RESEARCH PAPERS IN EDUCATION
https://doi.org/10.1080/02671522.2025.2507597

Routledge
Taylor & Francis Group

Deconstructing the rationalities and epistemologies of school accountability in England

Sevda Ozsezer-Kurnuc[a] and Agnieszka Bates[b]

[a]Presidency of Inspectorate Board, Turkish Ministry of National Education, Ankara, Türkiye & School of Education, University of Glasgow, Glasgow, Scotland, UK; [b]School of Education, Bath Spa University, Bath, UK

ABSTRACT

Schools' accountability for student performance outcomes is an established policy lever in many education systems. However, the meaning of 'accountability' and the mechanisms through which schools are held to account often escape scrutiny. Based on a conceptual analysis and empirical data from expert interviews, this article identifies three problems with how accountability has been deployed as a driver of school improvement in England. First, accountability is assumed to serve the overarching purpose of school improvement through enforcement rather than justification of schools' decisions and actions. Second, accountability is aligned with the rationalities and technologies of government, which problematise schools' conduct to warrant external intervention. Third, the reductive epistemology of calculation and control, which makes schools accountable for student performance in high-stakes tests and uses punitive measures when schools do not measure up, makes accountability opaque, defeating the idea of accountability as giving a transparent account of one's actions to provide information, justify and explain. Coupled with competition, accountability is yet to deliver the significant improvement promised by successive governments. This article, therefore, argues for a reconfiguring of accountability, with implications for policy and practice in England and other education systems which rely on accountability as a driver of improvement.

ARTICLE HISTORY
Received 3 May 2024
Accepted 24 April 2025

KEYWORDS
Accountability; education; rationalities of government; technologies of government; enforcement; answerability

Introduction

Schools' accountability for student performance outcomes is taken for granted as a driver of school improvement in many education systems. Reforming education through accountability for student performance and more efficient use of taxpayers' money was established in Australia, USA and England in the 1980s (Cumming 2012). In Italy, performance management and measurement of 'added value' in national standardised tests are key elements of the National Evaluation System of schools (Milner, Mattei, and Ydesen 2021). In 2006, in response to criticism of a lack of evaluation culture, national

CONTACT Sevda Ozsezer-Kurnuc Sevda.OzsezerKurnuc@glasgow.ac.uk Presidency of Inspectorate Board, Turkish Ministry of National Education, Ankara, Türkiye & School of Education, University of Glasgow, 11 Eldon Street, Glasgow, Scotland G3 6NH, UK

tests were made mandatory in Denmark, as an 'expression of accountability mechanisms prior to their materialization in policy' (Milner and Ydesen 2024, 224). The idea of improving schools through accountability for student outcomes has been amplified by the OECD's international comparisons and rankings developed within the Programme for International Student Assessment, PISA (Breakspear 2012). Test-based accountability for student performance outcomes, combined with a focus on core curriculum subjects and corporate management models, are the essential ingredients of the Global Educational Reform Movement (Sahlberg 2011).

The central role of accountability in education reform has been associated with the spread of New Public Management (NPM) and its underlying assumption that public education can be improved through market mechanisms of competition and parental choice, business values and efficient use of resources (Norris and Kushner 2007; Wilkins et al. 2019). NPM shifted schools' accountability measures from evaluating inputs and activities, for example the number of certificated teachers (Maroy and Voisin 2017), to prioritising results and outcomes (Levin-Rozalis, Rosenstein, and Cousins 2009), rendering schools as measurable 'entities' within global comparative and competitive frameworks (Wilkins et al. 2019, 157). NPM also entails 'intensifying control' (Norris and Kushner 2007, 3) through constant monitoring of performance targets and 'league tables' (Purdue 2005, 123), within an overarching belief that emphasis on accountability is central to good governance (Maroy and Voisin 2017).

Whilst NPM is a common denominator in education reform in diverse countries, it is 'taken up and resisted or revised within different political-administrative settlements to complement existing political structures and value systems' (Wilkins et al. 2019, 151), which means that accountability regimes evolve in different ways. In Anglo-Saxon countries, accountability is deployed to promote school competition and choice (Verger and Parcerisa 2017) and these, in turn, rely on a widespread use of school rankings. By contrast, in Nordic countries and Switzerland, emphasis is placed on outcomes-based management and standards-oriented curricula (Verger and Parcerisa 2017). In some contexts, NPM has been successfully resisted. For example, attempts to establish quasi-markets and school choice in Switzerland failed because citizens voted against marketisation in education, whilst Swiss teachers and researchers rejected managerial aspects of NPM over concerns about business values in education (Wilkins et al. 2019). The evolving nature of school accountability was also evident during COVID-19 pandemic, when national tests were postponed or reconfigured during lockdowns and teachers assumed policymaking roles due to central policy inertia (Milner, Mattei, and Ydesen 2021).

