

Exploring the Changing Face of School Inspections

Martin BROWN*
Gerry MCNAMARA**
Joe O'HARA***
Shivaun O'BRIEN****

Suggested Citation:

Brown, M., McNamara, G., O'Hara, J. and O'Brien, S. (2016). Exploring the changing face of School Inspections *Eurasian Journal of Educational Research*, 66 1-26
<http://dx.doi.org/10.14689/ejer.2016.66.1>

Abstract

Background: A number of countries have had school inspection for many years. The origins of these systems date back to the nineteenth century when mass public schooling was introduced, and education and other emerging public services were required to comply with centrally mandated rules and programmes. In contrast, many countries across the world have only introduced school inspection over recent decades as the perceived importance of educational quality as a driver of economic competitiveness has become influential in state policy. International bodies such as the OECD and, in particular, comparative evaluations of education systems such as PISA have led to a constant stream of interventions and reforms designed to deliver higher student performance outcomes. These factors have driven the growth of inspection.

Purpose of Study: The purpose of this study is to provide an analysis of the factors that have led to the rapid rise of inspection as a school governance mechanism. It goes on to examine how developing conceptualisations of the ways in which inspection can be employed to achieve the range of outcomes with which it is tasked are leading to an evolving toolkit of inspection approaches and models. A number of these are examined in detail with a view not only to description but in terms of whether some of the demands that they place on schools are, in fact, realistic in practice.

* Corresponding author: Dr., Martin Brown, School of Policy and Practice, Institute of Education, Dublin City University. e-mail: martin.brown@dcu.ie

** Prof. Dr., Gerry McNamara, School of Policy and Practice, Institute of Education, Dublin City University. e-mail: martin.brown@dcu.ie

*** Prof. Dr., Joe O'Hara, School of Policy and Practice, Institute of Education, Dublin City University. e-mail: joe.ohara@dcu.ie

**** Dr., Shivaun O'Brien, School of Policy and Practice, Institute of Education, Dublin City University. e-mail: shivaun.obrien@dcu.ie

Sources of Evidence: This study used document analysis of policy documents and existing research to deconstruct factors relating to the changing face of school inspection since the late 1990's.

Main Argument: Formal processes of school inspection have become virtually universal. It is also argued that inspection, as it is now widely understood and practiced, has moved quite far from its historical roots and purposes. Inspection is now a complex component of wider modern concepts of public sector management and governance including quality, improvement, accountability, transparency and cost effectiveness.

Conclusions: Historically inspection was largely about compliance with rules and to an extent to judge the work of individual teachers. Now, at least, in theory, it is as much concerned with creating a regulatory framework within which schools as organisations can enjoy greater autonomy while simultaneously being held responsible for student performance outcomes.

Keywords: school inspection, accountability, school self-evaluation, governance, school development

Introduction and background

A number of countries, particularly England, Ireland and the Netherlands, have had school inspection for many years. The origins of these systems date back to the nineteenth century when mass public schooling was introduced, and education and other emerging public services were required to comply with centrally mandated rules and programmes. In contrast, many countries across the world have only introduced school inspection over recent decades as the perceived importance of educational quality as a driver of economic competitiveness has become influential in state policy. International bodies such as the OECD and, in particular, comparative evaluations of education systems such as PISA have led to a constant stream of interventions and reforms designed to deliver higher student performance outcomes. These factors have driven the growth of inspection.

Somewhat paradoxically, while many of these reforms have sought to decentralise autonomy and responsibility to local bodies or individual schools and teachers, this policy direction has also led to a perceived need for greater accountability to counterbalance the increased autonomy given to schools. As a result, formal processes of school inspection have become virtually universal, a remarkable development in a short space of time. However, it is argued in this paper that inspection, as it is now widely understood and practiced, has moved quite far from its historical roots and purposes. Inspection is now a complex component of wider modern concepts of public sector management and governance including quality, improvement, accountability, transparency and cost effectiveness. In fact, it is suggested here that school inspection is an interesting example of the ways in which changing political theories of governance impact on the management and delivery of public services. Historically inspection was largely about compliance with rules and to an extent to judge the work of individual teachers. Now, at least, in theory, it is as much concerned with creating

a regulatory framework within which schools as organisations can enjoy greater autonomy while simultaneously being held responsible for student performance outcomes.

At first it seemed that new theories of public sector governance such as 'new public management' (NPM) were primarily driven by lack of trust in autonomous professionals and the desire to impose accountability through inspection, sanctions and rewards. However, since NPM theory is also concerned with reducing state bureaucracy, decentralisation and privatisation of services, consumer and 'stakeholder' voice and choice, and improvement by both regulation and competition; external monitoring systems such as inspection are serving several ends. Accountability remains central, but the improvement of organisational performance to the extent that the organisation can be trusted to be more autonomous is also a key goal. In this conception, inspection becomes a lever or mechanism not only to judge schools but to improve them to a satisfactory level and then leave in place processes through which they can monitor and maintain the required standards. Thus, current conceptualisations of inspection suggest that Inspection models need to be adapted as education accountability systems mature. Schools and their stakeholders develop evaluation literacy and innovation capacity to improve education on their own and thus have less need of being driven by top-down inspections and reform initiatives (see, for example, Barber, 2004). In consequence school inspection theory and practice is in a constant state of flux.

Reflecting this wider concept of inspection, the paper will explore how inspection models and types have changed over recent years (e.g. from full to proportional inspections, from checking compliance to evaluating the educational practices and output of schools and from external monitoring to increased emphasis on school self-evaluation to enable schools to respond to, or support changes in national education systems and educational reforms). The paper begins by placing school inspection in the context of 'reform' of public sector governance. It examines some of the factors that have led to the rapid rise of inspection as a school governance mechanism. It goes on to suggest that developing conceptualisations of the ways in which inspection can be employed to achieve the range of outcomes with which it is tasked are leading to an evolving toolkit of inspection approaches and models. A number of these are examined in detail with a view not only to description but in terms of whether some of the demands that they place on schools are, in fact, realistic in practice.

Method

Research Design

Document analysis was used as the research method for this study. According to Bowen (2009), document analysis can be utilized as a stand-alone method and can serve varying purposes such as providing background information and historical understandings. It can also be used for the purpose of 'tracking change and development, and verification of findings from other sources' (30). Atkinson and Coffey (2009) also states that 'we have to approach documents for what they are and what they are used to accomplish' (79). In this study document analysis was used to deconstruct factors relating to the changing face of school inspection.

Research Sample

This study included an analysis of 621 documents on school inspection policy and practice in Europe and elsewhere. To ensure authenticity and reliability of data and to form overarching themes for the analysis; inclusion of documents was initially limited to peer-reviewed literature on school inspection that has been published since the 1990's. However, to triangulate results and to form an overall interpretation of the study, the inclusion of documents was also extended to official government policies and publications produced by international organizations.

Research Instrument and Procedure

Documents were initially selected based on the author's prior knowledge of the field. Following on from this, key terms were identified for the search strategy. Using these terms, databases were searched for relevant literature. The initial search resulted in the collation of approximately 800 articles. Based on the inclusion criteria, all articles were read to determine if they would be included in the analysis of which, approximately 179 articles were excluded from the final analysis.

Data Analysis

Thematic analysis was used to identify patterns (Bowen, 2009) emerging in the literature. Within this, all documents were coded using a data extraction form. This involved the production of a summary outline detailing the purpose, method, conclusion and key themes emerging for each document contained in the sample. This process of analysis allowed the researchers to form an overall interpretation of the study.

Inspection and the discourse of public sector reform

The roots of all inspection systems, whether in health care, education or other public services, are primarily consistent with Bentham's assertion that 'the more strictly we are watched, the better we behave' (as cited in Etzioni, 2010: 393). As Donaldson (2013: 9), in reference to Sparrow (2008), affirms, 'inspection concerns itself, either implicitly or explicitly, with requiring deliverers of services or citizens themselves to conform to certain expectations'. The importance of such inspection systems also has been expressed as 'a greater commitment to values' (Mackiney and Howard 1998: 471) and 'the legal obligation to respect the legitimate interests of others affected by decisions, programs, and interventions' (Considine 2002: 21).

