptual framework of the special issue and, second, presents an overview of the papers included in the issue and their main contents.

High-stakes Accountability and Teachers

Teachers are currently at the center of education reform agendas globally. During the past two decades, teachers—and especially teachers in primary and secondary education—have been subject to unprecedented interest and scrutiny. Main policy actors nationally and internationally conceive teachers as key factors of education systems' effectiveness (Mourshed et al., 2010), and have sought to align teachers in the pursuit of economic competitiveness (Robertson, 2012). Under the master narrative of the knowledge-based economy, the Organisation of Economic Co-operation and Development (OECD) has touted teachers as '*front-line workers responsible for engaging students and promoting their learning*' (OECD, 2014, p.32), which raises the issues of how to conceptualize, measure and promote 'teacher' or 'teaching effectiveness' (OECD, 2013) – and how to hold teachers accountable for what they do in their classrooms and more explicitly for the outcomes of their work.

Stefan Hopmann argued 10 years ago that we live in an "age of accountability", yet he also pointed out that "*accountability concepts change over time and are different in different places*" (Hopmann, 2007, p.369). In this respect, we should note that the accountability term remains ambiguous and multifaceted, with policy actors employing various and often implicit definitions of accountability, its elements and its limits. Overall, accountability is a concept with multiple meanings, and a policy programme that covers a broad range of policy options and models including political, legal, bureaucratic, or market forms of accountability (Verger & Parcerisa, 2017). Capturing the meaning and implications of accountability becomes even more challenging *in* – and not least *across* - other non-English language contexts since the term "accountability" tends to be difficult to translate, with no single term corresponding to the English/American concept of accountability in most other languages.

Despite this conceptual complexity, there is consensus among scholars that procedures related to accountability increasingly permeate societies and key policy domains. Whilst it should be emphasized that these processes unfold differently depending on the context, the spread and impact of accountability policies indicate fundamental changes in the "feedback loop" (Thrift, 2005, p.6) by which societies deal with themselves and with ill-defined "wicked problems" (Rittel and Webber, 1973), such as those related to education, teaching and learning. The reinforced emphasis on accountability thus has significant implications for education provision and the redistribution of resources and responsibilities within educational systems (Hopmann, 2007).

Melvin Dubnick's (2006) discussion of "orders of accountability" takes this point further and highlights the semiotic richness of the accountability term. Dubnick distinguishes between four such orders of accountability:

1. *Performative accountability*: arising in face-to-face relations involving direct and explicit acts of account giving.
2. *Regulatory accountability*: the "control of conduct" characterized by how well one follows the guidance, rules and operating standards set by a resource giver, often on the basis of law and constrained and directed by the "code" and dominant rationales of the task environment. This might not involve direct and explicit account giving, but implies the potential "threat" of being called to performative account.
3. *Managerial accountability*: the use of accountability as a means to motivate and elicit purposive behaviour, such as better service and effectiveness. Centered on the use of incentives and/or sanctions, the focus of managerial accountability is on designing task environment conditions that encourage actions leading to improvement rather than controlling or constraining those actions *per se*. As a modality of power seeking to frame individual and collective actions, managerial accountability has been harnessed politically as "promises" for a wide range of socially desirable aims, including equity, democracy, ethical behaviour, and, especially, improved performance and quality in the production of goods and services.
4. *Embedded accountability*: centered on the internalization of the norms, values and expectations to a degree that the embedded sense of "being accountable" will guide behavior without necessarily having to resort to the orders of performative, regulatory or managerial accountability. Embedded accountability is related to professionalism and a sense of moral responsibility, integrity and authority that come from the commitment to live up to commonly shared expectations to the capabilities and qualifications of professionals. Dubnick points out that embedded accountability stands out as both foundational and aspirational for the modelling of accountability because it—like contemporary governance overall—rests on a foundation of legitimacy and trust, rooted in the belief that the various actors (policy-makers, managers, educators, administrators, etc.) and components of a given system are aligned in terms of having and reflecting common expectations and operating under the assumption that the actors are accountable for what they do in their distinctive roles.

