

An inspector calls! School leaders' perceptions of inspection in Wales

Gareth Evans 

Yr Athrofa: Centre for Education, University of Wales Trinity Saint David, Swansea, UK

Policy Futures in Education
2025, Vol. 0(0) 1–21
© The Author(s) 2025



Article reuse guidelines:

sagepub.com/journals-permissions
DOI: 10.1177/14782103251320826
journals.sagepub.com/home/pfe



Abstract

Accountability in education takes many forms. One of the most prominent in the Welsh context is school inspection, which is considered an important measure of school performance and a way of holding schools accountable for the public money they receive. However, inspections are not accepted as adding value to the system without exception, and much has been written internationally about the negative impact of inspection on teachers and leaders. This small-scale study was interested in better understanding the perception of school leaders, as those with ultimate responsibility for the direction and performance of schools, based on their experience of the inspection process in Wales. It used a series of in-depth, semi-structured interviews to explore school leaders' feelings about and responses to inspection, using Foucauldian insights on power and discourse as a frame. Findings from the study suggest that the conduct of inspectors was more of a concern to participants than the judgements they awarded. While school leaders accepted inspection was a necessary process linked to the need to build and maintain public confidence, they were in broad agreement that the inspectors themselves did not communicate appropriately with school staff, were too authoritarian, and were at times disrespectful.

Keywords

Accountability, school inspection, Estyn, performativity, professionalism

Introduction

Broadly speaking, when a person or an organisation is accountable for something, there is an expectation that they will be held responsible for their actions and decisions (Normore, 2004). This is nothing new and can be considered both a moral and ethical positioning to which we have become accustomed over time. Accountability is important, argues Penzer (2011: 6), 'because without it any system risks losing legitimacy and public support'. Inspection is employed in many European

Corresponding author:

Yr Athrofa: Centre for Education, University of Wales Trinity Saint David, Swansea, SAI 8EW, Wales, Swansea, UK.
Email: gareth.evans@uwtsd.ac.uk

countries as a way of maintaining system legitimacy and public support for its education institutions (Gustafsson et al., 2015). In Wales, this is the responsibility of Estyn (which translates from Welsh ‘to reach out, stretch or extend’), the nation’s publicly funded education and training inspectorate.

A process through which the performance of institutions can be assessed (Kemethofer et al., 2017), inspection involves the visit of inspectors to schools to quantify how well they are functioning against a set of nationally agreed criteria. Typically, of particular interest to inspectors in Wales are matters relating to standards, wellbeing, teaching and learning, care and support, and school leadership (Estyn, 2020). Under the most recent, complete inspection framework (2016–2022), upon which this research is based, schools in Wales were given one of four summative judgements for each of the main inspection areas, which reflected their current performance and areas for improvement. Depending on their inspection outcomes, schools were given different levels of support and monitoring. As such, schools returning predominantly ‘Excellent’ and ‘Good’ judgements would be given greater autonomy than those considered ‘Adequate and needs improvement’ or ‘Unsatisfactory and needs urgent improvement’, which were often subject to more sustained scrutiny.

The results of inspection are made publicly available through Estyn’s website, which is used by parents and other key stakeholders to gauge schools’ performance relative to others (this is true of the current framework, as well as that employed from 2016 to 2022). Inspections are therefore considered an important mechanism of accountability, despite heavy reliance on inspectors’ own professional judgement to inform their final decisions (Bezem et al., 2022). This raises a recurring tension, between the drive for standardisation using inspection frameworks and the subjective view of inspectors adjudicating, and has been the focus of much debate related to inspection reliability (Baxter, 2013; Sinkinson and Jones, 2001). For example, a recent study on the efficacy of inspection in England found a causal relationship between inspection outcomes and the characteristics of lead inspectors, with male inspectors found to have made more lenient judgements about primary schools than females (Bokhove et al., 2023).

In Wales, as in many other ‘high-stakes’ accountability systems (Sims, 2016), the outcomes of inspection are significant and can have a lasting impact on the day-to-day running of a school. For example, a school that was considered by Estyn to be ‘Unsatisfactory and needs urgent improvement’ could suffer serious consequences, such as job losses and/or a change of leadership (Eyles and Machin, 2015; Jones and Tymms, 2014). In the most extreme cases, ongoing concerns about school standards could lead to the imposition of ‘special measures’ – recognition that a school is failing to give its pupils an acceptable standard of education (Welsh Government, 2017a). Such interventions are rare, however, and while inspection outcomes can prove problematic for some (e.g. those considered to be underachieving), they can also give cause for celebration.

It is common for schools deemed by the inspectorate to be performing well to attract attention within their communities and, in much the same way as negative outcomes invite publicity, those that receive positive feedback are often rewarded with favourable reporting in the press (Rosenthal, 2004). As such, a school in receipt of an ‘Excellent’ judgement would not only succeed in enhancing teachers’ sense of efficacy and self-worth, but likely make the school more appealing to new recruits (McNatt and Judge, 2008; Meyer et al., 2002). There is an expectation on these high-performing schools that they will share what they have done with others, either formally through brokered school-to-school collaboration or more informally, via ‘effective practice’ case studies (Estyn, 2021). Nevertheless, there is a perception that inspection in Wales has historically been ‘too punitive’, a criticism recently acknowledged by a former chief inspector (Keane, 2023: 88).

This paper aims to shed new light on school leaders’ feelings towards and responses to inspection, from a specifically Welsh perspective. A series of semi-structured interviews were

conducted with four school leaders, as those with ultimate responsibility for the direction and performance of schools, and reflected variation in inspection outcomes and diversity in school size, type and geographic location. Inspired by Foucault's conceptions of discourse and power (Foucault, 1972, 1977), a Foucauldian Discourse Analysis (FDA) was undertaken and provided a framework for the exploration of school leaders' perceptions of inspection and how they affect their subsequent actions. Although small in scale and thus not representative of the wider system, the study brings forth important messages related to inspector conduct and the way stakeholders interact during the inspection process. The research was carried out before recent changes to Estyn's inspection framework, which forms the basis of its work during 2024–2030 and includes the removal of one-word judgements to describe inspection performance.

Literature review

Inspections form a key part of school accountability architecture in many countries and are considered an important tool in the school improvement armoury (Jerrim and Jones, 2024). There is however strong evidence to suggest that, in some cases at least, inspection does not impact positively on school standards or how a school is perceived within its wider community. For example, in their analysis of so-called 'stuck' schools in England, described as those subject to a continuous cycle of poor inspection outcomes, Munoz-Chereau et al. (2022) found that receipt of below 'Good' grades carried reputational damage that made improvement more difficult. In other words, rather than acting as the catalyst for favourable change, they argued that negative Ofsted (Estyn's English equivalent) judgements can play a contributory role in the onset of increasingly challenging circumstances, and make it more likely that a school will experience further poor inspection grades in the future. However, inspection outcomes are not guaranteed to impact in the same way, irrespective of locale. Drawing on their work with schools in Germany, Gaertner et al. (2014) note that principals' and teachers' perceptions of school quality were highly stable, whether they had been inspected or not.

Emotional impact of inspection

As well as debate regarding the efficacy of inspection, there has been much written about its emotional toll on staff and the potential for those being inspected to feel heightened tension and strain (e.g. Jones and Egley, 2006; Nichols et al., 2006; Verger and Parcerisa, 2017). Research has found that at its most extreme inspection can be 'damaging emotionally and professionally' (Hopkins et al., 2016: 59) and, in some cases, 'profoundly toxic, damaging trust between staff, pupils, parents and policymakers' (Park, 2013: 120). Alternatively, teachers have reported feelings of relief, elation and pride in the aftermath of inspection (e.g. Barnes, 2012; Ofsted, 2007), particularly if it is deemed to have gone well. Regardless of one's own personal experience, there is broad acceptance that inspection plays an important role in providing schools with valuable analyses of their strengths and weaknesses, thus acting as a stimulus for support and resources provided externally (Matthews and Sammons, 2004).