A number of researchers (see, for example, McNamara & O'Hara 2008a; Ball 2012; Beckmann, Cooper, & Hill 2009) also share the view that certain formerly private-sector concepts have gained prominence in the discourse about the management of public services in recent decades. These include accountability, benchmarking, deregulation, decentralisation, value for money, quality assurance, risk-based analysis, targets and so forth. The lexicon of school inspection policy and practice certainly illustrates this trend. Indeed, Lindberg's (2013) stylised timeline of accountability serves as a useful means of summarising the new public management reform initiatives that can be found in most public services.

P must first transfer decision-making power over a particular D to A. Then A acts in this capacity and P can thereafter require A to

provide information and justification for these actions; and if A fails to do so, P can apply sanctions. (Lindberg 2013: 212)

In terms of decision-making power being transferred from government P to organisation A (decentralisation), Ball (2003) draws this conclusion:

Thus, the reforms are presented as giving 'managers and organizations greater freedom in operational decisions and remove unnecessary constraints in financial and human resource management' (OECD 1995: 29). However, it is misguided to see these reform processes as simply a strategy of deregulation; rather, they are processes of re-regulation. It is not a matter of the state abandoning its controls. Instead, it is the establishment of a new form of control, what Du Gay (1996) calls 'controlled de-control'; it is, in fact, a new kind of state. (Ball 2003: 217)

The government, or the agencies working on behalf of the government (inspectorate), then require the organisation to provide, in accordance with externally devised self-evaluation frameworks, information about the quality of services it provides and the justification for its actions. If the organisation fails to do so, the government has the right to sanction or remedy these actions in order to maintain and improve the quality of services provided (sanctions). This process, in short, is accountability, decentralisation and 'new inspection'.

Indeed, it is evident that since the turn of the last century, the development of school inspection theory reflects the same patterns of internationally mandated public sector compliance. This is true in most countries and regions and also in other publicly-funded bodies, such as the health care sector and prison services. Shewbridge et al (2014) affirm the following: 'A key recommendation coming out of the OECD Reviews of Evaluation and Assessment in Education is to situate school system evaluation in the broader context of public sector performance requirements' (164).

Furthermore, the ever-growing influence of organisations such as the World Bank, UNESCO and the OECD has also had a profound impact on school evaluation and inspection, 'As "lower" levels of education policy-making are being overlaid by a "higher level" of international policy-making, causing profound changes in education policy introduced in any locality' (Thompson and Cook 2014: 702). This change has resulted in most established (England and the Netherlands) and newly-formed inspectorates (Chile and Germany) aligning with a much broader global public policy reform agenda that captures new public management theories and international organisations' approval for normative international public sector educational performance. Shewbridge et al (2014: 164) point out the following:

When Ministries and other bodies with specific responsibilities for system evaluation need to show accountability for their performance, this stimulates demand for procedures to monitor progress in the school system and, where necessary, to establish adequate systems to collect evidence on progress. This is particularly the case in systems where high-level targets are set by the government related to productivity, competition or general economic and social improvement.

Similarly, Santiago (2013a: 35) is also of the view that:

Evaluation and assessment have become increasingly important as a result of greater levels of school devolution; a stronger role for market-type mechanisms in education; the emergence of new public management; the growing imperative of an efficient use of public resources; the need to focus on 'quality for all' and the rising importance of education in a global world'.

Specifically, with respect to the influence of public sector reform on present-day school inspection, Karsten (1999: 309) asserts that NPM philosophies eventually became embedded in the Dutch education sector during this period:

The central idea was the desire to reduce government bureaucracy, increasing autonomy for schools to regulate themselves. In the backwash of these plans for administrative changes it was often argued that the parents should be more involved, particularly with the choice of school and that there should be more competition.

Indeed, Van Bruggen (2010a: 22) made the following statement in the Standing International Conference of Inspectorates (SICI) profiles of inspectorates in Europe:

As has been said: "education is too important to leave it to teachers only..." The same accountability is asked from other public services like hospitals, libraries, local and national administrations, etc. In many countries, public evaluation of these public services has become commonplace in the last two decades.

However, although new public management can be seen as the catalyst for a significant change in school inspection policy and practice, it also should be recognised that there are other factors at play in driving change in this area at national and regional level. These factors include the length of time that each inspectorate has existed; the political agenda of governments; the strategic influence of organisations, such as the OECD and SICI; and the local evaluative infrastructures and support (for example, school self-evaluation capacity, data warehouses, etc.). On the other hand, a very different perspective on the rise and function of inspection is offered by Donaldson (2013: 8), who argues that:

inspection is often associated in the public mind with a rather narrow set of activities which involve notions of compliance and audit. In fact, it is a very plastic concept which takes and has taken many forms and which can serve many different purposes.

The author goes on to provide a list of potential contributions that school inspection can make to education policy and practice: 'enforcer, assurer, mitigator of risk, catalyst, knowledge broker, capacity builder, partnership builder, agenda setter, and preserver/creator of the space for innovation' (Donaldson 2013: 8).

The variety of roles and purposes which inspection can fulfil means that the particular emphasis given to it depends enormously on contexts both political and educational. In practice for example, Dederling and Müller (2011) comparing school inspection in the federal states of Germany, with the United Kingdom and the Netherlands affirm that in all of the federal states of Germany; 'quality development' is a priority of school inspection. In comparison, the authors go on to state that school

inspection in the United Kingdom and the Netherlands focuses on the public accountability of schools and teachers through various mechanisms, such as the publication of school inspection reports. 'The school inspection thus leverages pressure onto a school to improve its quality in order to acquire clients, thus triggering the schools to enter into competition' (Dedering & Müller 2011: 307). In the case of the federal states of Germany, however, the publication of inspection reports is left to the discretion of each school: 'Hence, there is no increased competition among schools. Rather, the school inspection provides the schools with information they can use to optimize their governance processes which they previously lacked' (ibid).

Indeed, the varying purposes of inspection policy and practice found in countries like the Netherlands and Germany parallels Landwehrs (2011) study on the potential effects of school inspection where school inspection may not only be used 'for the purpose of traditional school accountability but rather, within a complex network of educational discourse. The potential effects of school inspection may also be for the purpose of 'gaining knowledge', 'school development' and 'enforcing standards' (as cited in Gaertner et al 2013: 491).

In summary, many scholars in the field see recent developments as constituting a fundamental paradigm shift in school inspection to a conception that has now become the conventional mode of thinking about inspection in most countries or regions. By way of explanation, Van Bruggen (2010b), refers to inspectorates in Europe: 'In the last 20 years or so, the roles and tasks of inspectorates of education have changed radically in most countries where inspection was already in existence since the early decades of the 19th century' (87). Indeed, Thomas (1998) refers to this period as 'the genesis of the New School inspection system', stating that it saw a significant change in school inspection policy and practice in England and elsewhere (see, for example, MacBeath 2006; McNamara and O'Hara 2008b; Hall and Noyes 2009). As Martin (2008) in reference to the evolution of school inspection in England puts it: 'while the HMI of the 1950s and 1960s had a role to promote and disseminate good practice and inform government about the quality of education, the role of Ofsted was to evaluate and challenge schools to improve and make public their judgements' (54). In the case of England, 'the key point of the new system was to be that all schools would be inspected within a short time scale, and then at regular intervals thereafter' (Thomas 1998: 42). Under the influence of 'New School inspection', almost all inspectorates, whether in Europe or elsewhere, have moved towards a model of what Van Bruggen (2010b) refers to as 'full', 'whole' or 'complete' modes of inspection.