In England, accountability has been associated with the Education Reform Act (ERA) 1988 and the formation in 1992, of the schools' inspectorate Ofsted (Office for Standards in Education). However, references to 'accountability' and 'holding schools to account' significantly increased in the last twenty years (Brighouse and Waters 2022). An analysis of Schools White Papers by successive governments reveals that an explicit focus on accountability started with the 2005 White Paper 'Higher Standards Better Schools for All' (DfES 2005). Although the earlier White Paper (DfE 1997) included a whole section titled 'Standards and accountability', it refers to standards and targets rather than accountability as drivers of school improvement. Later references to accountability (DfES 2005, 1) cite 'greater accountability through national testing and regular

independent inspection' and associate accountability with the introduction of the national curriculum back in 1976. The focus on accountability in Schools' White Papers increased from 2010, as seen in the frequency of references to 'accountability'. From 2016, accountability has also been described with a range of adjectives, for example: 'fair, stretching accountability'; 'robust and proportionate accountability'; 'intelligent accountability'; 'fair, robust, ambitious accountability' (DfE 2016) and 'clear accountability' (DfE 2022). References to improving accountability also appeared, e.g.: 'fairer accountability measures'; 'new accountability measures'; 'our reformed accountability system', and 'we are improving the accountability framework' (DfE 2016). The 2010 Schools White Paper supports 'formal external assessment as the basis of accountability for performance', predicated on 'strong evidence' by the OECD (DfE 2010, 67–68). However, the links between accountability and performance are not explained. The following exemplifies a lack of conceptual linkages to explain *how* holding professionals to account improves student outcomes:

> An effective accountability system ensures that professionals are held accountable for the outcomes of their decisions using fair, intelligent, reliable and carefully-balanced measures of success or failure. These measures must avoid creating perverse incentives or unduly hindering innovation. . . And they should be proportionate, giving schools and groups time to improve while reacting in time to avoid chronic failure that irredeemably damages any child's education. (DfE 2016, 21)

Despite its assumed role as a major policy lever, the meaning of educational 'accountability' remains elusive. This issue extends beyond policy texts to academic literature, as evidenced by Dubnick's (2014) characterisation of accountability as 'ambiguous', 'expansive' and 'underexplored', and Schedler's (1999) point about its 'evasive meaning', 'fuzzy boundaries' and 'confusing internal structure'. In its basic form, 'accountability' denotes an account-giving relationship, whereby an individual or an institution is obliged to explain their decisions or actions to another party (Brundrett and Rhodes 2010; Rasmussen and Zou 2014). 'accountability' can be further deconstructed into two dimensions: 'answerability' and 'enforcement' (Schedler 1999, 4). 'answerability' entails the obligation to provide reliable facts, i.e. information (the information dimension of accountability) or valid reasons, i.e. justification for one's actions (the argumentative/explanatory dimension) to ensure transparency. Answerability forms a 'benign inquiry' aimed at establishing a dialogic relationship between the power holders and those who hold power to account (Schedler 1999, 15). 'enforcement' involves rewarding desirable behaviour and punishing misconduct, thus holding individuals or institutions to account for their actions and the ensuing material consequences. The execution of sanctions is seen as the condition of effective institutions. Importantly, accountability concerns agents (those who exercise power) and not subjects (those subordinate to power), it:

> concerns subjects only as far as we ascribe some degree of freedom to them . . . it concerns public employees only as far as we envision administrative organizations not as mechanical conveyor belts of decisions from top to bottom, but rather as loci of decision making at all hierarchical levels. (Schedler 1999, 20)

As this article demonstrates, limited freedom to engage in local decision-making calls into question 'high-stakes' accountability measures imposed on schools in England, enforced by Ofsted and sanctions for schools which fail to improve. Based on

a conceptual analysis and empirical data from expert interviews, we argue for a reconfiguration of accountability to develop more complex, nuanced understandings of 'school accountability' to inform policy, practice and further empirical research.

Literature review

Significant differentiation in the use of 'accountability' in conceptual and empirical studies on accountability in education resonates with Schedler's (1999) point about its 'evasive' nature, 'fuzzy' boundaries and 'confusing internal structure'. This is reflected in the use of adjectives to define 'accountability', similar to those used in English policy texts discussed above (DfE 2016, 2022). These range from 'felt' accountability (Trask and Cowie 2022) denoting individuals' choices and responses to being held to account, in anticipation of positive or negative consequences; 'test-based' accountability (Leader and Pazey 2023) to denote high-stakes accountability as a driver of school improvement, and 'grassroots' accountability (Banghart 2021) as a bottom-up approach to addressing local and individual student needs. Scholars referring to 'accountability systems' (Leader and Pazey 2023; Moss 2022) focus on national-level policies and education systems. As Maroy and Voisin (2017), 5) explain, under NPM, 'accountability' became either a 'synonym for loosely defined political goals (such as "good governance" or "democracy")' or a term for 'mechanisms' for controlling and ensuring the quality of public services. To systematically capture its 'polysemic character', Maroy and Voisin (2017) developed a typology of accountability policies and tools, identifying distinct configurations of policy regulation through 'strong' accountability, neo-bureaucratic accountability, reflexive responsibilisation and 'soft' accountability. Their analytical categories also consider: (high, medium or low) stakes for actors involved in accountability relations; the relative strength of these relations; conceptions of actors as either utilitarian or reflexive and socially situated, and external or internal mediation for the expected change set in accountability policies.