Overall, the global inspection regime has attracted mixed reviews (Rosenthal, 2004), albeit there is strong evidence to suggest that teachers' perceptions of the process tend to be more negative than positive (Park, 2013; Tunç et al., 2015; Vass and Simmonds, 2001). Reflecting on what she calls the 'underlying power dynamics' inherent in education inspection, Tian (2024: 5) notes that an upgrade or downgrade in inspection results can significantly impact the wellbeing of those under scrutiny. In their study of teachers' response to inspection in England, Brunsdon et al. (2006) reported unhealthy

levels of anxiety in teachers at all times; that is, before, during and after inspections. Focussing on staff at a single primary school, they found that the process of inspection (e.g. the visit of inspectors to their school) rather than its outcome had generated most anxiety.

A similar study interested in the effects of Flemish school inspection (Penninckx et al., 2016) found a strong increase in stress and anxiety amongst teachers, as well as a decreased level of professional enthusiasm before and during inspection. Drawing from a much larger sample, it concluded that respondents from schools with an unfavourable inspection judgement reported more severe emotional effects post-inspection, whereas more positive outcomes elicited ‘a substantial positive impact on the self-efficacy of individual staff members’ (Penninckx et al., 2016: 738). Irrespective of outcome, Ouston and Davies (1998: 19) noted in the aftermath of inspection the propensity of schools to ‘slow down’ while recovering from the ‘ordeal’ of having inspectors visit – a phenomenon they described as the ‘post-inspection blues’.

A profession under surveillance

The subjecting of teachers to ongoing scrutiny of performance via inspection (Brix et al., 2014), has created a climate in which they feel they are being continuously watched and challenged to achieve. This ‘surveillance’, considered by Foucault (1977) a tool of social control, is in some part owing to what Hardy and Lewis (2016: 671) describe as the ‘fetishisation of school performance data’, and an apparent obsession within the educational sphere to quantify and measure what schools, and therefore pupils, are able to do. Roberts-Holmes (2015: 1) calls this focus on measuring schools the ‘datafication’ of teaching which, he says, means that ‘teachers’ work is increasingly constrained by performativity demands to produce “appropriate” data’. Stevenson’s (2017: 537) depiction of datafication as transforming the educational process ‘into numbers that allow measurement, comparison and the functioning of high-stakes accountability systems’ feels particularly relevant in this context.

Burusić et al. (2016: 7) offer as a definition of educational effectiveness the ‘degree in which an educational system and its components and stakeholders achieve specific desired goals and effects’. But it is who or what determines those desired goals and effects that is of real interest here, given the intense pressure on schools to perform in line with Estyn expectation. Within its 2016–2022 inspection framework, effectiveness in the eyes of Estyn was determined by a school’s performance against its five key inspection areas (Estyn, 2020), of which ‘standards’ formed an important part. One is mindful however that effectiveness is not a neutral term and instead dependent on the lens through which performance is viewed. Firestone (1991: 2) reminds us that ‘defining the effectiveness of a particular school always requires choices among competing values’ and that ‘criteria of effectiveness will be the subject of political debate’. Richards (2020: 512) concurs, noting that judgements ‘can only be mediated through inspectors’ past experience’, which results in inferences and forecasts, rather than indisputable fact.

The implication for teachers and leaders is that there does not exist an objective classification of performance, with practitioners having no option but to accept ‘the reality given to us by our patterns and structures’ (Radford and Radford, 2005: 76). In some cases, this can lead to a fear amongst teachers that their idea of innovation, so often encouraged by policymakers (Welsh Government, 2017b), might conflict with criteria promoted by school inspection (Dederig and Müller, 2011). According to Cunningham (2019: 55), whose research considered the relationship between inspection and school improvement (which is itself a contested space, given what constitutes improvement for one school might be considered normal or stagnation for another) in Australia, how a

school performs in inspection will determine whether it considers itself ‘a winner or loser within the policy framework’.

Heightened pressure on school leaders

Whilst all staff within a school are in some way susceptible to the outcomes of inspection and what happens as a consequence, it is arguably school leaders that have the most to gain or lose (Segerholm and Hult, 2016). As Wallace (2001: 156) makes clear, ‘heads alone are charged with legal responsibility for running the school within the oversight of the governing body’, and thus ultimate responsibility for inspection performance rests with those appointed by governors to lead. Bolívar et al. (2013: 20) take this thinking a step further, considering it a key responsibility of school leaders to create ‘the institutional conditions that promote the effectiveness of schools as organizations’. Earley (2017: 162) considers the social and practical aspect of the school leader’s role as their ability to ‘influence others’ actions in achieving desirable ends’, which might be considered an oversimplification of the dynamic between teacher and leader. However, in the case of inspection, it is widely acknowledged that what school leaders understand by inspection outcomes will likely impact on what action they take subsequently (Quintelier, 2020).

The heightened pressure on school leaders to perform in inspection was brought into sharp focus recently, and towards the completion of this research, when it was widely reported in the media that a headteacher in England had committed suicide after receiving a negative inspection report. Ruth Perry’s family were quoted as saying she had taken her own life in January 2023 as a ‘direct result’ of the pressure put on her by the process and outcome of her inspection, which downgraded her school from ‘Outstanding’ to ‘Inadequate’ (Weale, 2023: no page). Ms Perry’s tragic death and the events leading up to it have been the subject of intense scrutiny and attracted heavy criticism, most notably from education trade unions who believe inspection to be disproportionate and counter-productive (Ng and Kingsley, 2023). There has since been pressure on England’s inspectorate, Ofsted, to reform its inspection processes (Adams, 2023; Perryman et al., 2023; Santry, 2023).

Netolicky (2020: 392) describes school leaders as having to tread a fine balance between accountability and autonomy: that is, school leaders are at once constrained by accountability regimes, rankings, comparisons and an emphasis on external testing data, but at the same time have some freedom to make their own decisions within tight parameters. This study explores these ideas further, and considers the extent to which accountability regimes – specifically those related to school inspection – constrain school leaders and, conversely, how much freedom they allow school leaders to make their own decisions. In particular, it is interested in the factors contributing to school leaders’ emotional and physical responses to inspection, and whether it is the *process* or *outcome* of inspection that is of most significance.

Theoretical framework

The study adopted an interpretivist paradigm that relies on the subjective experience of individuals to develop meaning (Arthur et al., 2012) and was rooted in the assumption that what we know to be true is a construct of the material world in which we live (Cresswell, 2009). That school leaders do not interpret or respond to inspection in a uniform way, owing to their own assumptions and encounters (e.g. Courtney, 2013; Hopkins et al., 2016), served to justify this orientation. The interest of interpretivist researchers in understanding, explaining and demystifying social reality through the eyes of participants (Cohen et al., 2018) was a central tenet in this regard. The research adopted a post-structuralist stance and drew heavily on the work of Foucault, which allowed for identification

of key themes pertinent to inspection (e.g. the impact of inspection on school staff; inspection as a mechanism of control; and the unintended consequences of inspection) and how they are manifested in a Welsh context. In particular, the study was interested in Foucauldian conceptions of ‘discourse’ and ‘power’, and the relationship between the two.

Foucault’s (1972: 54) construct of discourse, or ‘discursive formations’, as bodies of knowledge that ‘systematically form the objects of which they speak’, is useful in exploring why school leaders might respond in a particular way. According to Foucault (2002: 42), individuals are bound by the ‘rules of formation’ that determine the objects, modes of statement, concepts and thematic choices within a particular discourse. And so we can never fully separate ourselves from discourse and discursive formations, which form the basis of articulation and action (Fadyl and Nicholls, 2013). They are what privilege certain ideas and values, and make school leaders think and act in a given way. Discourses mobilise and influence, and as far as inspection is concerned, are what separates ‘good’ from ‘bad’ and gives weight to the inspectorate as purveyors of ‘truth’ or what Foucault (1977: 304) might call ‘the judges of normality’.