New tools of inspection for a changing landscape

In this changed conception of inspection, the purpose of school inspection is not merely that of 'inspection for compliance'. Rather, inspection is conceptualised as a catalyst for continuous school improvement through a process of externally-regulated school self-governance. As the then Minister of State for school standards in England put it, 'when it comes to external evaluation, the key is to make the process of inspection as useful to schools as possible, supporting self-improvement where it is present, spurring it where it is not' (Miliband, 2004: 15). According to van Bruggen (2010b), 'the rationale behind the rapid growth of full inspections as an important instrument in national educational policy is that these conditions enable the inspections to work with two rather different educational governance philosophies'

(95). New school inspection is not merely about guaranteeing a minimum set of standards; it also hinges on the view that there is a continuous need for improvement regardless of the level of quality attained. Van Bruggen (2010b: 96) states:

The second educational governance philosophy is enacted where the government wants the “best possible” schools. This is the improvement push that inspections and all related work must give; not only for the schools with quality problems, but for all schools.

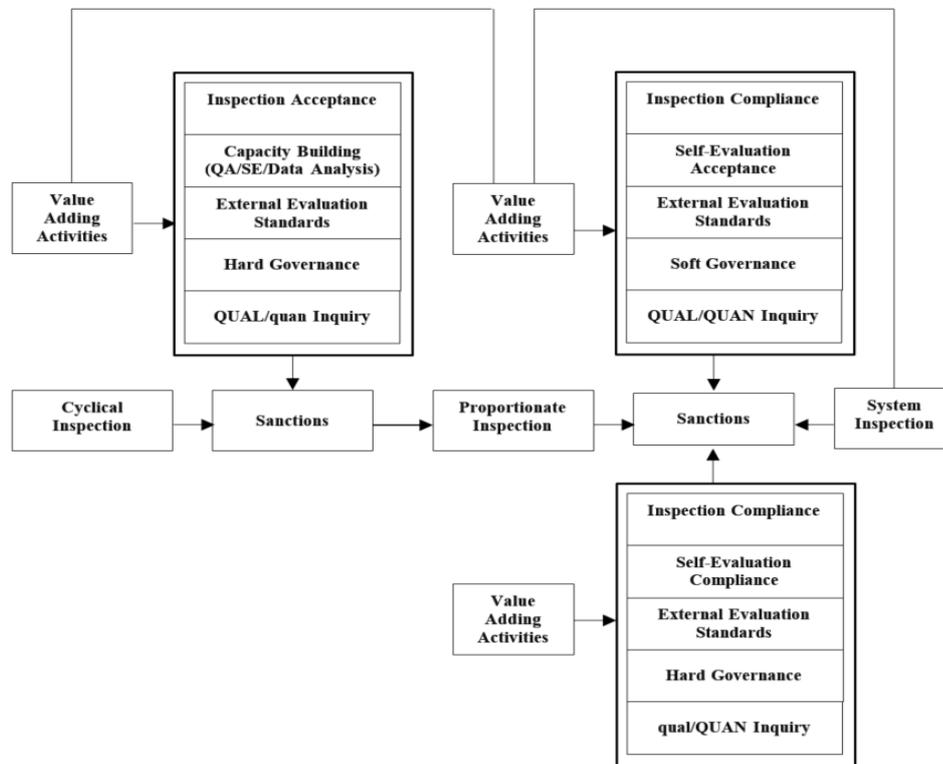
The introduction of regular whole school inspections has resulted in the initiation of other, related, evaluation activities: school self-evaluation; the increasing use of statistical conjectures of quality to inform inspection judgements; and, more recently, a move from cyclical to proportionate, risk-based inspection. There has also been an increasing production of system-wide thematic inspection reports in countries like Ireland.

As a result of what has now become the continuous improvement function of inspection, schools in most countries are inspected at regular intervals. Schools are evaluated through three inspection domains (cyclical, proportionate and collective system inspections) that consist of a series of overlapping evaluation activities, irrespective of the length of the various inspectorates' existence (Figure 1).

In addition, in this century there is also a discernible trend towards the use of shorter, risk-based inspections in many countries for a variety of reasons that include: the increasing use of quantitative data; the regulatory requirement for schools to carry out their own school evaluations in accordance with an externally mandated framework of quality indicators (e.g. England and Ireland); and the financial burden associated with school inspections.

We have argued to this point that inspection has become part of the educational landscape in very many countries either as a reformed version of previous iterations or in many cases a new feature of educational governance. We have also suggested that while inspection may certainly be seen as a tool for achieving compliance, ensuring accountability and assuring quality it is also perceived in most cases as having a developmental and improvement function and a role in equipping schools to analyse performance and drive change. This conceptualisation and the array of roles that flow from it clearly requires a varied set of approaches and tools to be employed by inspectorates and we turn now to some interesting developments in this regard.

Figure 1.
New school inspection domains. Source: The authors



From external monitoring to school self-evaluation

In the 1990's, there was a strong political agenda for decentralisation leading to the proliferation of regular whole school inspections. In the case of England, 'after a decade of antagonism between agencies of governments and schools' (Macbeath, 2006: 1), a new model of school evaluation, self-evaluation, began to take prominence. As Nevo (2010) states, 'parallel to the almost universal phenomenon of external school evaluation, many countries have more recently tended to apply newly developed evaluation methods at the school level in the form of internal evaluation or self-evaluation' (781). Further, MacNamara & Nayir (2014), in reference to the self-evaluation research of MacBeath (1999; 2008; 2009) affirm: '... even he could hardly have foreseen the current popularity of self-evaluation as a key element of the inspection regime in so many countries' (51). Indeed, with the appeal of less centralised control and the prospect of increased cooperation between schools and the inspectorate, a dual system of internal/external quality assurance began to emerge and is now being used in most OECD countries. For example, in their discussion of the emergence of school self-evaluation in England, Hall and Noyes (2009) state, 'in the late 1990s, in response to widespread distress and frustration amongst teachers, the National Union of Teachers (NUT) commissioned work on school self-evaluation' (850).

Self-evaluation also was appealing in that evaluation would no longer consist of command and control accountability mechanisms via the Inspectorate. The foundation of the new relationship between state and school was to be based on the concept of 'intelligent accountability' and was perceived by some (Nevo 2002) as being able to counteract the inherent limitations present in both systems of evaluation if used in isolation. For school self-evaluation, Huber (2011) makes the following statement in reference to school governance in Switzerland: 'through this process, so-called 'blind spots' may be uncovered that a school was unable to identify by means of an internal evaluation' (473). For school inspection:

The participants in school self-evaluations are usually more familiar with the specific nature of the local school context and communicate better with the (local) school community. By focusing attention on additional data which reveal the unique character of particular schools, they can broaden the focus of external evaluations. (van Hoof & van Petegen 2007: 106)

In other words, school self-evaluation and school inspection could be mutually beneficial to serving the school accountability agenda on the one hand and school improvement agenda on the other (see, for example, Nevo 2002; 2010). It also must be noted, however, that the rapid promotion of school self-evaluation also occurred at a time when more responsibility (in the form of decentralisation) was deemed necessary to be transferred to the school and away from the state. This is based on the view that: 'decentralization with appropriate accountability systems and teacher quality are essential elements of any strategy to improve quality of school education' (World Bank 2007: ii). As Gordon and Whitty (1997) put it, 'the single largest change in schooling under the neoliberal project has been the push towards the atomisation of the control of schools' (456).

However, while appeals for school self-evaluation grew, and governments rigorously promoted the concept of decentralised school autonomy, extensive debate surrounded schools being allowed to take more responsibility for their own evaluations. Indeed, to this day, there remains much concern about finding the correct balance between school self-evaluation and inspection (see, for sample, MacBeath 2006; Van Amelsvoort & Janssens 2008; MacBeath 2010; Nevo 2010; Brown 2013; Simons 2013). With the new relationship between state and school, a dichotomy of concerns emerged among inspectorates regarding the effective deployment of self-evaluation. Concerns among inspectorates relating to how best to combine school inspection and self-evaluation centred on the importance of maintaining the accountability function of inspection while simultaneously allowing schools to form their own evaluative judgments regarding the quality of education provided. The varying views on how best to form an amalgam between school self-evaluation and inspection, tended to reverberate within Alvik's (1996) typology of three distinct models of decentralised evaluation: 'parallel' ('often disconnected' (Durrant & Holden 2005: 110)), 'sequential' ('in which external bodies follow on from a school's own evaluation and use that as the focus of their quality assurance system' (Cambridge & Carthew 2007: 289)) and 'co-operative' evaluation ('The two parties discuss and negotiate the process and different interests and viewpoints are taken into account simultaneously' (MacBeath 2005: 85)).