Literature on schools' accountability also focuses on *accountability mechanisms* and *accountability measures*. Some confusion about accountability (Schedler 1999) arises in discussions of '*accountability mechanisms*', i.e. actors, institutions, structures and technologies which make accountability work in particular ways in education systems. For example, Earley and Weindling (2004, 78) identify two key actors in two types of accountability: government and its agencies (Ofsted) in 'contractual accountability' and parents in 'market accountability'. 'market accountability' frames parents as 'clients' choosing the best school for their child. Whilst 'contractual' and 'market' accountability entail that schools are accountable to Ofsted and parents-as-customers, Earley and Weindling 2004 explain that teachers are also accountable to colleagues ('professional accountability') and students ('moral accountability'). In seeing 'responsibility' to pupils and colleagues as 'accountability', Earley and Weindling 2004 thus conflate these two terms, which could be read as 'conceptual stretching' (Lindberg 2013, 205) if not misuse of the term.

Much debate on accountability focuses on *accountability measures* within results-driven forms of accountability, which also refer to 'standards', 'outcomes' and 'performance', and often use these terms interchangeably within a logic of metrics and standards (Dubnick 2014). The key *accountability measures* deployed in many education systems

are high-stakes test results (Glassow 2023; Perryman and Calvert 2020), despite mounting concerns over 'performative' accountability they engender (Perryman and Calvert 2020). For example, in the US context, standardised test data have been used to monitor teachers and 'coerce' them into 'accepting normalising judgements of their practice' (Webb 2005, 189). Within the Global Educational Reform Movement (Sahlberg 2011), the rationale for test-based accountability is often justified by disappointing PISA results, as happened in Denmark (Rasmussen and Zou 2014). In Australia and other countries, *accountability measures* in the form of high-stakes test data are central to global accountability regimes and, despite their 'perverse', 'anti-educational' effects on teachers' pedagogical practices and their students, they are so widespread that they can be seen as 'almost the 'new norm'' rather than an 'aberration' (Lingard and Sellar 2013, 634). Anti-standardisation movements have also emerged, not only in high-stakes accountability systems but in systems of 'soft' accountability such as Catalonia (Parcerisa et al. 2022), opposing the increasing use of high-stakes tests to sort students, evaluate teachers and rank schools.

Confusion around accountability also stems from the assumption that accountability serves the purpose of school improvement. Accountability has been widely advocated for enhancing students' performance (Barber 2005; Schleicher 2014), facilitating effective utilisation of public funds (Earley and Weindling 2004), and reassuring the public about school quality (Gilbert 2012). However, empirical evidence to support this assumption, particularly regarding the enforcement dimension of accountability (Schedler 1999) is weak. For example, 'accountability pressure' was found to positively influence improvement efforts of school headteachers (Altrichter and Kemethofer 2015, 32). On the other hand, fear of punishment was found to make teachers 'reproduce' surveillance, resulting in a 'paranoiac enactment' of accountability that maintained the status quo rather than taking risks to improve pedagogical practice (Webb 2005, 204). Similarly, empirical evidence linking accountability policies to improved educational equity is weak (Maroy and Voisin 2017; Verger and Parcerisa 2017). Verger and Parcerisa (2017, 28) identified a 'variegated effect' of accountability on learning outcomes, inter-school dynamics, teachers' professional identity, curriculum and inequalities. A National Foundation for Educational Research report on six jurisdictions (New South Wales, England, Japan, New Zealand, Singapore and Wales) noted a paucity of rigorous evidence on the impact of accountability on the curriculum, standards and teacher and pupil engagement (Brill et al. 2018).

There is also no consensus in the literature on whether accountability to schools' inspectorates drives school improvement (De Wolf and Janssens 2007). School inspection visits in England focus much more on accountability than in other countries, for example German school inspections aim at fostering school development (Bitan, Haep, and Steins 2014). However, evidence of improvement achieved through the mechanism of school inspections remains inconclusive (De Wolf and Janssens 2007). As De Wolf and Janssens (2007, 2020) explain, this is mainly due to the intended, 'window dressing' type of behaviour, and unintended strategic behaviour such as a myopic focus on short-term solutions, engendered by inspections. In England, whilst critics emphasise that Ofsted inspections are not 'the best system for engendering long-term improvement' (Perryman 2009, 628), advocates see Ofsted as an agency that secures improvement through holding schools to account

(Gilbert 2012), as reflected in its strapline 'improvement through inspection' (Chapman 2000, 57). In a comprehensive exploration of the nature of accountability in education in England, West, Mattei, and Roberts (2011) identify several types of accountability present in schools, and the effects of sanctions associated with each type.

Much research evidence points to adverse effects of accountability, particularly when it involves high-stakes consequences for schools. These include: diminished staff well-being (Brady and Wilson 2021); erosion of teacher solidarity and student trust in educators (Levatino et al. 2024); high staff turnover and teacher attrition (Glassow 2023; Perryman and Calvert 2020); stifled creativity and professionalism (Olivant 2015), and gaming strategies to artificially raise test outcomes (Maroy and Voisin 2017). Mandel and Pendola (2021) found that the Texas Education Agency's 8.5% cap on special education enrolment, introduced in 2004 to cut costs and improve overall performance figures, resulted in a significant, state-wide decline in special education provision. The cap was subsequently deemed illegal by the US Department of Education. Amid conflicting empirical findings on the impact of accountability *measures* and *mechanisms* on school improvement, the COVID-19 pandemic brought into sharp relief weaknesses in the rationales for existing approaches to accountability (Milner, Mattei, and Ydesen 2021; Moss 2022).