Accountability, Governance, and Power

Given the political implications of accountability and the ways in which accountability might be adopted as a policy instrument, we must also consider governance and power. Of all the orders of accountability defined by Dubnick, the managerial order, which usually relies on the use of standardized tests, is the one that has become more predominant globally. This is consistent with the broader governance scholarship, which conceives accountability as a tool of New Public Management (NPM) meant to ensure that organisations and individuals provide services according to the goals set for them or agreed upon with them (Hood, 1995; Hood & Margetts, 2007). Managerial accountability has gained international popularity and attraction for many different reasons, including the consolidation of NPM as a hegemonic public sector reform paradigm. Among other implications, NPM has contributed to education reformers increasingly promoting managerial

school governance styles and educational services oriented toward the achievement of tangible and measurable results (Kalimullah, Ashraf, & Ashaduzzaman, 2012).

In this sense, the emphasis on accountability as a managerial policy instrument has shifted the reliance from implicit values and norms towards more explicit and specific standards and contracts. This change has had major repercussions for established notions of professionalism, not least in education sectors across the world. In this respect, Dubnick's (2006) discussion of U.S. education reform and school accountability is particularly interesting. Dubnick points out that during the 1980s and 1990s the emphasis moved from regulatory to managerial orders of accountability—as a reaction to an allegedly excessive bureaucratization of the post-World War II era—accompanied by the 2002 No Child Left Behind (NCLB) Act, which enforced high-stakes testing regimes to improve school performance (as measured by student test scores), thereby challenging established notions of teacher professionalism. Effectively, the particular combination of managerial accountability and high-stakes testing led to an excessive focus on first order performative accountability, resulting in the adoption of "gaming strategies" among schools and teachers to improve scores and meet performance standards. On this basis, Dubnick (2006) speculated that, by relying on managerial accountability to resolve problems associated with excessive regulatory accountability, reformers undermined the power and authority of professional norms that were previously thought to underpin teacher commitment and accountability.

In a similar manner, the post-World War II era was defined by a time when social matters and risks (e.g., health, education) were entrusted to the professional institutions that held the specific expertise related to the associated complex (and often ill-defined) problems that society presented. These organizations maintained considerable autonomy in terms of professional judgment, having no fixed boundaries to how they approached the various complex problems that emerged (Lingard, Sellar & Lewis, forthcoming). Through broadening the scope and differentiating the means of such approaches, fields like education underwent a massive expansion. These developments eventually resulted in what many deemed the "crisis of the welfare state", where fears about resource scarcity brought about mistrust and anxiety related to whether comprehensive aid would be sustainable in the future. In response, the complex, ill-defined problems were transformed into more narrowly and clearly defined—yet easily adjustable and volatile—expectations as to what can be achieved with a given amount of resources through standard-setting, indicators and benchmarking. Thus, for teachers, the late 1990s and early 2000s brought demands for new forms of accountability due to concerns about the quality and effectiveness of teaching, as well as teacher education (Tatto, 2011).

Hopmann (2007) argued that this "management of expectations" is, in principle, more target-oriented, yet it comes with the price that whatever does not fit into the specific expectation regime becomes marginalized (pp. 370-373). In education, "good instruction" thus became defined primarily by its measurable outcomes. In cases of ambiguity and conflicting goals, the balance would tend to tip towards more well-defined expectations. One major consequence of the shift towards "management of expectations" is that more complex and contested issues (e.g., issues such as bullying, gender equity, migration inclusion, second language learning, etc.) become targeted through intervention programs of limited duration, scope, and impact, that is, programs which are meant to ensure the public that no ill-defined problem is left behind but do not necessarily disturb the status quo. Moreover, in accordance with the ideological leanings of NPM, the shift towards an accountability regime based on "management of expectations" had important political implications. Among other things, it entailed that educational provision was to be opened up to the private sector since, as long as expectations are met, education can be provided outside traditional public institutions and previously established notions of professionalism.

High Stakes Accountability

Standards-based testing frameworks are usually considered indispensable to bring about the data underpinning the accountability regime. Tests—and the related accountability policies—are deemed high-stakes if they have "real or perceived consequences for students, staff, or schools" (Chapman & Snyder, 2000, p. 458). Accountability policies and practices have increasingly focused on measurable performance outcomes, which has, in turn, provided the conditions whereby teachers have become subjected to mechanisms for quantifying their performance and incentivizing their abilities to increase student test scores (Ball, 2015; Smith, 2014; Verger & Parcerisa, 2017).