Some countries such as New Zealand tend to operate within a collaborative model of school evaluation. As Nusche et al (2012) point out, 'New Zealand's approach is

collaborative in the sense that both parties attempt to work together to agree on a rounded picture of the school in which there is mutual recognition of its strengths and consensus on areas for development' (95). However, in the case of European inspectorates, Ferrer (2010) states that the 'modalities of such combinations did vary from one country to another, but in general terms there was an explicit interest in sequential models' (128). MacBeath, in reference to inspectorates in Europe, affirms:

In Europe, the Standing International Conference on Inspection (SICI) has for a number of years argued for a 'sequential' approach (Alvik, 1996) in which the school's own self-evaluation provides the focus and centrepiece for external review and in which initiative lies with school leaders to place self-evaluation at the heart of school and classroom practice. (2008a: 386)

As the Chief Inspector of Ireland put it, 'Our ultimate goal is for schools to conduct their own evaluations transparently and accurately and for inspectors to visit these schools to evaluate the school's own self-evaluation' (Hislop 2012). However, cognisant of the view that 'the languages interweave, like shoals of fish, their boundaries indistinct and their participants seemingly inhabitants of more than one grouping' (Barnett 1994: 166); the foundations for the new relationship between state and school also paralleled the development of inspectorate-devised self-evaluation guidelines and report templates that would subsequently be used by schools and evaluated by inspectors. The logic of this mode of evaluation can be described by means of Macbeath's use of Perkins' (2003) metaphor of 'taming the wild' and 'wilding the tame'. On the one hand, there is a need for schools to demonstrate elements of externally created evaluative best practices where 'the wild is tamed by clear targets, predetermined outcomes and focusing of teachers' attention on templates of good practice' (MacBeath 2012: 131). On the other hand, however, "wilding of the tame" suggests recognition of domestication and a conscious attempt to loosen the ties that bind teachers to mandated practice. This is explicitly stated in relation to the process of learning to jump through the hoops before 'going wild' (ibid). In practice, an inspector participant in Brown (2013) when referring to inspectorate devised guidelines to support school self-evaluation in Northern Ireland provides logic for this mode of evaluation.

We provide Together Towards Improvement, which is a framework of quality indicators and questions and descriptors, and many schools wish to use it, but my view would be, and I think that would be the inspector's views, it's not compulsory to use it...What we expect is that schools are engaging in self-evaluation, and if they have developed other tools, then we're very interested to know what those tools are. We want to know how effective they are in giving insight. We'll tell other schools about them...unashamedly...but we would always say to that, with Together Towards Improvement, if you want to add other criteria and questions to it, to reflect the reality of your school because every school is different, then you should do that, if you want to modify it, as long as you're not taking away from it. (170)

Indeed, when referring to schools being explicitly driven by external mandates, MacBeath (2008b) earlier cautioned that 'schools that play safe, driven by external mandates set tight parameters around what can be said and what can be heard. Such

schools are antithetical to the notion of a learning organization which, by definition, is always challenging its own premises and ways of being' (145). This perspective resonates with McNamara & Nayir (2014) who add, 'One thing all the literature in the field stresses, however, is that school self-evaluation will not just happen because it is mandated in policy documents' (53). This may also be the case when policy documents and instruments are transposed from one country to the next. As an inspector participant in Grek (2014: 55) puts it:

I've just been to Mexico. I was part of an OECD group looking at the evaluation of the system in Mexico. What they did was to take materials from Scotland and translated them into Spanish and suggested that all schools do that. What happened? Nothing really. Any system has to be supported not just by printed material but face to face discussion and good examples.

Nonetheless, almost all countries in the world where school inspection exists now have in place a set of inspectorate devised school self-evaluation instruments that schools are required or strongly recommended to use. In the case of Malaysia, for example, schools are required to use an evaluation instrument called 'Standard Quality of Education Malaysia' (SQEM). According to Hamzah & Tahir (2013: 55):

The evaluation has to be based from an abundance of data, for example, the students' academic results in both schools and public examinations, students' performance in co-curricular activities and also the quality of teaching and learning. All these must be documented for the verification of the Schools Inspectorates.

Comparing the evaluation of Chinese calligraphy to that of school appraisal in Shanghai, Tan (2013) highlights what has now become common evaluation practice in almost every country where school inspection exists: 'Like Chinese calligraphy, school appraisal involves the schools showcasing their subjective characteristics within the rules set by the Shanghai authorities' (99). It is apparent that the logic of this framework is the avoidance of unregulated governance by schools. Donaldson (2013: 11) provides a plausible explanation for this value-adding activity:

The powerful relationship between external and internal evaluation is central to stimulating improvement. Each can make a particular contribution, but the synergies arising from the combination of the two can bring particular benefits. Inspectorates are increasingly emphasising the importance of effective self-evaluation as a driver of improvement. But self-evaluation can become self-delusion or worse and must operate within a framework of accountability which both encourages its rigour and validates its authenticity.

In consequence, from the initial act of full inspection, through a series of elaborate value adding activities such as the decentralisation of externally mandated mechanisms for self-evaluation, almost all established inspectorates (England, Ireland and the Netherlands); what a Swedish inspector participant in Grek et al refers to as 'First-generation inspectorates' (2013: 498) tended to be positioned within a cyclical model of evaluation where schools were inspected over a set period using various mechanisms to ascertain the quality of education provided. This format also has been adopted by newly-established inspectorates. However, the value placed on repeated

whole school cyclical inspections has been questioned, as pointed out by Van Bruggen (2010b: 109):

But a repeated inspection of the same 'superficial' character – and certainly if this comes only after five or six years – does not bring much new knowledge. This is, of course, for many inspectorates a budgetary problem: not having sufficient staff and money. But this is not the only cause: many influential bodies (political parties, teacher union executives and universities) find it dangerous to build too powerful an inspectorate that has too large an influence on schools, and one that is going to dictate what 'the best quality' is in terms of quality statements and indicators.

Accordingly, in this century, there also has been movement away from cyclical inspections and a move towards shorter, more focused, risk-based and system-wide thematic inspections in some countries (in particular for those countries that have completed a full set of system-wide inspections). This change has occurred in parallel to inspectorates' increasing use of quantitative data to form value judgements regarding the quality of education provided in schools; in consequence, it also is used as a significant determinant to ascertain the frequency of school inspections. The following section will provide specific examples from various European inspectorates and analyse the logic propelling this upward trend.