To critically examine the rationalities and epistemologies underpinning the government notion of 'school accountability' and its deployment to drive school improvement, educational experts' views were sought, guided by two questions:

- How do educational experts evaluate the deployment of school accountability in England?
- What are their recommendations for improving the accountability system?

The following section draws on the work of political scientists (Dubnick 2014; Lindberg 2009, 2013; Schedler 1999) who offer more precise understandings of 'accountability', characterised by conceptual clarity and analytical rigour, in contrast to less rigorous accounts of accountability found in some policy texts and empirical studies discussed above.

Conceptual framework

As traditional accounting expanded beyond recording and reporting financial transactions (Dillard and Vinnari 2018) to public administration and other applied fields, the concept of accountability acquired an 'undesirable semantic confusion' (Lindberg 2009, 3). It is, therefore, important to examine both its conceptual core and the rationalities underpinning the deployment of the *mechanisms* and *measures* of school accountability. At the core of accountability is a relationship whereby:

> A is accountable to B when A is obliged to inform B about A's (past or future) actions and decisions, to justify them, and to suffer punishment in the case of eventual misconduct. (Schedler 1999, 7)

Drawing on Schedler (1999) and others, Lindberg (2013, 209) identifies the following characteristics at the core of accountability:

(1) An agent or institution who is to give an account (A for agent);
(2) An area, responsibilities, or domain subject to accountability (D for domain);
(3) An agent or institution to whom A is to give account (P for principal);
(4) The right of P to require A to inform and explain/justify decisions with regard to D; and
(5) The right of P to sanction A if A fails to inform and/or explain/justify decisions with regard to D.

Of importance here are two points: first, that an institution, for example a school, is required to inform and explain/justify its decisions and actions within the answerability dimension of accountability (Schedler 1999). Second, within the enforcement dimension (Schedler 1999), the school may face material consequences including rewards for desirable behaviour and sanctions for misconduct. The use of accountability in England extends to holding schools to account 'for outcomes for every child' (DfE 2016, 21), not just the decisions which may have led to these outcomes. Based on the empirical data presented below, the overarching purpose of accountability in England is that of school improvement through enforcement rather answerability, punishing schools for low student performance in national tests, as if low performance constituted misconduct (Schedler 1999).

When applied to new contexts, accountability can undergo 'conceptual stretching' (Lindberg 2013, 205) diluting its conceptual core as methods of limiting power. In England, for example, successive governments positioned school leaders and teachers as the 'implementers' of government decisions rather than decision makers (Bates 2016), yet it is school practitioners who are held to account for student outcomes, shaped by government decisions. By contrast, policymakers appear to escape accountability for their own decisions, even though accountability to the public is one of the *Seven Principles of Public Life* (gov.uk 1995) that all government and public service officials are expected to abide by. For example, between 2010 and 2023, the Department for Education saw ten Secretaries of State for Education (Logan 2023), with ministers being moved around government departments at a pace which left little time for answerability and enforcement (Schedler 1999). This also exemplifies 'severe confusion about the core meaning of accountability' (Lindberg 2009, 4) as methods of constraining the political power of education policymakers.

To offer tools for empirical analyses of accountability in its many contexts of application, Lindberg (2013) developed sub-types, based on the following characteristics:

- Source of control (internal or external source)
- Strength of control (high or low strength)
- Direction of relationship (vertical: upward or downward or horizontal)

Lindberg (2013) uses these characteristics to identify 12 types of accountability, ranging from business, fiscal and legal to societal and political. Table 1 maps these types, across Schedler's (1999) dimensions of accountability.

Table 1. A mapping of the government approach to holding schools to account against the dimensions (Schedler 1999) and sub-types (Lindberg 2013).

		Accountability mechanism					Deployment of accountability	
		Source of control		Direction of relationship			Strength of control	
		Internal	External	Upward	Downward	Horizontal	High	Low
Accountability measures	Information dimension of *answerability*		✓	✓			✓	
Deployment of accountability	Justification dimension of *answerability*							
	Enforcement		✓	✓			✓	

To further illuminate Lindberg's (2013) types, we also draw on Dubnick's (2014) 'promises' related to the deployment of accountability: democracy, control, justice and performance, framed around four narratives. The *promise of democracy* narrative implies that accountability can constrain power by fostering the answerability and responsiveness of power-holding officials. The *promise of control* narrative involves organisational arrangements for eliciting obedience and efficiency through administrative control, bureaucratisation, reporting and auditing. The *promise of justice* narrative incorporates rules and procedures designed to deal with unacceptable behaviour. The fourth narrative rests on the role of accountability in realising a *promise of performance*, whereby accountability rests on standards and metrics for measuring performance to incentivise desirable behaviour (Dubnick 2014, 30). Table 1 maps the government approach to accountability, based on the analysis of Schools White Papers presented above (DfE 2010, 2016, 2022).