Evans (2022) argues that ‘performativity’ and ‘professionalism’ are two of the more prominent discourses relevant to education in contemporary Wales. The former, a discourse that pervades teachers’ work (Jeffrey, 2002), is described by Ball (2003: 216) as ‘a mode of regulation that employs judgements, comparisons and displays as means of incentive, control, attrition and change’, and can be related to the way in which teachers and leaders are compelled to evidence performance, using standardised benchmarking criteria as a guide. In the case of inspection, this means that those being inspected are required to reach certain targets, set and agreed by the inspectorate as a ruling authority on education, and provide evidence that they have done so as a way of validating their practice.

The discourse of professionalism, meanwhile, is a much more contested terrain (Mockler, 2005). It can, on the one hand, relate to the way in which members of a profession perceive their knowledge, beliefs and skills as being somehow inherent to their professional practice. It is what makes them a professional, and the currency through which they can showcase, legitimately, their proficiency in a given field. The alternative is a more reductive top-down, institutional and managerial form of professionalism (Dehghan, 2022; Leung, 2009). In essence, professionalism of this kind is more prescribed, and requires the professional to do *what is expected* of someone in their position. Troman’s (1996: 476) perception of professionalism as that being ‘defined by management and expressed in its expectations of workers and the stipulation of tasks they will perform’ would align with this thinking.

Defined as a ‘multiplicity of force relations’ (Foucault, 1976: 92), Foucault’s conceptualisation of ‘power’ relates, in simple terms, to the way in which one individual or organisation influences another. This does not mean, however, that power resides only in the hands of a few and is instead manifested in ‘localized episodes’ that have ‘effects on the entire network in which it is caught up’ (Foucault, 1977: 27). Deacon (1998: 113) moves this discussion forward, arguing that ‘multiple, local and unstable relations of power are seen as inherent in all human interactions’, and thus power is considered intangible and omnipresent, and an expression of the everyday connections we make. In this case, and building on discursive principles laid out earlier, the inherent and diffuse nature of power (Gaventa, 2003) means that a teacher is only a teacher if they have somebody to teach, and an inspector only an inspector if they have something to inspect.

Foucault’s (1976) conceptualisation of ‘power’ as boundaryless and pervasive was particularly helpful when exploring the relationship between inspectors, as those *doing* the inspecting, and school leaders, as *subjects* of inspection. Pertinent in this case is what Foucault (1977) calls ‘discipline’, and the way in which school staff are monitored and surveilled by the establishment on

behalf of its citizens. A means through which discipline and disciplinary power is exerted (Foucault, 1977), the idea of surveillance takes inspiration from Bentham's (1791) 'panopticon', a circular-shaped prison that allowed watchmen to observe occupants without them knowing they were under observation. Creating a dichotomy between the 'watchers' and the 'watched', such surveillance is, according to Fiske (1999: 218), 'the most efficient form of power, the most totalitarian and the hardest to resist'. And so panopticism, as understood by Foucault, is representative of a wider societal ordering that controls behaviours and compares performance using both visible and invisible forces (Hancock, 2018).

In his paper on the surveillance of teachers, Page (2017: 4) argues that watching the profession in this way derives from the high-stakes nature of inspection, and is now commonplace, noting that 'where once surveillance was temporal, focused on specific times and activities, teachers now work within an environment of normalised visibility'. Subsequently, knowing that what they do and how they behave is under constant scrutiny, teachers respond by customising their practice to suit what they consider to be inspector preference in an activity commonly known as 'window dressing' (Ehren et al., 2016: 87), a process through which school leaders and their staff 're-story themselves in and against the audit culture' (Stronach et al., 2002: 130). It is what prompts these acts of fabrication (Ball, 2000), and how such fabrications play out in the classroom, that is of interest here.

However, the surrender to surveilling forces is not hopeless and Foucault (1976) makes clear that power can be undermined, exposed, made fragile and thwarted in an action commonly known as 'resistance'. This resistance plays out in the attempt by individuals 'to evade, subvert or contest strategies of power' (Gaventa, 2003: 4), albeit there are pluralities of resistances present everywhere in power relations (Oksala, 2014). Haugaard (2022: 356) argues that social life is characterised by two levels of 'conflict' (a version of resistance); those that are relatively shallow and take place *within* existing social structures, and those that occur 'when actors do not wish to collaborate in structural reproduction'. In these 'deeper' conflicts, individuals are 'not bound by mutually shared perceptions of what are reasonable structural constraints' (Haugaard, 2022: 357) and, often through means of coercion or violence, seek to delegitimise the very processes on which societal norms are based. Haugaard's theorisation provides a useful mechanism for the exploration of resistance in the context of inspection.

Research design

The study's overarching aim was to better understand how school leaders in Wales interpret and respond to school inspection outcomes. To this end, the following three research questions were explored:

- (1) How do school leaders understand inspection as a concept?

This question was considered an important foundation for subsequent questions related more specifically to inspection outcomes. It worked on the assumption that school leaders interpret the outcomes of inspection in different ways, based on their own prior experience of inspection and interpretation of what inspection is designed to do (Bennett, 2003; Hopkins et al., 2016; Ouston et al., 1997).

- (2) How do school leaders interpret inspection outcomes?

This question was designed to gauge what school leaders understand by key Estyn judgements and was inspired by an earlier pilot study, involving semi-structured interviews with two school leaders, that revealed contrasting interpretations of what ‘Good’ meant in the context of inspection. For instance, one of those interviewed described ‘Good’ as ‘not good enough’, while the other was sanguine about the same judgement he considered to be an accurate reflection of his school’s accomplishments.

(3) How do school leaders respond to inspection outcomes?

The study’s third and final question was interested in the aftermath of inspection and what school leaders do as a consequence of judgements awarded. This allowed scope for exploration of participants’ behavioural responses, as well as the more practical interventions made by school leaders in response to inspection outcomes.

Participants

Four school leaders were chosen to take part in the study, on the basis that: they held a senior position within their school’s staffing structure (e.g. headteacher, deputy headteacher or assistant headteacher); their school had been inspected during the last complete inspection cycle (2016–2022); they were in post at the time of the last inspection; and there had been no significant material change to the school (e.g. merger/federation) since the last inspection.

In addition, and in order to gain as broad a range of perspectives as possible, participants from primary and secondary schools of different sizes and geographic locations were chosen. Variation in schools’ performance in their last inspection (drawing from secondary data publicly available via the Estyn website) was reflected in the inclusion of two schools that had fared particularly well, and another two that were rated relatively poorly (see [Table 1](#)). Anonymity was agreed as a pre-requisite for participation, and involved the omission of all contextual detail (e.g. exact location of school, actual report outcomes and number of pupils, etc.) that could have led to identification.

Methods and analysis

The research was conducted in three stages. The first involved a retrospective review of inspection outcomes in each school leader’s school. This was undertaken ahead of data generation to provide

Table 1. Participating school leaders and school characteristics (anonymised).

School leader	School context	School inspection outcomes
SL1	Primary of fewer than 500 pupils; area of relatively high deprivation	A mix of ‘Excellent’ and ‘Good’
SL2	Primary of more than 500 pupils; area of relatively high deprivation	A mix of ‘Excellent’ and ‘Good’
SL3	Secondary of fewer than 1000 pupils; area of relatively low deprivation	Mainly ‘Adequate and needs improvement’
SL4	Secondary of fewer than 1000 pupils; area of relatively high deprivation	Mainly ‘Adequate and needs improvement’

context and better understand what research participants had experienced through the process of inspection. In stage two, semi-structured interviews were employed and although interview questions were predominantly pre-planned for consistency, there was some nuance on the basis of schools' performance in Estyn inspection and school leaders were given opportunity to elaborate or provide additional information as the interviews progressed (Bailey, 2014). Interviews were conducted online during summer 2022 (before the introduction of new inspection arrangements) and recorded to support transcription.