From cyclical to risk-based inspections

In this and the latter parts of the last century, established inspectorates, such as England, Ireland, and the Netherlands, tended to be inspected through a process of regular whole school inspections. This trend occurred in tandem with the development of externally devised school self-evaluation guidelines where schools in almost all countries where inspection exists (e.g. Korea, Malaysia, and Singapore) were obliged or strongly advised to operate within an externally mandated self-evaluation framework of quality indicators. Within this framework, school self-evaluation reports, normally consisting of a 3-year cycle for improvement, are validated either through cyclical (for example, the Czech Republic and the Styrian Inspectorate) or proportional/risk-based inspections (for example, Ireland, Northern Ireland, Flemish Community of Belgium, the Netherlands, New Zealand and Scotland). In many ways, the case of Shanghai's school evaluation policy and practice sums up what is common inspection practice in most countries:

Under the appraisal system, every school is required to formulate a 3-year development plan that comes with a yearly implementation plan...The Shanghai Municipal People's Government Educational Supervisory office will then conduct an on-site inspection, and the supervisory experts will cast votes on whether the plan passes inspection. (Tan 2013: 101)

As previously stated, there recently has been an upward trend towards proportionate or risk-based inspections, the frequency of which is based on inspection value judgements relating to deviant cases identified as being outside of acceptable school practice. These value judgements include, but are not limited to, an examination of school self-evaluation reports and school development plans; on-site observations of the quality of teaching and learning; and the change capacity of schools coupled with various desk checking exercises, such as an analysis of externally devised

examination results. Other countries, such as the Netherlands and Sweden, also use value judgements, such as the media, in order to ascertain the component of risk associated with the quality of education provided. Plausible explanations as to the varying use of cyclical and proportionate-based models among countries include the collective school self-evaluation capacity of each country, local evaluative infrastructures and support (Self-evaluation guidelines, data warehouses, etc.), and also the length of time that each inspectorate has been in existence. For example, 'MacBeath et al. assert that it is England and The Netherlands that have the most developed external evaluation systems' (Whitby 2010: 10). In consequence, inspectorates of this type (Australia, England, and Northern Ireland) have the required infrastructure to initiate risk-based and system-wide inspections.

For those countries that have not undergone a full round of country-wide inspections and/or have limited access to system-wide school performance assessment data, inspectorates tend to operate within the domain of repetitive cyclical inspections. In the case of the Czech Republic, for example, Strakovaa & Simonová (2013: 477) state:

Assessment and evaluation in the Czech education system have been repeatedly highlighted as a weak point of the country's education policy... In 1996, an OECD review team made specific recommendations related to the development of student assessment at the end of compulsory and upper secondary education (OECD, 1996); however, the system of assessment and evaluation has not yet been established.

The move towards risk-based inspection also can be seen in more recently established inspectorates, such as Sweden. Initially, with the division of the National Agency for Education (NAE) into two distinct agencies in 2003, all schools were to be inspected over a six-year period. However, with the rebranding of the Swedish inspectorate in 2008, there also has been a move towards proportionate inspections: 'Overall, the development has been regarded as implying a 'harder 'inspection, focusing only on deviances' (Ronnberg et al 2013: 180). Indeed, the upward trend towards risk-based inspections appears to be commonplace in almost all continents. In the case of Australia, for example, Santiago et al (2011) affirm that 'there does appear to be a move towards a risk-based determination, using available data to allocate schools to categories of risk which in turn determine the frequency, depth and degree of externality of reviews' (109). For those systems that have gone through a process of cyclical school inspections, there also has been a tendency to move towards targeted risk-based inspections for a variety of reasons. Worryingly, full inspections were abandoned altogether in the case of Estonia during the initial phase of whole school inspection in this country. Van Bruggen (2010b: 110) states:

Estonia's inspectorate is very clear about this political choice: after a couple of years of full inspections, it could not be demonstrated that all schools were better, and so the full inspections were abolished and only in high-risk cases is there a general inspection (conducted by a small national inspectorate), apart from the thematic inspections.

In many ways, the case of Estonia's move away from cyclical inspections is similar to that of other countries such as Ireland, where initially, all schools were inspected over a set period of time that normally consisted of a week's duration. However, this

mode of evaluation was abandoned in 2011 and replaced with risk-based inspections using various modes of judgement to target deviant cases falling outside the realm of acceptable quality indicators. The Chief Inspector of Ireland put it this way:

We have abandoned the traditional cyclical approach to planning the inspection programme and instead, we use a range of criteria to decide where inspections should take place. These criteria include information from previous inspections, from State Examinations and from the Schools Division of the Department, in addition to consideration of the length of time since the previous inspection. (Hislop 2012)

Many established inspectorates have followed suit after going through the process of whole school inspection, including England and Northern Ireland.

Justifying the movement towards proportionate inspections seems to be common for governments who are of the view that inspection resources could be better used elsewhere, such as carrying out more system-wide thematic inspections. Also, due to the change capacity of some schools, there is less of a need to inspect schools that are able to carry out their own evaluations compared to those that are not. Ehren et al (2013:15) provide further reasoning on the trend towards risk-based inspections in Europe:

Differentiated school inspections are generally implemented to increase the efficiency of school inspections by targeting inspection resources to potentially weak schools. Also, the increasing scale and professionalism of some schools allow for differentiated inspections when Inspectorates of Education can use results of self-evaluations to target inspection visits and potential areas for improvement. Both types of school inspections are, however, also often used as parallel methods by the same Inspectorate of Education.

There are more obvious reasons for the trend towards risk-based inspections, however. Risk based inspections are also aligned with inspectorate reforms in other publicly funded services. For example, Davis & Martin (2008a:16) make the following observation relating to the English governments concerns relating to the cost of inspecting public services as well as, 'the administrative burdens' placed on service providers.

The government at first turned a 'blind eye' to the mounting concerns about the cost of inspection and the administrative burdens that it placed on schools, NHS trusts, police services, councils and other local service providers. But from around 2004 onwards inspectorates came under pressure to reduce their running costs and co-ordinate their hitherto largely separate activities. They responded with promises of more 'proportionate' or 'risk-based' approaches...

In the case of school inspection, there is a need to reduce the high costs attached to school inspection practice. For example, Ehren et al (2011) studying the impact of school inspection in eight European countries (Austria, Czech Republic, England, Ireland, Netherlands, Norway and Sweden) state that 'annually an average amount of 68.559.103 euros is spent on inspecting schools in every country in this project'. In practice, Van Bruggens (2010b) avowal relating to issues associated with repeated

cyclical inspections, coupled with the financial burden to the exchequer, resonates with Simons (2013) observation on the promotion of Self-Evaluation in England and Ireland.

It is perhaps worth noting in relation to this initiative and that of Ofsted in England, that the promotion of school self-evaluation systems in both countries may have something to do with the economic downturn and the realization that an external inspection system of a week-long duration of all schools in the country, even if only once every four to five years, is too costly to sustain. (Simons 2013: 10).

However, concerns relating to the move from cyclical to targeted risk-based inspections also have been expressed. For example, Perry (2013) in reference to the move from cyclical to risk-based inspections in Northern Ireland, highlights a number of issues raised by the Teaching Council of Northern Ireland: 'The General Teaching Council for Northern Ireland (GTCNI) has recently highlighted concerns around this risk-based approach, suggesting that this may have a "potentially in-built socio-economic bias", and noting an "excessive reliance" on quantitative data and examination outcomes' (11). Indeed, initial concerns seem to be justified when it comes to the potential for more intense inspections to be carried out without taking into account the overwhelming effect of social deprivation on student learning outcomes. According to Leithwood et al:

Best estimates suggest that everything schools do within their walls accounts for 20 percent of the variation in students' achievement is based on what happens in schools (Creemer and Reetzigt, 1996) - the maximum difference a school can make because external factors are so powerfully stacked for some schools against others. (2010: 249)

In many ways, GTCNI concerns relating to the move towards risk-based data driven inspections without taking into account the context in which schools operate resonates with Perryman's (2006) observations of school inspections in England and Wales: 'The inspection system in England and Wales seems to pathologize the 'sick' school, without taking into account the society in which it is positioned' (151).

More recently, in order to rectify anomalies that could unintentionally target schools that have no control of contextual factors that inhibit learning outcomes, there has been considerable government interest in the use of value-added statistical conjectures of quality. Such measures could level the playing field between socio-economic school types by quantifying the varying contexts in which schools operate.