The thickening of the global education policy field over the past decades is closely associated with the spread of a "global testing culture", centered on the regular administration of high-stakes standardized tests (Smith, 2016). For example, student test score results on national and international tests, such as the Programme for International Student Assessment (PISA) and Trends in International Mathematics and Science Study (TIMSS), have become of particular concern to subnational, national, and international stakeholders who use such outcomes to compare educational systems and actors, set policy agendas, and/or identify areas in need of attention or improvement. While carried out in various ways and to various degrees around the world, a heightened focus on student testing is nevertheless a global phenomenon, as major transnational agencies such as the OECD, the United Nations Educational, Scientific and Cultural Organization (UNESCO), and the World Bank have invigorated a global interest in school performance measures that rely on standardized tests to measure, evaluate, and compare national and subnational education systems. During the last 15-20 years, these organizations have been engaged in the development of statistical indicators and the quantification of teacher performance to an unprecedented degree (Rutkowski, 2008; Sellar & Lingard, 2013; Verger & Curran, 2014). Moreover, the more recent adoption of a performance-oriented Education 2030 Framework for Action under the UN Sustainable Development Goals is likely to further promote this line of thinking in development contexts (UNESCO, 2017).

Despite the various ways in which national systems have engaged with this trend, similar logics and goals appear to be driving the development of teacher accountability policies and instruments that rely on statistical measures of teacher quality. Specifically, an increased focus on outputs (e.g., performance scores, comparative rankings) over inputs (e.g., teacher preparation, credentialing) is consistent with a broader (re)articulation of schooling as a process that can *and should* ideally be quantified and measured, or with the idea that teachers can be materially dis/incentivized to perform better and to increase student- and school-level performance results. We see this in merit pay schemes, as well as threats of termination or loss of tenure. Some countries such as the United States and England have begun explicitly linking student test scores to teacher evaluation via value-added models (VAMs), which are statistical instruments designed to measure school and teacher effects on student learning over time (AERA, 2015; Leckie & Goldstein, 2017). While most countries have heeded caution with VAMs specifically, many have embraced other high-stakes teacher accountability policies, often predicated on the same logics that undergird VAM-use (OECD, 2014; Sahlberg, 2011). These "high-stakes" approaches to teacher accountability have been extensively critiqued in the literature for various consequences related to the weakening of teacher morale (Collins, 2014), the narrowing of the curriculum (Au, 2007) and the problems associated with the reliability, validity and fairness of such accountability systems (Amrein-Beardsley, 2014; Darling-Hammond, 2015). However, some scholars have also urged us to question the notionally "low-stakes" approaches to teacher accountability (e.g., the use of test results for instructional purposes) and the extent to which these approaches produce similarly narrow views of education and teacher

quality, a point that is raised in two of the articles included in this special issue (see Lewis, 2017, and Thiel, 2017).

Nonetheless, again, managerial accountability (cf. Dubnick, 2006) is not a monolithic category. Managerial accountability policies differ according to the level and nature of the consequences for the educational actors that are being held accountable, mainly teachers, principals and schools. To understand these policies, analysts must consider the ways that the "problem" of accountability are represented by policy actors (the "problem-setting"), the contextual dynamics that produced the "problem" in the first place (Peters, 2015), as well as the contextual conditions of the policy implementation and enactment (Verger & Parcerisa, 2017). In Ray Pawson's (2002) words: "… it is not 'programmes' that work; rather it is the underlying reasons or resources that they offer subjects that generate change. … Whether the choices or capacities on offer in an initiative are acted upon depends on the nature of their subjects and the circumstances of the initiative," (p. 342). Hence, the articles included in this special issue offer distinctive and context-sensitive perspectives on the adoption of high-stakes teacher accountability policies and a range of explanations for the often disappointing and adverse outcomes, when stated intentions of such policies fail to materialize or lead to undesired or unexpected behaviours. In this sense, the special issue provides theoretical elements and empirical evidence for challenging the common assumption that "high stakes" as a component of accountability policies will produce, under all conditions, better student learning outcomes and more responsive and innovative working environments for teachers.