A Foucauldian Discourse Analysis (FDA) was used in stage three to analyse data and provide greater depth of understanding of participants' views (Khan and MacEachen, 2021). Building on the study's post-structuralist stance, FDA offered a framework for the exploration of popular perceptions of inspection and how school leaders interpret and respond to inspection outcomes. The interest of FDA in power relations and power effects inherent in discourses (Powers, 2007) made it particularly useful in this context and I was drawn to its capacity to explore 'how people think, what they know and how they speak about the world' (Raby, 2002: 30). Willig's (2008) six-stage application was considered most applicable, as it provided a mechanism for the mapping of data generated against key areas of interest (as outlined in the aforementioned research questions).

Willig (2008: 115–117) lists 'discursive constructions', 'discourses', 'action orientation', 'positionings', 'practice' and 'subjectivity' as being central to FDA, and the identification of these stages was useful in that they allowed a richer analysis of the way in which both individuals and institutions were framed by participants. The simplicity of Willig's approach was another influencing factor, with the 20-stage orientation offered by Parker (1992) considered unnecessarily complicated and too political for the specific needs of this research. Building on the structure of similar studies (e.g. Phillips, 2024), the use of FDA within an interpretivist paradigm provided a means through which subjective experiences and interpretations could be properly explored, and opened up a 'window of understanding' about what individuals *really* thought about inspection (Khan and MacEachen, 2021: no page). An adaptation of Willig's model of FDA, informed by the pilot study and pertinent to the overarching research aims and questions, is offered in Table 2.

Findings

In this section, key findings related to the adapted model of Willig's (2008) six-stage FDA are presented in turn, using the questions presented in Table 2 as a scaffold for discussion. School leaders are labelled SL1-4, in accordance with Table 1, for ease of reference.

Stage 1: How do the school leaders describe inspection? On what are these descriptions based?

School leaders described inspection in myriad ways, although there were a number of recurring themes. The most obvious, evident in all participant interviews, was the role of inspection as a form of accountability. Participants described the inspection process as an 'audit' (SL1) and a mechanism by which the inspectorate could 'inform' various stakeholders about how their schools are faring (SL4). SL2 considered it important that schools demonstrate 'value for money' in their offer to children. Other descriptions included the perception of inspection as being intrusive and something to be feared. SL4 used the following analogy to describe the process from his perspective:

I've said a few times that in the run-up to that inspection it was like watching a slow motion car crash – I knew it was coming and there wasn't anything I could do about it. And then the car crashed and you kind

Table 2. Adaptation of Willig's (2008) six-stage FDA.

Stages	Description	Adaptation to research questions
1. Discursive constructions	How is the discursive object constructed?	How do the school leaders describe inspection? On what are these descriptions based?
2. Discourses	Similarities and differences in constructions – location in wider discourses. What is their relationship to one another?	What are the similarities and differences in these descriptions and the wider discourses in which they are located?
3. Action orientation	Context – what constructions achieve; what are their purpose? The functions and gains generated through constructing the object in a certain way	What do the school leaders do, based on their understanding of the inspection (action), and why do they do it (orientation)?
4. Positionings	Subject positions in discourse – rights and duties. How are different people positioned?	Where are the school leaders positioned in relation to the discourse around inspection? How does this effect what they can and cannot do?
5. Practice	How subject positions open or close opportunities – what can people say and do given the positions made available in discourses?	How do school leaders' subject positions open up or close down opportunities for action?
6. Subjectivity	Feelings, thoughts and experiences – social and psychological effects on the subject	What can be felt, thought and experienced by school leaders? How does this impact on their behaviour and interactions with others?

of think you're done, but then you realise, no it's not done because the car crash is going to last for ages. (SL4).

SL3 described inspection as a 'very traumatic experience', which he said stemmed from the possible ramifications linked to poor outcomes. He made reference to the 'fallout' from inspection, citing the potential for reputational damage to both schools and staff, and resolved that he 'wouldn't want to put other people through what we went through'. Despite the broadly negative descriptions of inspection, there was a feeling widely shared amongst participants that inspection is a necessary process, closely connected to the need to build and maintain public confidence in schools as educational institutions.

Stage 2: What are the similarities and differences in these descriptions and the wider discourses in which they are located?

This section explores how the discursive constructions presented in Stage 1 align with the prominent discourses of performativity and professionalism, outlined earlier. For example, the suggestion by SL3 that inspection is designed 'to make sure that they (schools) reach the required standards' can be considered performativity in action, as schools work to meet the requirements of the inspection framework.

There are also performative connotations within the construction of inspection as intrusive. This stems largely from the idea that inspectors are let into school to do as they wish, regardless of

whether school leaders want them there in the first place. The process of granting inspectors access to classrooms and other points of interest is accepted practice, albeit there was a feeling that inspectors do so without staff's full approval, as demonstrated by reference to the inspectorate 'ransacking' (SL3), 'rifling through drawers' (SL1) and looking through cupboards (SL2).

You know what, it's almost like someone coming into your house and ransacking it, [and] then telling you what's wrong with it and leaving it in a mess. It's almost personal – and it shouldn't be – because the whole premise of inspection is that it's supposed to be a good thing. It's more about the way it's done – there's too much pressure and too much devastation. (SL3).

The most obvious manifestation of professionalism was that school leaders prepared diligently for inspection so as to present the best possible version of themselves and their staff. They apparently did this both to respond to the performative culture that surrounds the inspection process (i.e. because the discourse of performativity requires that they have to), and as a symbol of their professional pride in what they are employed to do (i.e. because of their inherent professionalism, they want to prepare well). Take, for example, SL2's admission that she encouraged her staff to 'plan in advance' in readiness for the visit of inspectors, as she 'didn't want to leave anything to chance'.

Two of the school leaders interviewed made explicit reference to a lack of opportunity for 'professional dialogue' with inspectors. SL1 and SL3 both expressed disappointment that they were not given a chance to discuss in greater detail their inspection outcomes, nor sufficient opportunity to feed into scrutiny of school processes and practice during the inspection itself. SL1 said 'there was a lack of professional dialogue' which led to 'a kind of die in the ditch conversation where one side is fighting for one word and the other side is fighting for the other'. She described how one of her senior leadership team 'had her hand pushed into her face' after asking inspectors for their attention. SL3 made similar representation about the conduct of those inspecting, recalling that the inspection process 'wasn't ultra friendly' and 'wasn't supportive'. He said the inspection had not been undertaken in a collegiate way and 'was done to us, not with us'. SL2 said the way in which inspectors 'conduct themselves personally' was an issue:

It's the way that they're framing the questions – the body language, the facial looks coming back. You know, it's not as an equal, you're not having a conversation with an equal professional. They make you feel that 'what I say can affect your career'. (SL2).

In what could be seen as an attack on her professional identity, SL2 describes in this excerpt a clear and *unequal* separation of school staff from those inspecting, born out of a hierarchical model that affords inspectors control over teachers' time and practice when in school. Frostenson (2015: 20) considers such actions 'the hallmark of de-professionalisation, a process wherein professional actors lose the ability to influence and the power to define the contents and forms of their own work'.

Stage 3: What do the school leaders do, based on their understanding of the inspection (action), and why do they do it (orientation)?

School leaders presented contrasting views of what happens in the aftermath of inspection, which appeared almost entirely dependent on how their school had performed and what judgements had been given. For example, those who returned favourable inspection results felt a sense of liberation.

It did give us the confidence to take the school off in the direction that we wanted to go in. It just gave us that clean break almost to go and do something different. Once we'd seen the back of them, we could move forward without having to have to answer that scrutiny. (SL2).

Those in receipt of weaker judgements described a very different set of influences; SL3 and SL4 said the inspection process had consumed their schools for many months after the publication of outcomes, and spoke of the profound emotional effect of their school's inspection result on them personally. SL4 described feeling 'really isolated' as the person responsible for making key decisions post-inspection, while SL3 called his school's inspection outcomes 'a major distraction' and something that dominated the entire school's work. However, things changed quickly once the school was removed from Estyn monitoring:

We're now a professional learning school, [and] we're getting money thrown at us by the local consortium. They wanna work with us a lot more... And it's almost immediate – as soon as you've got out of monitoring, you're invited back in to be a professional learning school because they know the stuff that you're doing. And that's the thing, once you're out of it, you can really go for it. But if you're stuck in it, you're stuck; it's almost like the rich get richer and the poor get poorer. (SL3).