As can be seen in Table 1, the government approach rests predominantly on performance based, information-oriented accountability (to Ofsted), characterised by strong control and upward hierarchical direction. The information dimension (Schedler 1999) of school accountability relies on measures of children's progress, school performance tables and information offered to parents and school governors in 'easy-to-navigate formats' (DfE 2016, 22). It also relies on enforcement, for example 'intervention in coasting and failing schools' (DfE 2016, 15). The distinctions made in the policy discourse through adjectives referring to accountability as 'clear', 'intelligent' and 'robust', preclude debate about enhancing answerability or a focus on narratives that go beyond performance, to include the democratic promise (Dubnick 2014). Given this tinkering with adjectives and inconclusive evidence of the effects of high-stakes, performative accountability on school improvement (Lingard and Sellar 2013; Perryman 2009; Webb 2005), a question arises whether the persistence with accountability in education policy serves purposes other than school improvement.

The dual concepts of 'rationalities' and 'technologies of government' help answer this question by revealing the interactions between policymakers, experts, advocacy networks, investments, discourses and technologies through which government 'gets

done' (Lewis, Savage, and Holloway 2019, 7) and which enable 'government', i.e. governing the conduct of individuals, groups and whole populations (Savage 2019, 2). Government gets done by activating the 'rationalities of government' and 'technologies of government' (Miller and Rose 2008). The 'rationalities of government' rely on problematising individuals' or groups' conduct, i.e. render conduct problematic to justify intervention (Lewis, Savage, and Holloway 2019, 6). In this account, an accountability system relies on presenting particular phenomena as problematic and devising policy solutions to govern the population. The 'technologies of government' are assemblages of techniques, organisations, actors, instruments and processes which facilitate government (7). For example, accountability is used to problematise schools' performance to warrant the regime of Ofsted (with its 'technology' of accountability *mechanisms* and *measures*) and further policy interventions. As a 'technology of government', accountability serves the dual purpose of governing schools and governing populations. For example, for Gilbert (2012, 9), former Head of Ofsted (2006–2011), governing schools relies on the 'best school leaders' who 'feel greater ownership of accountability'. For DfE (2016, 65), owning accountability is also expected of parents who, play a 'crucial' role in the education system, 'from supporting their child to holding schools to account'. 'Owning' accountability and holding others to account are framed here as something to be espoused both in professional and family life. Accountability as a 'technology of government' is deployed to govern the conduct of education professionals, parents and other actors in the education system.

However, as Ball, Maguire, and Braun (2012, 5) emphasise, teachers and school leaders are not 'cardboard cut-out' policy implementers but complex professionals who have the capacity to disassemble and reassemble policies, individually and collectively. Faced with high-stakes accountability, they can teach to the test (Bates 2016), 'nudge' students towards subject 'choices' to improve student test scores (Maguire et al. 2020, 505) or fabricate performance for Ofsted inspectors (Perryman 2009), withholding true accounts of their successes and failures and rendering accountability opaque. It is this complexity and nuance that our study has sought to examine.

Methodology

To evaluate the deployment of accountability in the education system in England, 15 purposively selected (Cohen, Manion, and Morrison 2018) educational experts were interviewed. Each expert interview (Harvey 2011; Van Audenhove, Donders, and Puppis 2019) was between one and two hours in duration and focused on the two questions related to accountability cited above, as well as questions about approaches to school evaluation in England as part of a larger study. The experts included esteemed academic researchers, school inspectors, policymakers, policy advisers and experienced school practitioners in leadership positions (Table 2). Each participant's expertise was derived from working in education between two and five decades. Nine participants were in post in 1988, the year in which Education Reform Act (ERA) was passed. Four participants were in post in 1992, when Ofsted was established. Being the witness of, and participant in, such historical milestones gave the participants privileged knowledge and understanding. In addition to historical knowledge, their expertise extended to theoretical and practical knowledge. Four participants held professorships in UK

Table 2. Research participants and their expertise.

Category of expertise	Pseudonym	Area of expertise
Policy Actor	Richard	Inspector, policy maker- international and the UK
	Cooper	LA management roles, academic
	Dennis	HMI, policy adviser, researcher
	Gabriella	Inspector from a devolved system
	Bella	HMI, adjudicator, policy adviser
	Martin	Policy adviser, policy maker
Policy Influencer	Nora	Researcher, academic in the field of assessment
	Gordon	Researcher, specialised in alternatives to Ofsted inspection
	Eduardo	Researcher, policy adviser
	Orion	Researcher, academic in the field of evaluation
School Practitioner	Kelvin	Experienced headteacher
	Torr	Senior teacher with academic background
	Felicia	Deputy head- international and the UK
	Ned	Headteacher, lead headteacher, national leader of education
	Kent	Principal, Ofsted Inspector

universities. Of school practitioners, two had Master's degrees. Eleven experts undertook senior leadership roles, including headship, or served as school governors. Five experts took policy advisory roles and three worked for the DfE. All participants had experience in England and seven also had in-depth knowledge of other education systems (Wales, Scotland, New Zealand, Dubai, United Arab Emirates, United States, Greece, and Berlin jurisdiction in Germany).