None the less, the interchangeable use of value-added measures can be seen in many countries inspection frameworks, such as England, Canada and Ireland. For example, in Canada, the Literacy and Numeracy Secretariat of the Ontario Ministry of Education developed a benchmarking module, called 'Schools Like Ours'. Its purpose is to find 'similar schools to any selected school using any combination of the school demographics, program information, and performance indicators by province, region, or board.' (Department of Education Ontario 2007: 4). Indeed, large scale organisations often use extensive single method quantitative studies to influence decision making in a country (see, for example, PISA 2009; OECD 2012). For instance, Santiago, in reference to the Danish Education system, states:

The first results of PISA published in December 2001 jolted Denmark and its education community. They provided evidence that one of the most expensive education systems in the world was performing at a level that, when compared to the outcomes observed in other OECD countries, was only average. (2013b: 44)

Further, it would be reasonable to suggest that the persistent drive to focus heavily on this single method of enquiry is problematic, especially when value-added measures are used for punitive purposes, such as the deployment of scarce resources and the attrition of teacher credibility. This perspective resonates with Simons (2004) who states:

Government seeks a closer relationship with the research and evaluation community and a more prescriptive role in determining the 'what' and the 'how' of policy-related inquiry. It is the 'how' – the methodology – that is problematic. Simplicity and certainty are what governments seek. Complexity and uncertainty are what we habitually deliver. (410)

Certainly, the trend towards value-added indicators has been questioned for various reasons, such as the accuracy of results being used to make judgements about school practice. Nonetheless, Scheerens, Glas, & Thomas (2003) assert that although the calculation of a school's effect on student performance is a complex process due to the wide variety of factors that inhibit student progress. However, they also state that '...the more information it is possible to have about individual students, sub-groups of students, and all students in a school as well as comparative data across a whole population (or representative sample) of schools, the more reliable and informative any subsequent analysis is likely to be' (Scheerens, Glas, & Thomas, 2003: ch.13.3, para.1). On the other hand, it seems that almost all inspectorates quite naturally agree that forming judgements using only quantitative conjectures of quality is inadequate. As an inspector participant in Brown (2013) puts it, 'I mean, if inspection is only data-driven, just send me the details by email, and I'll write you the report by e-mail' (127). Still, it is apparent that without the use of school and system-wide school performance data, the trend towards risk-based and system-wide inspections is limited. This issue is undoubtedly a key element of the variable rate of inspection change in countries, and it is this issue that forms the next part of the study.

From valued to value added

Many view the need for value for money indicators as being part of the modernisation of all public service entities. For example, Santiago (2013b), referring to Ireland's OECD country background report on Evaluation and Assessment (2012), states: 'The value for money imperative has been a fundamental part of public service modernisation, and this has given an additional importance to evaluation and assessment in the educational context' (47). Indeed, the increasing use of complex data sets as a means of judging the component of risk attached to the quality of education provided in schools has become common place for almost all countries that have moved towards risk-based and system-wide inspections. Moreover, against a backdrop of increasing school transparency and the deployment of publicly available school league tables by governments; the increasing use of value-added measures by schools may also be attributed to the need for schools to justify the quality of education provided within their contextual constraints. For example, Coe & Vischer (2013), in

reference to Murdoch and Coe (1997), state that there is evidence to suggest that the rapid use of value-added assessment by schools in England was spurred on by the belief that the publicly available statistical conjectures of school quality deployed by the English government (in the form of school leagues tables) was unfair for a variety of reasons. These reasons include, for example, the fact that such performance indicators are unable to highlight students' achievement before entering the school. In consequence, '...perceptions of the public judgements of their effectiveness...were often a factor in their choice to implement a confidential value added school monitoring system. The published school performance information included average raw achievement of a school's students which did not adjust for relevant features of the student intake' (xiii). Indeed, with the initiation of whole school regular inspections; as schools have developed their capacity to carry out their own evaluations, so too have established inspectorates when reference to the use of value-added measures are concerned. As an interview participant referring to England in Ozga (2012) states: 'Because we have all this Key Stage Data and because it is longitudinal, we are practically, without boasting, we are probably the leading administration in the world as far as value-added measures and schooling are concerned' (447).

Through the encouragement of international organisations, such as the OECD, many governments now have developed (or in the case of Ireland, are in the process of developing) complex data sets in the form of 'value added' indicators that can be used as part of the inspectorates' risk-based assessment process. In discussing the Netherlands, Shewbridge et al (2014) state, 'Similar to the free school meal bands in Northern Ireland, student achievement data are classified into separate performance bands on the basis of level of disadvantage (mainly using parental educational level). In this way, the risk assessment takes account of school context' (147). The use of value-added models is not limited to European countries, however. In the case of North America, for example, value-added models are used to ascertain the impact that individual schools and teachers have on the quality of education provided. According to the American Statistical Association (ASA), 'many states and school districts have adopted Value-Added Models (VAMs) as part of educational accountability systems. The goal of these models, which are also referred to as Value-Added Assessment (VAA) Models, is to estimate effects of individual teachers or schools on student achievement while accounting for differences in student background' (2014: 1).

As alluded to earlier, the promotion of value added indicators within school inspection frameworks is also centred on the recognition that, the component of error attached to using such indicators is reduced when using value-added indicators. This assumption holds true when compared to methods comparing a school's performance to the average raw score for a population. Therefore, these measures should be welcomed and encouraged. Value-added indicators also may be viewed as a method for the re-regulation of teachers and schools, a concept that resonates with Ball (2003) who states, 'Within this ensemble, teachers are represented and encouraged to think about themselves as individuals who calculate about themselves, 'add value' to themselves, improve their productivity, strive for excellence and live an existence of calculation' (217). Regarding the OECD's Review of Evaluation and Assessment of Portugal, Donaldson et al (2012) state that there is a greater degree of uncertainty relating to the value that schools add to student learning in countries such as Portugal, which, at the time of writing, compare externally devised school test-based scores against the total population of schools.

Another difficulty concerns the comparison of student outcomes across schools. The average results of national examinations (both in the 9th Grade and secondary education and, as of 2011/12, in the 6th Grade) at the school level are publicly disclosed with no account for the socioeconomic context of each school (or the characteristics of schools' student population). This can considerably distort considerations about the effectiveness of each school as average results do not reflect the value added by schools to student results. (132)

Ironically, adding trepidation to issues surrounding the use of value-added indicators, Donaldson et al (2012) further state, 'It is important to note that value-added models are still under development, and therefore, they are prone to error' (189). This caution regarding value-added indicators and their use as measures of teacher and school performance has been affirmed by the ASA (2014), who state, 'A VAM score may provide teachers and administrators with information on their students' performance and identify areas where improvement is needed, but it does not provide information on how to improve the teaching' (7). Yet, despite almost every inspectorate in the world stating that one of the key functions of school inspection is to ensure that a schools own evaluation results are both valid and reliable; quite paradoxically some inspectorates are now using comparative value added measures to ascertain the quality of teaching provided in comparison to, for example, schools of a similar socio economic grouping, for instance. The logic of introducing value-added indicators as value judgements within school inspectorate frameworks and ascertaining the component of risk across all school types is further described by an inspector in Brown (2013):

The other element that is vitally important is to know that there is the value added standard as well as the actual standard. So, if I am teaching in a [Name of non-disadvantaged area] and all of my students come in at X level, am I really adding value even if 90% of them go to third-level institutions? Is it my doing as a school, or is it that they would reach that anyway because of a whole lot of other factors? So, I think that value added is one thing that you have to take into account, and then the other thing is that the national norms, but the national norms on a contextual basis. So, there are all those skills and systems that need to be put in place. (189)

None the less, for all its perceived benefits and connotations of quality in school inspection policy and practice, the use of value-added data presents a dilemma to most countries and will continue to do so in the future. However, given the ever-increasing need to introduce value-added data to formulate inspection judgements, the following supposition by O'Neill (2011) on the potential misuse of data in education is compelling:

Every time you use a measure of pupil attainment for some extraneous purpose you risk creating a perverse incentive. So every time you find yourself doing that or participating in a system that requires you to do that, ask what incentive is this creating? Whom is it damaging? Pupils, head teachers, schools, employers and I think that would be a very useful first step.