Stage 4: Where are the school leaders positioned in relation to the discourse around inspection? How does this effect what they can and cannot do?

The positioning of school leaders as being subordinate to visiting inspectors was a feature of all four interviews. However, subordination was not absolute and, in many cases, did not endure for the entirety of the inspection process. For example, SL3 described a higher level of subservience during his senior leadership's final meeting with inspectors:

It was a bit surreal – we all had to stand up as they (the inspection team) walked in and they sat down and just went through it, and there was no interaction... It was almost like a verdict in a court and then off they went and left us to, like, pick up the pieces. It was a very surreal experience, I gotta say, it left me flat. Actually, I was devastated. (SL3).

In court, the way in which those present rise to their feet as the judge delivers their verdict is both a display of respect to the higher authority, and a practice that has become normalised through the discourse of law and order. That SL3 and his colleagues saw fit to stand in such a fashion is a symptom of the performative culture in which they operate. SL3's reflection that 'we all had to stand up as they walked in' is not a literal representation of what happened; the staff did not *have* to stand up as inspectors entered the room, but they did so presumably because they *felt they had to*.

SL4 described what could be considered a similarly high level of subordination, although manifested in a different way. He spoke of his regret that visiting inspectors had focussed so heavily on his school's pupil behaviour statistics, despite them having seen little evidence of poor behaviour during their time in school.

We had this recommendation around behaviour, that behaviour had to improve because it was 'Adequate and needs improvement', on the basis that we'd had an increase in exclusions. But, you know, exclusions go up and down – that's kind of the point of exclusions – and I remember saying at the time 'have you seen any instances of poor behaviour?' and they said 'none... none at all'. (SL4).

SL4 later reflected, in reference to behaviour, ‘that decisions had already been made before they came in’ and ‘I don’t think there was anything they would have seen that would have changed it (their judgement)’. Instead, he said ‘it (Estyn’s interpretation of his school’s performance) was done from the minute they walked in; you just knew and so it was just damage limitations’. SL1, whose school had performed well by comparison, was of a similar opinion, noting that in her experience inspectors had ‘a fixed view’ about schools ‘that they then have to find the evidence to back up’.

Another less prominent, although no less significant positioning was that of school leaders as social outcasts. This was best demonstrated by SL4, who spoke of ‘being the person lots of people talk to, to the person nobody talks to’ following publication of his school’s disappointing inspection outcomes. A perceived loss of confidence in his ability as a school leader manifested in a reluctance to support others, either directly (via school-to-school collaboration) or indirectly (by presenting at conferences). When asked for more specific detail, he said his inspection outcomes meant ‘I don’t want to share as much’ and ‘I’m less collaborative’. SL4 said he no longer volunteered to present publicly about education as ‘you don’t want to set yourself up for a fall’ and described being ‘more protective’ of himself and his school in the aftermath of inspection.

Stage 5: How do school leaders’ subject positions open up or close down opportunities for action?

There is a material consequence to the positioning of school leaders as subordinate, which was encapsulated by SL3 in his reflections on his school’s restricted access to professional opportunities post-inspection (see *Stage 3*). He said the school was much healthier financially since its reinstatement as a ‘professional learning school’, given it was now able to provide paid-for professional support to other schools.

We’re back being a professional learning school, which is brilliant because we’re now on the cutting-edge of, um, things like Curriculum for Wales, [and] local and national priorities. But when you’re not a professional learning school, and you can’t have [initial teacher education] students or whatever, you almost feel cut off so it’s almost unfair in a way, I think, that that’s put on the school. (SL3).

In this case, the reinstating of SL3’s school also appears to have led to more active participation in curriculum reform, which was ongoing at the time of writing and includes a comprehensive overhaul of Wales’ national curriculum offer for the entire 3–16 age range (Donaldson, 2015; Evans, 2023). This implies that schools with a professional learning designation have an advantage over those who do not, and casts doubt over the capacity of schools deemed by the inspectorate to be failing to properly engage in curricular developments (a point apparently supported by SL3’s earlier reference to *the rich getting richer and the poor getting poorer*).

Also of note in *Stage 5* is the potential for school leaders to resist inspector representations, as demonstrated by the decision of two participants to make formal complaints about the conduct of those inspecting. In what could be considered an illustration of *shallow conflict* (Haugaard, 2022), one school leader complained that an inspector had been ‘power mad... to the point it was unpleasant’, whereas another felt compelled to share with the inspectorate ways in which they believed inspectors had been disrespectful.

One of the inspectors was quite aggressive in some respects, and so dismissive of some staff. They were determined that they were in the driving seat – and that this was no kind of partnership. And for them to

put a hand in the face of somebody who had worked so hard – I couldn't imagine allowing anybody else to come into our school and to, you know, to treat us in that way. (SL1).

Such complaints can be considered a more practical and, to some extent, scaffolded form of resistance, based on inspection protocol (and thus respecting of existing social structures). Interesting in this case, however, is that both complainants appeared less confident in their ability to challenge successfully what the inspectors decided following their inspection visit, than they were inspectors' behaviour. In other words, they were more comfortable criticising formally the *conduct* of inspectors (i.e. what they did and how they behaved) than what the inspectors actually *had to say* (i.e. their judgements). Drawing again on Haugaard's contrast in conflicts, it could be argued that challenging inspector judgements is of 'deeper' concern than the criticism levelled here, as it implies a level of distrust in established structures; in this case, the inspection framework.

The idea that challenging, and overturning, inspection judgements is a forlorn hope was nevertheless a recurrent irritation. SL1 said that while she 'took them (inspectors) on repeatedly', her protestations were ultimately futile ('it fell on deaf ears') and she felt certain that 'no matter what I said or did, it wouldn't have made any difference'. SL4 framed a disagreement with inspectors over his school's decision to enter all pupils for a particular GCSE qualification as 'a real fight' that 'we did challenge, but they didn't listen'.

Stage 6: What can be felt, thought and experienced by school leaders? How does this impact on their behaviour and interactions with others?

A common thread running through a number of interviews was an underlying feeling of failure on the part of participants, who considered their school's inspection outcomes a negative reflection of their own performance as school leaders. For SL2, this feeling was prevalent after *all* inspections, regardless of judgement awarded. In what might be considered symptomatic of the 'post-inspection blues' (Ouston and Davies, 1998), she described 'this sinking feeling after an inspection, no matter what the outcome is' and the tendency of teachers to hold onto even the smallest of criticisms, such as their professional pride in their work. She spoke of 'that tiny chunk of a development area' that 'hangs with you and makes you feel rotten for ages' as being reflective of a teacher's desire to do their best by pupils, even in the most challenging of circumstances.

Colman (2022: 18) notes how the disciplinary gaze of the inspectorate 'invokes in leaders feelings of anxiety and fear, amid relentless pressure to perform'. Similar feelings were surfaced during conversations with SL3 and SL4, who both questioned their own futures as school leaders as a result of their inspection experience. In a more acute articulation of this, SL4 recalled a series of exchanges with his wife, a local authority adviser, and his deputy, in which he questioned his capacity to 'turn around' his school's fortunes and speculated on what other jobs might be available to him if he were to step back as headteacher ('if I don't do this, what can I do?'). SL3 entertained similar thoughts, describing fallout from his school's inspection report as 'the lowest point in my career'. He later confessed to having considered leaving school to become a delivery driver in his partner's delivery firm.

That's how low it was, I honestly saw driving around delivering things as a... that would have been a complete relief [rather] than going back into school every day at a time when it was horrible. So that's how bad it got. (SL3).

The most extreme response to inspection that arose during interview was the prospect of school leaders taking their own lives as a result of their experiences. The shocking revelation that an experienced school leader known to one of the participants had contemplated suicide following their school's inspection was surfaced during an exchange with SL2, who said that the lead inspector's conduct 'was so awful' her colleague had been made to 'go into a very dark place'.