Table 2 categorises participants based on professional experience and role at the time of interview, in the areas of: policy development ('*policy actors*'); education research ('*policy influencers*') and policy implementation ('*school practitioners*'). To ensure anonymity and eliminate the risk of prominent participants being identified, all potentially sensitive details were removed, both from Table 2 and from interview transcripts. Each participant also received their interview transcript for suggested redactions.

Interview data analysis involved inductive and deductive analytic moves (Miles, Michael Huberman, and Saldaña 2013). The inductive phase included identifying codes and themes pertaining to experts' evaluations of the deployment of school accountability in England. NVivo-12 was used for initial coding, with the initial codes subsequently grouped into two recurring themes: the necessity of accountability and problems with accountability. Since all participants acknowledged the necessity of accountability, but also referred to problems in its deployment, further coding focused on the rationales underlying this necessity, and evaluations of high-stakes, performative accountability deployed in England. The codes pertaining to the rationales were then grouped to those cited by *policy actors*, *policy influencers* and *school practitioners*, yielding a notable pattern discussed below, summarised in Table 3. The codes designating experts' evaluation of high-stakes, performative accountability and its ostensible function as a driver of school improvement also pointed to variation between the *policy actors*, *policy influencers* and *school practitioners* (Table 3). In particular, predominantly positive evaluations were articulated by *policy actors* and *school practitioners* actively involved in the work of Ofsted. Other experts evaluated accountability from a critical stance. Recommendations for improving England's accountability system were derived from a deductive analysis, based on Schedler's (1999), Lindberg's (2013) and Dubnick's (2014) distinctions.

Table 3. Accountability as a necessity and the underlying rationale.

	Rationales for the necessity of accountability
Policy actors	**External accountability taken for granted**: *if there is no external spark, it gets worse* (Martin) before accountability was introduced *schools worked really well for some proportion of the population* (Martin) *there are schools, and there always will be, that are not serving the needs of the pupils properly* (Bella) **Accountability to Ofsted**: *a good thing* (Bella, Martin) *motivator* (Kent) *a reality check* (Eduardo) **Motivation to enforcement continuum**: Ofsted … *has been one of the motivators to ensure schools do improve* (Kent) *nudging … keeping you on your toes* (Martin) *a professional obligation the schools should be hassled* (Dennis) *some teachers who are lazy … need a rocket behind them* (Bella)
Policy influencers	**Accountability important in public services and the modern world**: *accountability is needed to question people in any sector including education* (Orion) *in the modern world … everyone must be accountable* (Gordon)
School practitioners	The need for an *external eye*, but it does not have to be Ofsted (Ned)

Whilst the participant sample ($n = 15$) could be seen as a limitation, it was determined by saturation, i.e. a point at which obtaining further data to add new insights is unlikely (Cohen, Manion, and Morrison 2018). The interview data set totalled approximately 140,000 words and, given the length of each interview, offered a number of important insights (Harvey 2011). Purposive sampling, utilised to identify expert participants, may lead to recruiting individuals with similar views and experiences (Cohen, Manion, and Morrison 2018). This indeed happened in this study, prompting an effort to invite more *policy actors*, after the first few interviews generated highly critical evaluations school accountability in England. Due to difficulties with access, *policy influencers* working in non-governmental organisations and unions did not participate in this research. Another limitation of the key findings presented below pertains to a loss of more detailed analyses of accountability by each participant, due to the requested redactions and word allowance constraints.

Experts' evaluations of the deployment of school accountability in England

The inductive analysis of interview data revealed that all experts saw accountability as a necessity. However, their evaluations of high-stakes performative accountability and its impact on school improvement diverged.

The necessity of accountability

All participants articulated the necessity of accountability in public education. For example, Orion emphasised that there is no 'desirable' alternative to accountability. Ned explained that schools need an 'external eye' to improve. For Gordon, accountability is a 'modern-day phenomenon', whereby 'everyone must be accountable' (Table 3).

The most notable finding was the conflation of accountability with Ofsted inspections by some *policy actors*. Ofsted inspections were justified as a 'reality check' and 'motivator' to do the 'right things' or in a punitive sense, as a means to bring people 'into line' through fear:

> . . . there's always been an inspector of sorts, but not on the scale that we have now. I think, when inspectorate was adopted, there used to be a sense of fear about it. And people felt frightened by it, and it did bring people into line. And I don't think that's a bad thing. (Kent)

Unlike Kent, who appeared to align with the enforcement dimension of accountability (Schedler 1999), Torr was critical of the punitive approach, as an 'aggressive, hard-nosed approach to managing people' which does not allow answerability:

> You're accountable for this and if it doesn't work, if you don't do what you're supposed to have done, we will sack you' . . . That's not me being given an opportunity as a classroom teacher to say: 'Well, actually, it didn't work the way you wanted it to. Because of all of these other reasons. (Torr)

Gordon's explanation of the necessity of accountability was accompanied by important epistemological considerations:

> Accountability is a good thing. Everyone must be accountable . . . So, the first reason, education is highly important. And second, it's very expensive. And MPs have a right to challenge teachers and say: 'You're getting millions from government. What are you doing with it?' . . . But the question then becomes what kind of accountability? What form should it take? How accurate is it? How valid? How reliable? (Gordon)

As discussed further below, for Gordon, a 'real account' of schools would include contextual data as vital to understanding student performance, within a combination

Table 4. Participants' evaluations of government approach to accountability.