Global Perspectives on Accountability

In documenting, understanding and explaining the background, working and impact of high-stakes teacher accountability policies, the special issue adopts a "global perspective", in the sense that the eight articles together cover very different social and material contexts in various regions of the world (North and South America, Europe, and East Asia). Moreover, the eight articles to various degrees seek to contextualize their object of study in the wider landscape in which international, national, sub-national, and local levels are engaged in the non-zero-sum game of multi-scalar education governance (Dale, 2005). The articles represent a remarkable diversity in this respect, with five articles focusing on subnational political entities (New Jersey, Rio de Janeiro, Shanghai, four German Länder, and three school districts in New York, Virginia and Texas), two articles on national entities (Chile and Mexico), and one article presenting an international comparative study based on data from the OECD Teaching and Learning International Study (TALIS). We should emphasize that our intention with adopting a "global perspective" is not to speculate in universal truth claims about the outcomes, and how and whether high-stakes teacher accountability policies might be beneficial or not. Rather, our objective is to show that the understanding and explanation of accountability policies requires close consideration of context-specific factors, and how these interact with the "program ontology" , that is, the theories of socio-economic dynamics underpinning the policy in question, and how it will bring about the intended changes (cf. Pawson, 2002).

Overview of the Articles

The impetus for this special issue is the widely held acknowledgement that there is a need for documenting, understanding and explaining the background and nature of high-stakes teacher accountability policies from a wide range of perspectives and within and across a broad range of territories, as well as more general analyses on the globalization of accountability in education and

the role of major transnational agencies like the OECD in this terrain. In the Call for Papers, we sought to capture this aspiration with the following prompts:

- What are the historical, political, and/or social backgrounds for accountability policies in different educational contexts?
- How have such policies developed and been enacted over time, and to what effect?
- How do teacher accountability policies merge and/or interact with market-based education reforms?
- Which policy actors are engaged in the formation of such policies, and how does this relate to the thickening of the global educational policy field?

The articles included in this special issue address one or more of these questions, offering compelling cases that bring to bear a critical lens on the multi-faceted approaches to teacher accountability that are currently in use around the world. An important point is that each of the eight articles is based on empirical inquiry relating to the particular spatial contexts in question. Moreover, two studies (Steven Lewis' piece on the OECD program PISA for Schools, and Corrie Thiel's study on side effects of accountability policies) discuss low- and high-stakes accountability policies, and thereby help to broaden and clarify the debate, as well as encourage reflection on the current state of research on the topic.

This section provides a summary of the articles, followed by a brief discussion of further provocations to consider. First of all, we might with reference to Dubnick (2006) point out that all articles, more or less explicitly, address the implications and outcomes following the shift from regulatory towards managerial accountability. Together, the articles show that the four orders of accountability in practice might be intertwined, complementary, or contradictory, depending on the case in question. In this way, the articles overall indicate that the orders of accountability are rarely complementary of each other, with the first three orders neatly nested within the fourth; that is, accountability, as an embedded rationale, directs managerial, regulatory and the performative orders of accountability. Several articles highlight the tensions, pressures, discontent, and disorientation resulting from the introduction of high-stakes managerial accountability policies, which challenge prevailing notions of teacher professionalism and the associated norms, values and practices. The reliance on student performance outcomes as the basis for teacher accountability stands out as a factor in this respect, as well as the issue of what gets left out and hence un-accounted for with such policies and practices. In many ways, the articles thus corroborate Dubnick's (2006) point that the shift from regulatory to managerial accountability results in "mission problems" due to the erosion of teachers' "embedded accountability".

Overall, the articles suggest a disconnect between the accountability policies in place and teacher perceptions or attitudes; either teachers do not approve of the policies (New Jersey and Rio de Janeiro), or they understand them differently than intended by the reformers (Shanghai), or they are left out of the policy formation process (Mexico), or they seek to 'talk back' to policy-makers by engaging with alternative and 'richer' forms of accountability (PISA for Schools in the US). To begin with, William C. Smith and Katarzyna Kubacka use data from the OECD TALIS program to provide an overview of high-stakes teacher accountability in 33 countries. Situating their study within the larger "Global Testing Culture", they present the following important findings: (1) the use of student test scores in teacher appraisals is nearly universal; (2) test scores are rarely included in isolation; and (3) a large number of teachers report not receiving any feedback on the component(s) used to make high-stakes decisions. The authors make the point that higher focus on students' test scores is closely associated with lower levels of perceived feedback utility. Perhaps this is a particularly important finding to help frame the seven remaining articles of the special issue that,

with various emphases, address and elaborate on these and related issues in particular socio-economic and cultural contexts.