Discussion and Conclusion

In this century, changes in inspection policy and practice in all public services have been brought about as a direct result of a closer alignment with new public management philosophies. As a result, inspections tend to operate within the two school governance philosophies as described by Van Bruggen, namely accountability and continuous improvement. (2010b). Within the framework of this 'new school inspection', inspection tends to operate as either cyclical or targeted risk-based. However, for a variety of reasons such as the high costs accrued from carrying out repeated cyclical inspections, there has been a significant trend towards risk-based inspection models in most countries. In either case, the current iteration of inspection in action consists of a series of overlapping activities including the increasing use of assessment data coupled with an analysis of a schools own internal evaluations. While it can be argued that these developments as described above have resulted in more flexible inspection better designed to achieve a variety of very different objectives, these methods and models also pose significant challenges.

These challenges include, but are not limited to: (i) the capacity of schools to carry out their own evaluations; (ii) the levels of trust and respect between schools and the inspectorate; (iii) the level of inspection power or influence deemed necessary to bring about school improvement; and (iv) how inspectorates can move towards a model of risk based inspections while at the same time, ensuring that all schools who receive, at least, a satisfactory inspection continue with their school improvement agenda. Solving these concerns relates to the pursuit of a workable, integrated inspection system as envisaged by Barber and Mourshed (2009) who state: 'if the challenge of the 1980s was describing what effective schools are, of the 1990s, how to make schools more effective and of the 2000s, describing what effective systems are, then the pressing question for the 2010s is how to make systems more effective' (7).

What does appear to be as certain as it is possible to be about any educational matter is that inspection in whatever evolving form is likely to be a feature of the global landscape of schooling for the long term. Speaking of the UK but surely more widely applicable is the observation of Davis & Martins (2008b)

Whatever the future holds, it is clear that a combination of tighter resources and rising expectations on the part of both the public and politicians will continue to drive demands for better, more personalised, more responsive, more efficient, and more cost effective public services. This may throw some of the tensions between national and local priorities into even sharper relief. But the inspectorates have demonstrated a remarkable capacity to adapt and respond to changing government priorities, and in our view the smart money is on continued refining and fine tuning of inspection of public services, rather than its wholesale abandonment. (150)

References

- Atkinson, PA & Coffey, A. (2011) Analysing Documentary Realities. In: Silverman D (eds), *Qualitative Research - 3rd edition*, United Kingdom: Sage Publications Ltd, pp. 77 – 93.
- Alvik, T. (1996) *Self-evaluation: what, why, how, by whom, for whom?* Dundee: Scottish Consultative Council on the Curriculum.
- American Statistical Association. (ASA) (2014) ASA Statement on Using Value-Added Models for Educational Assessment. Available at: http://www.amstat.org/policy/pdfs/ASA_VAM_Statement.pdf (accessed May 1 2014).
- Ball, S. (2012) *Global Education Inc.: New Policy Networks and the Neoliberal Imaginary*. United Kingdom: Routledge.
- Ball, S. (2003) The teacher's soul and the terrors of performativity. *Journal of Education Policy*, 18(2): 215–228.
- Barber, M. & Moursehead, M. (2009) Shaping the Future: How Good Education Systems Can Become Great in the Decade Ahead - Report on the International Education Roundtable, July 7. Available at: http://www.mckinsey.com/locations/southeastasia/knowledge/Education_Roundtable.pdf (accessed 20 April 2014).
- Barnett, R. (1994) Power, Enlightenment and Quality Evaluation. *European Journal of Education*, Wiley, 29(2): 165–179.
- Beckmann, A., Cooper, C. & Hill, D. (2009) Neoliberalisation and managerialisation of 'education' in England and Wales—a case for reconstructing education. *Journal for Critical Education Policy Studies*, 7 (2): 311–345.
- Bowen, G. A. (2009) Document analysis as a qualitative research method. *Qualitative research journal*, 9 (2): 27- 40.
- Brown, M. (2013) Deconstructing Evaluation in Education: The case of Ireland. PhD Thesis, Dublin City University, Ireland.
- Cambridge, J. & Carthew, C. (2007) Schools Self-evaluating Their International Values: a Case Study. In: Hayden M, Thompson J, and Levy J (eds), *The SAGE Handbook of Research in International Education*, United Kingdom: Sage Publications Ltd, pp. 283–298.
- Coe, R. & Visscher, A. (2006) *School Improvement Through Performance Feedback*. United Kingdom: Taylor & Francis Ltd.
- Considine, M. (2002) The End of the Line? Accountable Governance in the Age of Networks, Partnerships, and Joined-Up Services. *Governance*, 15(1): 21–40.

- Davis, H. & Martin, S. (2008a) The Rise of Public Service Inspection. In: Davis H, and Martin M (eds), *Public Service Inspection in the UK*, United Kingdom: Jessica Kingsley Publishers, pp. 7-19.
- Davis, H. and Martin, S. (2008b) The Future of Public Service Inspection. In: Davis H., and Martin, M. (eds), *Public Service Inspection in the UK*, United Kingdom: Jessica Kingsley Publishers, pp. 135-152.
- Dedering, K. and Müller, S. (2010) School improvement through inspections? First empirical insights from Germany. *Journal of Educational Change*, Springer, 12(3): 301-322.
- Department of Education, Ontario. (2007) Ontario statistical neighbours - Informing our strategy to improve student achievement. Available at: <http://www.edu.gov.on.ca/eng/literacynumeracy/osneng.pdf> (accessed March 19 2014)
- Donaldson, G. (2013) The SICI Bratislava memorandum on inspection and innovation. Available at: <http://www.sici-inspectorates.eu/About-us/Vision-mission/The-Bratislava-Memorandum-is-on-the-Website> (accessed October 20 2014).
- Donaldson, G., Looney, A. & Nusche D. et al (2012) *OECD reviews of evaluation and assessment in education: Portugal*. Paris: OECD Publishing.
- Durrant J. and Holden G. (2005) *Teachers leading change: Doing research for school improvement*. London: Paul Chapman Educational Publishing.
- Ehren, M. Altrichter H., McNamara G. et al. (2013) Impact of school inspections on improvement of schools – describing assumptions on causal mechanisms in six European countries. *Educational Assessment, Evaluation and Accountability*, 25(1): 343.
- Ehren, M. Altrichter, H. McNamara, G. et al. (2011) School Inspections: Studying the impact of school inspection on improvement of schools. Available at: <http://schoolinspections.eu/home/the-project/> (accessed 24 December 2013)
- Etzioni, A. (2010) Is transparency the best disinfectant? *The Journal of Political Philosophy*, 18 (4): 389-404.
- Ferrer, A. (2010) Teachers in charge? Internal school assessment and evaluation in Europe. In: Stoney S (eds), *Beyond Lisbon 2010: Perspectives from research and development for education policy in Europe (CIDREE Yearbook 2010)*, Slough: National Foundation for Educational Research, pp. 119-137.
- Gaertner, H. Wurster, S. & Pant, H.A. (2013) The effect of school inspections on school improvement. *School Effectiveness and School Improvement*, 25(4): 489-508.