I think if within Estyn, if you just shoot them (school leaders) down and just tell them that they're not doing enough and make them flustered... and actually the member of staff is exceptionally strong, you know, as a person. So when she said to me that it had made her, um, go into a very dark place, um, and that she had contemplated suicide, I just, I just couldn't believe it. (SL2).

SL2 said a formal complaint was submitted but the inspector continued in their role, despite the severity of the accusations made against them and the impact of the inspector's behaviour on her colleague's mental health.

Conclusion

This research adds to the existing weight of literature that views inspection as a necessary but complex mechanism of accountability. This complexity, which stems from the need of inspectors to make informed and fair judgement in what is by design an imperfect system, is not without consequence and whilst acknowledging the important role of inspection in maintaining standards, participants described a process of which they were fearful; a fear often deriving from the high-stakes nature of inspection and the potential consequences associated with a poor inspection result. However on this occasion, school leaders' overwhelmingly negative perceptions of inspection were not fuelled so much by their inspection outcomes (and whether or not they thought they had been harshly treated in the award of inspection judgements), but by the way the inspections themselves were conducted. Participants appeared particularly aggrieved about what visiting inspectors did and how they behaved during the inspection process, thus corroborating the findings of earlier research (Brunsden et al., 2006; Courtney, 2013).

Perryman (2007) argues that performativity is inextricably linked to the increased accountability and surveillance under which teachers find themselves and their schools being judged. She describes a process of 'jumping through hoops' to present one's practice in the best possible light, only doing so 'can lead to teachers' sense of emotional dissonance, as they lose their sense of professional independence' (Perryman, 2007: 176). At its most extreme, this can manifest in particularly distressing ways, as demonstrated by the revelation in this study that a school leader's colleague had contemplated suicide following their inspection. Indeed, school leaders appear as vulnerable as any to severe personal anguish, contrary to suggestion that stress levels reduce through experience (Hopkins et al., 2016). At the time of writing, debate regarding the future of inspection in England remains a live issue, although evidence presented here suggests that such extremities, involving the potential for intense emotional distress amongst those subject to inspection, are not unique to England and are of relevance to Wales also.

Notwithstanding the positive experiences of many in the school system (Penninckx et al., 2016; Quesel et al., 2020), the well-documented link between inspection and psychological trauma could in turn lead to further challenges and in some cases exacerbate ongoing difficulties with the recruitment and retention of teachers (Hutchinson et al., 2024). A recent survey of more than 6700 teachers and school leaders in England found that 76% thought inspection had a negative effect on retention, with 30% considering leaving the profession as a result of their most recent

inspection (Perryman et al., 2024). That two of the four school leaders interviewed in this research contemplated leaving the profession as a result of their inspection experience serves as a salutary reminder of the high-stakes nature of inspection, which requires that schools ‘hold themselves in a state of perpetual readiness’ (Perryman et al., 2018: 148) so as to satisfy the requirements of a pervasive audit culture.

There are other factors impacting on school leaders’ negative perceptions of inspection. For example, schools in receipt of weaker Estyn judgements are not typically afforded the same range of developmental opportunities as other, more successful schools in relation to, for example, professional learning, curriculum building or initial teacher education, which in the Welsh case stipulates that only schools identified by Estyn as being ‘effective’ should directly contribute to supporting aspiring teachers through their training (Welsh Government, 2023a). Thus it could be argued that the capacity of schools in monitoring to improve is further impeded by restricted access to such things which, somewhat perversely, contributes to the consolidation of already high-performing schools. Indeed, the suggestion that schools with a ‘professional learning’ designation are at ‘the cutting-edge’ of the Curriculum for Wales implies that those without are not. The implication here, therefore, is that denying schools in monitoring chance to broaden their horizons – through enrichment activities available to so-called stronger schools – could in fact advance their decline as other schools begin to feel the benefit of new knowledge and understanding that only they have access to.

Another unfortunate consequence of the panoptic imposition of power and surveillance over schools and school leaders is that it can sometimes make them more protective of their practice, as opposed to wanting to share the celebrated aspects of their work. This was demonstrated most obviously by the school leader that had, following his disappointing inspection outcome, declined to present at conferences for fear of making matters worse. At a time when schools in Wales are being actively encouraged to collaborate (Welsh Government, 2023b), that inspection makes some school leaders do the exact opposite seems entirely illogical. Buchanan (2015: 705) describes how the paradigm of accountability, with its emphasis on what can be measured and compared, ‘gives rise to particular professional practices that teachers and schools must engage in to remain legitimate’. It is this quest for legitimacy, in the eyes of the inspectorate principally, but also in the eyes of their peers and wider school communities, that can make school leaders behave and respond in undesirable ways.

Tian (2024: no page) describes inspection in England as being at a ‘critical tipping point’, with a fundamental decision to be made about whether inspections should be high-stakes (aimed at accountability) or low-stakes (aimed at fostering development). Recent changes to inspection in Wales, including the removal of one-word judgements in favour of more descriptive reporting of a school’s achievements, suggest that Estyn is already on a journey to becoming lower-stakes, although evidence presented here would imply that the inspectorate will need to encourage a change in behaviours as well as practice if it is to truly transform the inspection experience for those at the sharp-end of its activity.

In conclusion, the findings of this research should be taken for what they are: a snapshot of some school leaders’ perceptions of inspection, based on their own lived experiences. These insights have implications for both the inspectorate and the inspection regime more generally, and prompt further questions about the centrality of its position within the broader educational ecosystem. The study’s small sample size is an obvious limitation, but should not by itself preclude any remedial action from being taken and, in the absence of other such studies, it is important that the voices of these and other school leaders in Wales are heard.

Acknowledgements

I would like to thank the school leaders who participated in this study and spoke so honestly about their experiences. Thanks also to my doctoral supervisors, Jane Waters-Davies and Susan Jones, for their wise counsel and to the anonymous reviewers of this article, whose valuable comments have doubtless strengthened its presentation.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Ethical statement

Ethical approval

Ethical approval was gained from the University of Wales Trinity Saint David research ethics committee.

ORCID iD

Gareth Evans  <https://orcid.org/0000-0002-5646-2576>

References

- Adams R (2023) *Former Ofsted chief: school inspections should change after headteacher's death*. Available at: <https://www.theguardian.com/education/2023/jun/16/former-ofsted-chief-schoolinspections-should-change-after-headteacher-death> (accessed 20 June 2024).
- Arthur J, Waring M, Coe R, et al. (2012) *Research Methods and Methodologies in Education*. London: Sage Publications.
- Bailey L (2014) The origin and success of qualitative research. *International Journal of Market Research* 56(2): 167–184.
- Ball S (2000) Performativities and fabrications in the education economy: towards the performative society? *Australian Educational Researcher* 27: 1–23.
- Ball S (2003) The teacher's soul and the terrors of performativity. *Journal of Education Policy* 18(2): 215–228.
- Barnes JM (2012) *What Sustains a Life in Education*. UK: Canterbury Christ Church University. PhD Thesis.
- Baxter JA (2013) Professional inspector or inspecting professional? Teachers as inspectors in a new regulatory regime for education in England. *Cambridge Journal of Education* 43(4): 467–485.
- Bennett N (2003) Structure, culture and power in organisations. In: Bennett N, Crawford M and Cartwright M (eds) *Effective Educational Leadership*. London: Paul Chapman, 44–61.
- Bentham J (1791) *Panopticon: or, the Inspection-House*. Dublin: Thomas Byrne.
- Bezem P, Piezunka A and Jacobsen R (2022) School inspectors' decision-making: evidence from a comparative perspective. *Leadership and Policy in Schools* 23(2): 253–274.
- Bokhove C, Jerrim J and Sims S (2023) Are some school inspectors more lenient than others? *School Effectiveness and School Improvement* 34(4): 419–441.
- Bolívar A, López J and Murillo F (2013) Liderazgo en las instituciones educativas. Una revisión de líneas de investigación. *Revista Fuentes* 14: 15–60.