In the second article, Oren Pizmony-Levy and Ashley Woolsey help fill a much-needed gap in the teacher accountability literature by surveying teachers about their experiences with the New Jersey (USA) teacher evaluation system. This study is set against a politically contentious backdrop—one perpetuated by Governor Chris Christie's disparaging remarks against the local teachers' union. Ultimately, their findings point to an interesting paradox, whereby teachers generally agree that accountability is important and that teachers' effectiveness should be measured. However, the teachers overwhelmingly disapproved of the current directions the state was going with its teacher evaluation system. The authors draw attention to the close relationship between positive attitudes towards the system and the teachers' perceptions that the policy would have a direct and positive effect on teaching and learning. The teachers of this study also were more likely to support the system if they felt supported.

In the following article, Priya La Londe investigates teachers' perspectives on the aims of performance-based compensation in China. Due to outstanding rankings in the OECD PISA program, the education system in Shanghai has in recent years built a global reputation as a high-performing, accountability-driven education system. La Londe's analysis adds important nuances to this apparent success story by showing that there is a considerable gap between the intended aims of the policy on teachers' performance-based compensation currently in place, and teachers' perceptions of the same policy. Thus, whilst performance-based compensation was introduced to improve teaching quality, teachers thought that merit pay was meant to increase job satisfaction, and participation in teacher and student development activities. On this basis, La Londe puts forward the powerful argument that the policy of performance-based compensation in place in Shanghai represents a "failed marriage between standardization and incentivism". In this way, La Londe's findings indicate that the "management of expectations" (cf. Hopmann, 2007) was poorly executed—with teachers thinking that key notions, such as quality teaching, were vaguely articulated in the policy. More specifically, the program ontology (cf. Pawson, 2002) of the policy in Shanghai, based on creating incentives through coupling performance assessment with teacher compensation, appears mismatched towards those meant to be incentivized.

Lluís Parcerisa and Alejandra Falabella analyze the introduction of a new wave of accountability reforms in Chile, which is the country with the most marketized education system in the world since the government of the *Junta Militar* introduced an ambitious reform to promote freedom of school choice in the 1980s. In Chile, the adoption of a new wave of accountability policies in the educational system does not respond to the usual goal of promoting more competitive behaviors and incentives within the educational system; on the contrary, these policies aim at addressing the problems generated by an excess of market competition in education. In other words, accountability measures have been enacted, especially in the first decade of the 2000s, as a way to correct persisting market failures in a highly deregulated and unequal educational system. Nonetheless, as Parcerisa and Falabella show, the adoption and intensification of accountability measures in Chile in the 2000s was the result of a difficult political compromise between the center-left and the conservative political coalitions in a political context of intense social protests against marketization and profit in education. For the predominant political forces at that time, accountability was conceived as an attempt to promote equity in education through higher levels of schools control, but without having to alter the market logics and mechanisms that were –and still are—at the center of the governance of the Chilean education system.

Jaime Echávarri Valdez and Cecilia Perez Sanginés analyze the introduction and evolution of accountability measures in Mexico through different and successive waves of educational reforms.

Mexico started adopting education evaluation and accountability reforms focusing on teachers in the 1980s, in the context of a strong international debt crisis and in the name of promoting educational modernization. To some extent, Mexico has pioneered the adoption of teachers' evaluation policies attached to teachers' career development in the Latin American region. Echávarri and Peraza show how this is due to the presence of a strong epistemic community and a well-articulated network of policy actors that have been persistently behind the enactment of this type of accountability reforms in Mexico. Nonetheless, these reforms, apart from being very much contested in numerous Mexican states, have not delivered on their promise of raising educational quality standards and better educational results. According to Echávarri and Peraza, this failure is due to the fact that accountability reforms put an excessive focus on teachers and schools, and omit the structural and systematic problems that the Mexican educational system faces. Indeed, these reforms have evolved in parallel to an increasing stagnation of public education and to the impoverishment of students' families, schools' facilities and teachers' working conditions in Mexico. Thus, the main issue with accountability reforms is not necessarily related to the nature of the policies being enacted, but to the educational problems that accountability policies are not addressing and are not able to address by themselves.