- Gordon, L. & Whitty, G. (1997) Giving the 'Hidden Hand' a Helping Hand? The rhetoric and reality of neoliberal education reform in England and New Zealand. *Comparative Education*, 33(3): 453-467.
- Grek, S. (2014) Travelling inspectors and the making of Europe: Education policy learning and the case of the Scottish school inspectorate. *SISYPHUS Journal of education*, 2 (1): 40-61.
- Grek, S. Lawn, M. Ozga, J. et al (2013) Governing by inspection? European inspectorates and the creation of a European education policy space. *Comparative Education*, 49(4): 486-502.
- Hall, C. & Noyes, A. (2009) New regimes of truth: The impact of performative school self evaluation systems on teachers' professional identities. *Teaching and Teacher Education*, 25(6): 850-856.
- Hamzah, M., & Tahir, H. (2013) A Glimpse Into School Self-Evaluation in Malaysia: Are We Doing The Right Things? Or Are We Doing the Things Right? *Asian Social Science*, 9 (12): 50-51.
- Hislop, H. (2012) The quality assurance of Irish schools and the role of evaluation: current and future trends - Professor Seamus Ó Súilleabháin Memorial Lecture. Available at:
<http://www.education.ie/en/Publications/Inspection-Reports-Publications/Evaluation-Reports-Guidelines/The-Quality-Assurance-of-Irish-Schools-and-the-Role-of-Evaluation-Current-and-Future-Trends.pdf> .
(accessed 16 December 2014).
- Huber, S. (2011) School governance in Switzerland: tensions between new roles and old traditions. *Educational Management Administration & Leadership*, 39 (4):469-485.
- Karsten, S. (1999) Neoliberal education reform in the Netherlands. *Comparative Education*, 35 (3): 303-317.
- Leithwood, K. Harris, A. and Strauss, T. (2010) *Leading school turnaround: how successful leaders transform low-performing schools*. San Francisco: John Wiley & Sons.
- Lindberg, S. (2013) Mapping accountability: core concept and subtypes. *International Review of Administrative Sciences*, 79 (2): 202-226.
- MacBeath, J. (2012) *Learning in and out of school: the selected works of John MacBeath*. Abingdon: Routledge.
- MacBeath, J. (2010) School based evaluation: purposes, protocols and processes. In: Peterson P, Baker E and McGaw B (eds), *International encyclopedia of education* (3rd edition), Oxford: Elsevier, pp. 713-718.
- MacBeath, J. (2008a) Leading learning in the self-evaluating school. *School Leadership & Management*, 28, (4): 385-399.

- MacBeath, J. (2008b) Stories of compliance and subversion in a prescriptive policy environment. *Educational Management Administration & Leadership*, 36 (1): 123-148.
- MacBeath, J. (2006) *School inspection and self-evaluation: working with the new relationship*. London: Taylor and Francis.
- MacBeath, J. (1999) *Schools must speak for themselves: the case for school self-evaluation*. London: Routledge.
- McKinney, J., and Howard, L. (1998) *Public Administration: Balancing Power and Accountability*. United States: Praeger Publishers Inc.
- McNamara, G., and Nayir, K. (2014) The Increasingly Central Role of School Self-Evaluation in Inspection Systems across Europe: The Case of Ireland. *Turkish Journal of Education*, 3 (1): 48-59.
- McNamara, G., and O'Hara, J. (2008a) The importance of the concept of self-evaluation in the changing landscape of education policy. *Studies in Educational Evaluation*, 34 (3): 173-179.
- McNamara, G., and O'Hara J. (2008b) *Trusting schools and teachers: developing educational professionalism through self-evaluation*. New York: Peter Lang.
- Martin, S. (2008) Inspection of Education and Skills. In: Davis H, and Martin M (eds), *Public Service Inspection in the UK*, United Kingdom: Jessica Kingsley Publishers, pp. 52-68.
- Miliband, D. (2004) *Personalised Learning: Building a New Relationship with Schools*. London: Department for Education and Skills.
- Nevo, D. (2010) Internal and external evaluation. In: Peterson P, Baker E and McGaw B (eds), *International encyclopedia of education* (3rd edition), Oxford: Elsevier, pp. 781-785.
- Nevo, D. (2002) Dialogue evaluation: combining internal and external evaluation. In: Nevo (eds), *School-based evaluation: an international perspective*, Oxford: Elsevier, pp. 3-16.
- Nusche, D., Laveault, D., MacBeath, J. et al. (2012) *OECD Reviews of Evaluation and Assessment in Education: New Zealand 2011*. Paris: OECD Publishing.
- O'Neill, O. (2011) *Keynote interview with Baroness Onora O'Neill: British Educational Research Association (BERA) Conference*. Available at: <https://www.youtube.com/watch?v=orRhAmIysFQ>. (accessed 23 December 2014).
- Ozga, J. (2012) Governing knowledge: data, inspection and education policy in Europe. *Globalisation, Societies and Education* 10 (4): 439-455.
- Perry, C. (2013) *Approaches to school inspection*. Belfast: Northern Ireland Assembly.

- Perryman, J. (2006) Panoptic performativity and school inspection regimes: disciplinary mechanisms and life under special measures. *Journal of Education Policy*, 21 (2): 147-161.
- Power, M. (2003) Evaluating the audit explosion. *Law & policy*, 25 (3): 185-202.
- Rönnberg, L., Lindgren, J., & Segerholm, C. (2013) In the public eye: Swedish school inspection and local newspapers: exploring the audit-media relationship. *Journal of Education Policy*, 28(2): 178-197.
- Santiago, P. (2013a) The focus on evaluation and assessment. In: Nusche D, Radinger D, Santiago P., et al (eds), *OECD Review on Evaluation and Assessment in Education. Synergies for Better Learning. An international perspective on Evaluation and Assessment*, Paris: OECD Publishing, pp. 29-35.
- Santiago, P. (2013b) Trends in evaluation and assessment. In: Nusche D, Radinger D, Santiago P et al (eds), *OECD Review on Evaluation and Assessment in Education. Synergies for Better Learning. An international perspective on Evaluation and Assessment*, Paris: OECD Publishing, pp. 35-57.
- Santiago P., Donaldson G., Herman J., et al. (2011) *OECD Reviews of Evaluation and Assessment in Education: Australia*. Paris: OECD Publishing.
- Scheerens J., Glas C., & Thomas S. (2003) *Educational evaluation, assessment, and monitoring: a systemic approach*. [Kindle Edition 1.10.1]. London: Taylor & Francis.
- Shewbridge, C. (2013) School evaluation: From compliancy to quality. In: Nusche D, Radinger D, Santiago P, et al (eds), *OECD Review on Evaluation and Assessment in Education. Synergies for Better Learning. An international perspective on Evaluation and Assessment*, Paris: OECD Publishing, pp. 383-485.
- Shewbridge, C., Hulshof, M., Nusche, D., et al (2014) *OECD Reviews of Evaluation and Assessment in Education: Northern Ireland, United Kingdom*. Paris: OECD Publishing.
- Simons, H. (2013) Enhancing the Quality of Education through School Self-Evaluation. *Advances in Program Evaluation*, (14): 1-18.
- Simons, H. (2004) Utilizing evaluation evidence to enhance professional practice. *Evaluation*, 10 (4): 410-429.
- Strakova, J., and Simonová, J. (2013) Assessment in the school systems of the Czech Republic. *Assessment in Education: Principles, Policy & Practice*, 20(4): 470-490.
- Tan, C. (2013) *Learning from Shanghai: Lessons on achieving educational success*. Dordrecht: Springer.
- Thomas, G. (1998) A brief history of the genesis of the new schools' inspection system. *British Journal of Educational Studies*, 46 (4): 415-427.
- Thompson, G., & Cook, I. (2014) Education policy-making and time. *Journal of Education Policy*, 29 (5): 700-715.

- Van Amelsvoort, G., & Janssens, F. (2008) School self-evaluation and school inspection in Europe. *Educational Research and Evaluation*, 34 (1): 15-23.
- Van Bruggen, J. (2010a) *Inspectorates of Education in Europe; some comparative remarks about their tasks and work*. Brussels: SICI.
- Van Bruggen, J. (2010b) The role of school inspection in ensuring quality in education: past, present and future. In: Stoney S (eds), *Beyond Lisbon 2010: Perspectives from research and development for education policy in Europe (CIDREE Yearbook 2010)*, Slough: National Foundation for Educational Research, pp.119-137.
- Vanhoof, J. and Van Petegem, P. (2007) Matching internal and external evaluation in an era of accountability and school development: lessons from a Flemish perspective. *Studies in Educational Evaluation*, 33 (2): 101-119.
- Whitby, K. (2010) *School inspection: recent experiences in high performing education systems; literature review*. Reading: CfBT Education Trust.
- World Bank. (2007) Bulgaria - Education and Skills for the Knowledge Economy. A Policy Note. Available at:
http://siteresources.worldbank.org/INTBULGARIA/Resources/EducationPolicyNote_EN.pdf . (accessed 4 January 2014).