- Brix J, Grainger P and Hill A (2014) Investigating mandatory peer review of teaching in schools. *Australian Journal of Teacher Education* 39(4): 83–99.
- Brunsdon V, Shelvin M and Davies M (2006) Anxiety and stress in education professionals in relation to Ofsted. *Education Today* 56(1): 24–31.
- Buchanan R (2015) Teacher identity and agency in an era of accountability. *Teachers and Teaching* 21(6): 700–719.
- Burusić J, Babarović T and Ms V (2016) School effectiveness: an overview of conceptual, methodological and empirical foundations. In: Alfirević N, Burusić J, Pavčić J, et al. (eds) *School Effectiveness and Educational Management*. Cham: Palgrave Macmillan, 5–26.
- Cohen L, Manion L and Morrison K (2018) *Research Methods in Education*. 8th edition. London: Routledge.
- Colman A (2022) School leadership, inspection and visibility: panopticism and post-panopticism in an English coastal area of deprivation. *British Journal of Educational Studies* 70(1): 3–19.
- Courtney SJ (2013) Head teachers' experiences of school inspection under Ofsted's January 2012 framework. *Management in Education* 27(4): 164–169.
- Creswell JW (2009) *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. 3rd edition. Thousand Oaks, CA: Sage Publications.
- Cunningham C (2019) An investigation into school inspection policies in Western Australian state education performed by the expert review group. *Educational Research for Policy and Practice* 18(1): 39–58.
- Deacon R (1998) Strategies of governance: Michel Foucault on power. *Theoria* 45(92): 113–148.
- Dederig K and Müller S (2011) School improvement through inspections? First empirical insights from Germany. *Journal of Educational Change* 12(3): 301–322.
- Dehghan F (2022) Teachers' perceptions of professionalism: a top-down or a bottom-up decision-making process? *Professional Development in Education* 48(4): 705–714.
- Donaldson G (2015) *Successful Futures: Independent Review of Curriculum and Assessment Arrangements in Wales*. Cardiff: Welsh Government.
- Earley P (2017) Conceptions of leadership and leading the learning. In: Earley P and Greany T (eds) *School Leadership and Education System Reform*. London: Bloomsbury, 161–172.
- Ehren MCM, Jones K and Perryman J (2016) Side effects of school inspection: Motivations and contexts for strategic responses. In: Ehren MCM (ed) *Methods and Modalities of Effective School Inspections: Accountability and Educational Improvement*. Denmark: Springer International, 87–109.
- Estyn (2020) *Statistical first release: inspection outcomes of education providers in Wales inspected during the 2019-2020 academic year*. Available at: https://www.estyn.gov.wales/system/files/2022-01/Estyn_OfficialStatsPublication_Sept2020_en.pdf (accessed 10 June 2024).
- Estyn (2021) *Improvement Resources: Effective Practice*. Cardiff: Estyn.
- Evans G (2022) Back to the future? Reflections on three phases of education policy reform in Wales and their implications for teachers. *Journal of Educational Change* 23(3): 371–396.
- Evans G (2023) A new dawn or false hope? Exploring the early implementation of Curriculum for Wales. *Education Inquiry* 2: 1–15. DOI: [10.1080/20004508.2023.2297506](https://doi.org/10.1080/20004508.2023.2297506).
- Eyles A and Machin S (2015) *The Introduction of Academy Schools to England's Education: CEP Discussion Papers*. London: Centre for Economic Performance. LSE.
- Fadyl JK and Nicholls DA (2013) Foucault, the subject and the research interview: a critique of methods. *Nursing Inquiry* 20(1): 23–29.
- Firestone WA (1991) Introduction: Chapter 1. In: Bliss JR, Firestone WA and Richards CE (eds) *Rethinking Effective Schools: Research and Practice*. Englewood Cliffs, New Jersey: Prentice Hall, 12–27.
- Fiske J (1999) *Media Matters: Race and Gender in U.S. Politics*. Minneapolis: University of Minnesota Press.
- Foucault M (1972) *The Archaeology of Knowledge*. New York: Pantheon Books.
- Foucault M (1976) *The History of Sexuality Volume 1: An Introduction*. London: Allen Lane.

- Foucault M (1977) *Discipline and Punish: The Birth of the Prison*. London: Penguin Books.
- Foucault M (2002) *The Archaeology of Knowledge (AM Sheridan Smith, Trans.)*. Abingdon: Routledge.
- Frostenson M (2015) Three forms of professional autonomy: Deprofessionalisation of teachers in a new light. *Nordic Journal of Studies in Educational Policy* 1(2): 20–29.
- Gaertner H, Wurster S and Pant HA (2014) The effect of school inspections on school improvement. *School Effectiveness and School Improvement* 25(4): 489–508.
- Gaventa J (2003) *Power after Lukes: A Review of the Literature*. Brighton: Institute of Development Studies.
- Gustafsson JE, Ehren MCM, Conyngham G, et al. (2015) From inspection to quality: ways in which school inspection influences change in schools. *Studies In Educational Evaluation* 47(December): 47–57.
- Hancock BH (2018) Michel Foucault and the problematics of power: Theorizing DTCA and Medicalized subjectivity. *Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine* 43(4): 439–468.
- Hardy I and Lewis S (2017) The ‘doublethink’ of data: educational performativity and the field of schooling practices. *British Journal of Sociology of Education* 38(5): 671–685.
- Haugaard M (2022) Foucault and power: a critique and retheorization. *Critical Review* 34(3–4): 341–371.
- Hopkins E, Hendry H, Garrod F, et al. (2016) Teachers’ views of the impact of school evaluation and external inspection processes. *Improving Schools* 19(1): 52–61.
- Hutchinson J, Kelly J, Sibieta L, et al. (2024) *Incentives to recruit and retain teachers in Wales*. London: Education Policy Institute.
- Jeffrey B (2002) Performativity and primary teacher relations. *Journal of Educational Policy* 17(5): 531–546.
- Jerrim J and Jones A (2024) The strengths and limitations of using quantitative data to inform school inspections. *School Effectiveness and School Improvement* 35(2): 142–160.
- Jones BD and Egley RJ (2006) Looking through different lenses: teachers’ and administrators’ views of accountability. *Phi Delta Kappan* 87(10): 767–771.
- Jones KL and Tymms PB (2014) Ofsted’s role in promoting school improvement: the mechanisms of the school inspection system in England. *Oxford Review of Education* 40(3): 315–330.
- Keane A (2023) Changing features of school inspection: a perspective from Wales. *Forum* 65(3): 81–90.
- Kemethofer D, Gustafsson JE and Altrichter H (2017) Comparing effects of school inspections in Sweden and Austria. *Educational Assessment, Evaluation and Accountability* 29(4): 319–337.
- Khan TH and MacEachen E (2021) Foucauldian discourse analysis: Moving beyond a social constructionist analytic. *International Journal of Qualitative Methods* 20. doi: [10.1177/16094069211018009](https://doi.org/10.1177/16094069211018009).
- Leung C (2009) Second language teacher professionalism. In: Burns A and Richards JC (eds) *The Cambridge Guide to Second Language Teacher Education*. Cambridge: Cambridge University Press, 49–58.
- Matthews P and Sammons P (2004) *Improvement through Inspection: An Evaluation of the Impact of Ofsted’s Work*. London: Ofsted.
- McNatt DB and Judge TA (2008) Self-efficacy intervention, job attitudes, and turnover: a field experiment with employees in role transition. *Human Relations* 61(6): 783–810.
- Meyer JP, Stanley DJ, Herscovitch L, et al. (2002) Affective, continuance, and normative commitment to the organization: a meta-analysis of antecedents, correlates, and consequences. *Journal of Vocational Behavior* 61(1): 20–52.
- Mockler N (2005) Trans/forming teachers: new professional learning and transformative teacher professionalism. *Journal of In-Service Education* 31(4): 733–746.
- Munoz-Chereau B, Hutchinson J and Ehren M (2022) ‘Stuck’ schools: can below good Ofsted inspections prevent sustainable improvement? Available at: https://epi.org.uk/wp-content/uploads/2022/06/Final_report_stuck_schools.pdf (accessed 10 August 2024).
- Netolicky DM (2020) School leadership during a pandemic: Navigating tensions. *Journal of Professional Capital and Community* 5(3/4): 391–395.