Rolf Straubhaar's paper focuses on the educational and ideological gap between those designing accountability schemes in education, and those that are expected to implement these schemes. To this purpose, he has conducted ethnographic interviews with Ministry officials, policy-makers and teachers in Rio de Janeiro, one of the cities in Brazil that has experimented further with high-stakes accountability schemes and value-added models in the last years. The paper shows that policy-makers design accountability programmes informed by notions of management that come from business and the private sector; however, most teachers reject these notions and consider that the promoters of high-stakes accountability systems do not understand classroom and school dynamics sufficiently, and are not able to capture correctly what is good quality education with over-simplified test scores. Straubhaar's findings are relevant from the perspective of the realist evaluation approach we have referred to in this introduction to the special issue. The fact that the programme ontology of valued-added model interventions is not shared by the subjects of the intervention is a predictor of important implementation issues, and explains why, as shown in the paper, the final outcomes of the accountability interventions in Rio are quite far from the original goal of producing genuine improvements in instructional practices and in students' learning.

Steven Lewis shows that the OECD program PISA for Schools and the associated Global Learning Network is adopted by schools to broaden the debate and professional development on teaching and learning beyond the currently narrow focus on performance standards and student learning outcomes in U.S. education policy. Hence, schools and teachers in the USA are calling for alternatives to current high-stakes teacher accountability policies, and on this basis Lewis calls for more progressive reconceptualizations of accountability, such as the "rich accountabilities" model (Lingard et al. 2016), which combines "bottom-up" and "top-down" ways of holding systems, schools and communities accountable in order to acknowledge the broader societal purposes of education. Considering the substantial critique of the main OECD PISA program, it is remarkable that the schools included in Lewis' study embrace the OECD PISA for Schools program as means to challenge current teacher accountability policies in the USA. With his careful analysis of how PISA for Schools enables an intensification of global-local connectivity in education governance, Lewis highlights the capacities of the OECD in setting the agenda and shaping views of education (cf. Lukes, 2005), which might overlap, complement, or challenge national and subnational policies - and in some contexts potentially empower teachers to "talk back" to policy-makers. However, as Lewis also notes, it is rather privileged schools that have taken part in PISA for Schools and the

Global Learning Network in the USA, thus potentially further sidelining the voice of teachers working in more disadvantaged schools from the debate on how schools and teachers should be held accountable for what they do.

Finally, in the article, Corrie Thiel, Sebastian Schweizer, and Johannes Bellmann investigate the "side effects of accountability" in education, challenging the commonly accepted claim that such effects are only associated with high-stakes contexts. They use data gathered from interviews and surveys with teachers and school principals in four German federal states (*Bundesländer*) to provide evidence of side effects occurring in no- and low-stakes contexts. The authors distinguish between *adaptive behavior*, which is related to the alignment of teaching behavior with the accountability mechanisms, and *evasive behavior*, which is related to the circumvention of accountability consequences (e.g., cheating). They argue that adaptive behavior is pervasive across the four states and likely an inevitable systematic factor. Evasive behavior, however, is less prominent and more closely associated with the degree to which the educators are subjected to market and bureaucratic pressures. Their work urges us to look beyond the low- versus high-stakes debate and to consider the ways in which accountability regimes create (potentially negative) effects, regardless of the stakes attached.

Conclusion

The articles in this special issue present further evidence that we indeed live in an "age of accountability" (cf. Hopmann, 2007). Globally, managerial and performative accountability policies increasingly incorporate student test scores into teacher appraisal systems, though doing so in varied ways and with varied levels of stakes attached to the system outputs. But, as raised by some of the authors here, does the juxtaposition of low-stakes versus high-stakes which tends to dominate the debate on accountability limit the ways in which we can think about the performative nature of teacher accountability and the logics that undergird the systems in place? In other words, should we focus more on the ways in which accountability systems produce particular types of teachers (and working and learning conditions more broadly) rather than focusing on the degree to which low/high stakes matter?

Further, while these articles highlight the influence of the "global testing culture" on education systems world-wide, they also demonstrate the need for understanding accountability systems as context-specific. As such, we urge scholars to consider the social, historical, political and geographical contexts within which their research is situated and to promote a research agenda that looks at the specific responses and effects that accountability policies produce in different regulatory settings. We compel the same of policymakers as they develop teacher accountability policies that are most appropriate for their own education systems and better aligned with quality and equity in education more broadly.