	A predominantly positive evaluation	A critical evaluation
High-stakes, performative accountability	• necessary though it may create 'distortions' (Eduardo, Martin), lead to 'issues of gaming, unintended consequences' and decrease the quality of information gathered (Eduardo) • 'a double-edged sword' but he is 'confident of what we're doing' (Kent) • 'excessive focus on data' (Martin) • the above issues have been resolved through a rebalancing by Ofsted (Martin)	• creates 'perverse' incentives (Ned, Felicia, Kelvin, Cooper, Ned, Torr, Dennis, Gordon, Kelvin, Nora, Felicia, Orion) • schools can 'game the system' e.g. off-rolling pupils, 'all sorts of ways of boosting your scores by unprofessional practices (Gordon) • exodus of teachers (Nora, Cooper) and headteachers (Kelvin), closure of schools (Gordon) • proliferation of school inspections by Ofsted, DfE, MAT, RSC (Kelvin) • stifles more 'creative' approaches (Nora)
Accountability as a driver of school improvement	• being accountable to 'somebody external' is important, especially in the case of inadequate schools (Martin) • holding schools to account decreased the number of 'really bad' schools (Kent) • 'accountability means kids are doing better' (Kent) • in a system without accountability, practice becomes 'wishy-washy' and 'mediocre' (Bella)	• negative 'aggregate impact' of accountability on school improvement (Orion) • accountability 'counterproductive' to improvement, due to the narrowing of the curriculum and excessive levels of stress (Kelvin) • money and energy put into 'displacement' activities (Orion, Kelvin, Nora)

of horizontal and vertical accountability involving external inspection and self-evaluation respectively.

Problems with high-stakes, performative accountability

Despite a consensus around the necessity of accountability, the participants identified a range of problems with high-stakes accountability predicated on student performance outcomes (Table 4).

It can be expected that experts working for Ofsted are likely to evaluate the Government's approach in positive terms. However, these participants also highlighted some unintended consequences arising from high-stakes, performative accountability. For example, Eduardo and Martin acknowledged that high-stakes accountability may 'create distortions' and 'issues of gaming', because 'schools will naturally try and do whatever they can to get the best possible outcomes in terms of accountability' (Eduardo). However, high stakes go together with high autonomy:

> England has a high-stakes accountability system . . . but also has a great deal of freedom for schools alongside that as well. So, headteachers for example can take decisions about things that for a lot of other countries would be unthinkable like budgets, hiring teachers, tweaking curriculum etc. (Eduardo)

Kent described accountability as 'a double-edged sword' because 'some' schools are 'ditching subjects' which are not part of high-stakes accountability. However, he emphasised his confidence in Ofsted inspections because:

> . . . those persons whose inspection experience has been positive, you're going to get a reasonable response . . . And the experience of being inspected is nerve wracking, it's frightening . . . But you have to get over that a little bit . . . I think, we all big enough to cope with that. (Kent)

Similarly, Bella associated inspection-related stress with headteachers' reluctance to reveal weaknesses. Since Ofsted inspectors are 'objective', Ofsted should not be 'perceived as beating the schools with a big stick'; a 'sensible school embraces Ofsted' and 'uses it as a tool for their own improvement'.

Experts critical of performative accountability cited a range of negative consequences. For example, Ned referred to how 'people fixate on certain performance criteria' and 'will do everything they can to get up the league table', in the interest of the school rather than the interest of the child. Felicia claimed that 'as long as accountability exists in terms of exam results, you'll never get rid of the cheating'. Ned also noted the negative impact of high-stakes tests on students' learning and wellbeing:

> . . . the process we currently have is provoking mental health issues, is causing young people just to fixate on short term memory, is insisting that they are put through extreme pressure at the age of 16 and younger as they do mock exams, and do tests, and do a whole barrage of different examinations. (Ned)

Cooper described the current accountability system as 'pernicious' and 'wrongheaded':

> People do not have the courage to step outside the framework of the national curriculum. They don't do things that schools used to do like prepare children for society. They tend to do those things which will result in a successful Ofsted inspection next time. Because that's

> the accountability system. So, it's pernicious. It is wrongheaded. It is narrowing. It is not education. . . (Cooper)

Does accountability deliver school improvement?