- Ng E and Kingsley T (2023) *Ofsted report was 'deeply harmful' to late headteacher, family says*. Available at: <https://www.independent.co.uk/news/uk/homenews/ofsted-reading-national-education-union-ruth-perry-b2304848.html> (accessed 15 July 2024).
- Nichols S, Glass G and Berliner D (2006) High-stakes testing and student achievement: does accountability pressure increase student learning? *Education Policy Analysis* 14(1): 1–174.
- Normore AH (2004) The edge of chaos: school administrators and accountability. *Journal of Educational Administration* 42(1): 55–77.
- Ofsted (2007) *Monitoring Inspections of Schools that Are Subject to Special Measures*. London: Her Majesty's Stationery Office.
- Oksala J (2014) Resistance. In: Lawlor L and Nale J (eds) *The Cambridge Foucault Lexicon*. Cambridge: Cambridge University Press, 432–437.
- Ouston J and Davies J (1998) Ofsted and afterwards? Schools' responses to inspection. In: Earley P (ed) *School Improvement after Inspection? School and LEA Responses*. London: Paul Sage, 13–24.
- Ouston J, Fidler B and Earley P (1997) What do schools do after Ofsted school inspections or before? *School Leadership & Management* 17(1): 95–104.
- Page D (2017) The surveillance of teachers and the simulation of teaching. *Journal of Education Policy* 32(1): 1–13.
- Park J (2013) *Detoxifying school accountability: the case for multi-perspective inspection*. London: Demos.
- Parker I (1992) *Discourse Dynamics: Critical Analysis for Social and Individual Psychology*. London: Routledge.
- Penninckx M, Vanhoof J, De Maeyer S, et al. (2016) Effects and side effects of Flemish school inspection. *Educational Management Administration & Leadership* 44(5): 728–744.
- Penzer G (2011) *School Inspections: What Happens Next?* Reading: CfBT Education Trust.
- Perryman J (2007) Inspection and emotion. *Cambridge Journal of Education* 37(2): 173–190.
- Perryman J, Bradbury A, Calvert G, et al. (2023) *Beyond Ofsted: an inquiry into the future of school inspection at a glance*. Available at: <https://beyon dofsted.org.uk/reports/> (accessed 18 August 2024).
- Perryman J, Bradbury A, Calvert G, et al. (2024) A tipping point' in teacher retention and accountability: the case of inspection. *British Journal of Educational Studies* 1: 1–20. DOI: [10.1080/00071005.2024.2439791](https://doi.org/10.1080/00071005.2024.2439791).
- Phillips MJ (2024) A metamorphosis of sorts: Conceptualising women's academic identities in Australian academia. *Societal Impacts* 3: 100036.
- Powers P (2007) The Philosophical foundations of Foucaultian discourse analysis. *Critical Approaches to Discourse Analysis across Discipline* 1: 18–34.
- Quesel C, Schweinberger K and Möser G (2020) Responses to positive and negative feedback on organizational aspects of school quality: teachers' and leaders' views on a Swiss traffic light approach to school inspection. *School Effectiveness and School Improvement* 32(3): 345–362.
- Quintelier A, De Maeyer S and Vanhoof J (2020) Determinants of teachers' feedback acceptance during a school inspection visit. *School Effectiveness and School Improvement* 31(4): 529–547.
- Raby RC (2002) A tangle of discourses: Girls negotiating adolescence. *Journal of Youth Studies* 5(4): 425–448.
- Radford G and Radford M (2005) Structuralism, post-structuralism, and the library: de Saussure and Foucault. *Journal of Documentation* 61(1): 60–78.
- Richards C (2020) Viewpoint: thoughts on whole school inspection post-Covid. *London Review of Education* 18(3): 511–518.
- Roberts-Holmes G (2015) The 'datafication' of early years pedagogy: 'If the teaching is good the data should be good, and if there's bad teaching there is bad data. *Journal of Education Policy* 30(3): 302–315.
- Rosenthal L (2004) Do school inspections improve school quality? Ofsted inspections and school examination results in the UK. *Economics of Education Review* 23(2): 143–151.

- Santry C (2023) *Labour: 'Dreaded' Ofsted inspections 'ineffective'*. Available at: <https://www.tes.com/magazine/news/general/labour-dreaded-ofstedinspections-ineffective> (accessed 10 July 2024).
- Segerholm C and Hult A (2016) Learning from and Reacting to school inspection – two Swedish case Narratives. *Scandinavian Journal of Educational Research* 62(1): 125–139.
- Sims S (2016) *High-stakes Accountability and Teacher Turnover: How Do Different School Inspection Judgements Affect Teachers' Decisions to Leave Their School?* London: Department of Quantitative Social Science-UCL Institute of Education, University College London.
- Sinkinson A and Jones K (2001) The validity and reliability of Ofsted judgements of the quality of secondary mathematics initial teacher education courses. *Cambridge Journal of Education* 31(2): 221–237.
- Stevenson H (2017) The 'datafication' of teaching: can teachers speak back to the numbers? *Peabody Journal of Education* 92(4): 537–557.
- Stronach I, Corbin B, McNamara O, et al. (2002) Towards an uncertain politics of professionalism: teacher and nurse identities in flux. *Journal of Education Policy* 17(1): 109–138.
- Tian M (2024) Embracing complexity: rethinking education inspection in England. *Educational Review* 1–17. doi: [10.1080/00131911.2024.2325660](https://doi.org/10.1080/00131911.2024.2325660).
- Troman G (1996) The rise of the new professionals? The restructuring of primary teachers' work and professionalism. *British Journal of Sociology of Education* 17(4): 473–487.
- Tunç B, İnandı Y and Gündüz B (2015) Inspection and teachers' emotions: an emotional evaluation of inspection. *International Journal of Human Sciences* 12(1): 554–568.
- Vass P and Simmonds G (2001) *External review second 'updated' overview report*. London: Public Services Productivity Panel.
- Verger A and Parcerisa L (2017) A difficult relationship. Accountability policies and teachers: International evidence and key premises for future research. In: Akiba M and LeTendre G (eds) *International Handbook of Teacher Quality and Policy*. New York: Routledge, 241–254.
- Wallace M (2001) Sharing leadership of schools through teamwork: a justifiable risk? *Educational Management & Administration* 29(2): 153–167.
- Weale S (2023) *Pressure mounts on Ofsted amid outcry after death of headteacher*. Available at: <https://www.theguardian.com/education/2023/mar/21/ruth-perry-ofsted-regime-fatally-flawed-says-family-of-headteacher-who-killed-herself> (accessed 28 June 2024).
- Welsh Government (2017) *Schools Causing Concern: Statutory Guidance for Schools and Local Authorities*. Cardiff: Welsh Government.
- Welsh Government (2017) *Education in Wales: Our National Mission*. Cardiff: Welsh Government.
- Welsh Government (2023) *Criteria for the Accreditation of Initial Teacher Education Programmes in Wales*. Cardiff: Welsh Government.
- Welsh Government (2023) *Our National Mission: High Standards and Aspirations for All*. Welsh Government. Available at: <https://www.gov.wales/our-national-mission-high-standards-and-aspirations-all-html> (accessed 12 December 2024).
- Willig C (2008) *Introducing Qualitative Research Methods in Psychology: Adventures in Theory and Method*. 2nd edition. Maidenhead, UK: McGraw Hill/Open University Press.

Dr Gareth Evans is director of education policy at Yr Athrofa: Centre for Education, University of Wales Trinity Saint David. He specialises in the design, development and implementation of education policy, particularly in the Welsh context, and has written extensively on Wales' education system since devolution.