References

AERA [American Educational Research Association] (2015). AERA Statement on Use of Value-Added Models (VAM) for the Evaluation of Educators and Educator Preparation Programs. *Educational Researcher, 44*(8), 448–452. http://dx.doi.org/10.3102/0013189X15618385

Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher, 36*(5), 258-267. https://doi.org/10.3102/0013189X07306523

Ball, S. J. (2015). Education, governance and the tyranny of numbers. *Journal of Education Policy*, *30*(3), 299-301. https://doi.org/10.1080/02680939.2015.1013271

Chapman, D. & Snyder, C. (2000). Can High Stakes National Testing Improve Instruction: re-examining conventional wisdom. *International Journal of Educational Development*, 20, 457-474. http://dx.doi.org/10.1016/S0738-0593(00)00020-1

Dale, R. (2005). Globalisation, knowledge economy and comparative education. *Comparative Education,* 41(2), 117-149. https://doi.org/10.1080/03050060500150906

Darling-Hammond, L. (2015). Can value added add value to teacher evaluation? *Educational Researcher*, *44*(2), 132-137. https://doi.org/10.3102/0013189X15575346

Dubnick, M.J. (2006). Orders of Accountability. Paper presented at the World Ethics Forum, Oxford, UK. Retrieved from http://pubpages.unh.edu/dubnick/papers/2006/oxford2006.pdf

Hood, C. (1995). The "New Public Management" in the 1980s: Variations on a Theme. *Accounting, Organizations and Society, 20*(2/3), 93-109. https://doi.org/10.1016/0361-3682(93)E0001-W

Hood, C. C., & Margetts, H. Z. (2007). *The Tools of Government in a Digital Age.* Basingstoke, Hampshire: Palgrave MacMillan. https://doi.org/10.1007/978-1-137-06154-6

Hopmann, S. T. (2007). Epilogue: No Child, No School, No State Left Behind: Comparative Research in the Age of Accountability. In Hopmann, S.T., G. Brinek, and M. Retzl (Eds). 2007. *PISA zufolge PISA – PISA According to PISA. Hält PISA, was es verspricht? – Does PISA Keep What It Promises?* (pp.363-415). Wien: LIT.

Kalimullah, N. A., Ashraf, K. M., & Ashaduzzaman, M. N. (2012). New public management: Emergence and principles. *Bup Journal, 1*(1), 1-22.

Leckie, G., & Goldstein, H. (2017). The evolution of school league tables in England 1992-2016: 'contextual value-added', 'expected progress' and 'progress 8'. *British Educational Research Journal.* http://dx.doi.org/10.1002/berj.3264

Lewis, S., & Hardy, I. (2017). Tracking the topological: The effects of standardized data upon teachers' practice. *British Journal of Educational Studies, 65*(2), 219-238. https://doi.org/10.1080/00071005.2016.1254157

Lingard, B., Martino, W., Rezai-Rashti, G., & Sellar, S. (2016). *Globalising educational accountabilities*. New York: Routledge.

Lingard, B., Sellar, S., & Lewis, S. (2017). Accountabilities in schools and school systems. In G. Noblit (Ed.), *Oxford Research Encyclopaedia of Education* (pp. 1-28). New York: Oxford University Press. http://dx.doi.org/10.1093/acrefore/9780190264093.013.74

Lukes, S. (2005) [1974]. *Power: A radical view.* 2nd ed. Basingstoke: Palgrave Macmillan. https://doi.org/10.1007/978-0-230-80257-5

Mourshed, M., Chijioke C., & Barber , M. (2010). How the World´s most improved school systems keep getting better. London: McKinsey & Company.

OECD. (2013). OECD. 2013. *Teaching and Learning International Survey TALIS 2013 Conceptual framework.* Retrieved from http://www.oecd.org/edu/school/TALIS%20Conceptual%20Framework_FINAL.pdf

OECD. (2014). *TALIS 2013 Results: An international perspective on teaching and learning.* https://doi.org/10.1787/9789264196261-en

Pawson, R. 2002b. "Evidence-based policy: the promise of 'realist synthesis'." *Evaluation,* 8(3), 340–358. https://doi.org/10.1177/135638902401462448

Peters, B. G. (2015). *Advanced Introduction to Public Policy.* Cheltenham: Edward Elgar.

Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, *4*(2), 155-169. https://doi.org/10.1007/BF01405730

Robertson, S. L. (2012). Placing teachers in global governance agendas. *Comparative Education Review, 56*(4), 584-607. https://doi.org/10.1086/667414

Rutkowski, D. (2008). "Towards an understanding of educational indicators." *Policy Futures in Education, 6*(4), 470-481. https://doi.org/10.2304/pfie.2008.6.4.470

Sahlberg, P. (2011). Fourth Way of Finland. *Journal of Educational Change, 12*(2), 173–185. https://doi.org/10.1007/s10833-011-9157-y

Sellar, S., & Lingard, B. (2013). The OECD and global governance in education. *Journal of Education Policy*, *28*(5), 710-725. https://doi.org/10.1080/02680939.2013.779791

Smith, W. C. (2014). The global transformation toward testing for accountability. *Education Policy Analysis Archives*, *22*(116). http://dx.doi.org/10.14507/epaa.v22.1571

Smith, W.C. (2016). An Introduction to the Global Testing Culture. In *The Global Testing Culture: shaping education policy, perceptions, and practice*, edited by W.C. Smith, 7-23. Oxford: Symposium.

Tatto, M. T. (2011). Reimagining the Education of Teachers: The Role of Comparative and International Research. *Comparative Education Review*, *55*(4), 495-516. https://doi.org/10.1086/661769

Thrift, N. (2005). *Knowing Capitalism*. London: Sage.

Verger, A., & Curran, M. (2014). New public management as a global education policy: its adoption and re-contextualization in a Southern European setting. *Critical Studies in Education*, *55*(3), 253-271. https://doi.org/10.1080/17508487.2014.913531

Verger, A., & Parcerisa, L. (2017). A difficult relationship. Accountability policies and teachers: International evidence and key premises for future research. In M. Akiba & G. LeTendre (Eds.). *International Handbook of Teacher Quality and Policy* (pp. 241-254). New York: Routledge.

UNESCO (2017). *Sustainable Development Goal 4: Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all.* Retrieved from https://sustainabledevelopment.un.org/sdg4

About the Authors/Guest Editors

Jessica Holloway
Research for Educational Impact (REDI) Centre, Deakin University
Jessica.holloway@deakin.edu.au
http://orcid.org/0000-0001-9267-3197
Jessica Holloway is a post-doctoral research fellow at the Centre in Research for Educational Impact (REDI) at Deakin University. She draws on post-structural theory to understand contemporary modes of accountability and its production of new teacher and leader subjectivities. Her current project, entitled *Teacher Leaders and Democracy: An International Study,* looks at modes of distributive leadership in U.S. and Australian schools.

Tore Bernt Sørensen
University of Bristol
t.b.sorensen@bristol.ac.uk
http://orcid.org/0000-0003-0982-5685
Tore Bernt Sørensen completed his doctorate at the Graduate School of Education, University of Bristol, UK, in 2017 with the dissertation "*Work In Progress: The Political Construction Of The OECD Programme Teaching And Learning International Survey".* Tore's research centers on comparative studies of education governance in a global context. Tore has a background as teacher and teacher trainer in Denmark. Before starting his doctorate, he worked in the Analysis and Studies Unit of the European Commission's Directorate-General for Education and Culture.

Antoni Verger
Universitat Autònoma de Barcelona
Antoni.verger@uab.cat
http://orcid.org/0000-0003-3255-7703
Antoni Verger is associate professor at the Department of Sociology of the UAB. A former post-doctoral fellow at the University of Amsterdam, Antoni's research analyses the relationship between global governance institutions and education policy, with a focus on the study of public-private partnerships and accountability policies in education. Currently, he is coordinating the research project REFORMED - *Reforming Schools Globally: A Multiscalar Analysis of Autonomy and Accountability Policies in the Education Sector* (ERC StG, 2016–2021).

Please send errata notes to Audrey Amrein-Beardsley at [Audrey.beardsley@asu.edu]()

Join EPAA's Facebook community at [https://www.facebook.com/EPAAAAPE]() and **Twitter feed** @epaa_aape.

arquivos analíticos de políticas educativas
conselho editorial