Whilst Schools White Papers (DfE 2010, 2016, 2022) do not explain precisely how accountability could lead to school improvement, participants presented coherent explanations of both its positive and negative effects on improvement. Martin explained that being accountable to 'somebody external' is important because 'knowing that you'll be asked . . . keeps you on your toes':

> Lots of school improvement can happen without accountability, lots of it. But we all need somebody external to us whom we hate to make us get out of bed in the morning or to make us learn something that we don't want to learn, or to make us make a change that we don't want to make. . . There are four percent of schools out there where I would not want my child within a country mile of that school . . . Somebody needs to see it, and somebody needs to say it. And that's what accountability does . . . (Martin)

Kent argued that holding schools to account since 1988 decreased the number of 'really bad' schools:

> Before 1988 . . . there were some really bad schools, really bad. And I'm not saying there aren't still some really bad schools. But there are far, far fewer now than where there were . . . You know, before 1988, there was nothing. I mean there was nothing. There was obviously nothing. There was nothing to make sure schools were doing the job . . . You know, nobody was accountable. . . (Kent)

Eduardo considered accountability to be an effective driver of school improvement because of its function to make schools aware of their own weaknesses:

> . . . very often it is helpful to have an external neutral party taking a look at that from national perspective . . . in any organisation, you will be sometimes blind to some of your own weaknesses. (Eduardo)

Experts with opposing views argued that in exceptional individual cases, accountability 'may' have contributed to improvement (Cooper), however, the 'aggregate' impact was 'negative' (Orion) and 'counterproductive' (Kelvin). This was due to negative consequences such as: narrowing of the curriculum; diverting schools from their own, context-based needs, and causing excessive levels of stress. Kelvin described his experience as headteacher as 'being on a knife edge'. During her school visits in the 1990s, Nora found it 'very strange' that the role of educators was displaced with activities such as finding evidence of the school meeting accountability demands. Similarly, Orion pointed to the 'money and energy' put into 'displacement' activity:

> . . . does accountability drive up standards overall? No. No, in fact it probably has slightly the reverse impact. It's because it takes money and energy and puts it into a displacement activity . . . there are some instances where accountability has indeed meant that the institution has to improve and has been played a fairly vital role in establishing the basis for improvement. But that's quite different from what the aggregate impact is . . . I don't think [accountability] has played a significant role in improving the quality of education, rather the reverse. (Orion)

Orion also challenged the unfair distribution of responsibility for school improvement between policymakers and schools:

> . . . it is deeply unfair that schools get held to account for performance when the resourcing of education plays a significant part in the results that schools are able to achieve. And I don't just mean the money that the school itself gets. But we know, as a matter of fact that significant investment in early childhood education and preschool interventions has a lasting impact on the wellbeing including the educational performance of young people. So, what governments do in preschool and post-natal care and the like has an impact on what schools are able to do later on. So, you know, schools do not bear the full responsibility for underperformance . . . What I think is wrong is that we can't apportion blame correctly. So, it's the school that gets blamed . . . And the political and contextual variables that make a contribution to the success or failure of a school are not part of this school report. They're not part of the story that's told about the school. That's worrying because it's a partial story. (Orion)

Cooper emphasised that whilst accountability is a good idea, 'it's not the idea to produce school improvement'. This is because teachers do not need to be 'driven' to improve:

> Of course, any publicly funded system should be accountable to the public. But the question is 'How? Through what mechanism should it be accountable?' It is public money. So, what we have introduced in the name of accountability is inspection, performance tables and, alongside that, is parental choice which it isn't . . . I've no doubt that being accountable has an effect on schools. Schools are more alert, more aware than they were about the need to continuously improve. But [external accountability] presumes that professionals don't care . . . My view is that ninety nine percent of people who teach want to do a good job for children. So, they're already self-improving . . . They don't have to be driven . . . (Cooper)

Cooper's belief in the value of self-improvement as an intrinsic characteristic of teachers contrasted Bella's assertion that some teachers are 'lazy . . . need a rocket behind them', Martin's reference to inspectors 'keeping you on your toes' and Kent's assumption that 'we all need somebody external . . . to make us get out of bed in the morning'. These contrasting views resonate with Maroy and Voisin's (2017) analytical distinction about actors involved in accountability relationships as either utilitarian, and in need of 'strong' accountability, or reflexive and responsible and therefore responsive to 'soft' accountability.

Accountability as a political tool

The participants, who did not consider that accountability led to school improvement, saw it as a political 'tool' to enforce policymakers' agendas. Nora and Cooper pointed out that holding schools to account for exam results and position in school league tables appealed to middle class voters. Kelvin and Gordon noted that competition between schools extended marketisation and academisation. Academisation denotes a government policy since 2010, transferring schools from Local Authority control to create 'academy schools', directly accountable to the Department for Education (Bates, Choi, and Kim 2021). As Kelvin explained, accountability:

> . . . has become a tool that is used by various policymakers one way or another to force a particular agenda. And, particularly in this part of the world, there was a spate of bad

> inspections in 2012 to 2013 which led to a lot of those schools that were graded as 'inadequate' joining the academy system, which was of course government policy. (Kelvin)

Seeing school accountability as a political 'tool' opens up a perspective that goes beyond performative accountability as a mechanism for 'keeping you on your toes' (Martin) and sees through its deployment for purposes other than the Government's stated aim of driving school improvement. It can also offer insights into the rationalities of government (Miller and Rose 2008), as implied in Kelvin's point that the 'spate of bad inspections' problematised schools' outcomes to justify intervention in the form of academisation. A forced conversion of schools judged by Ofsted as 'inadequate' was at the time a flagship government policy (Bates, Choi, and Kim 2021). The technologies and rationalities of government underpinning accountability in England are further explored in the Discussion section below.

Experts' recommendations for improving the accountability system

While all participants acknowledged the necessity of accountability, their positions on how it could be better deployed in the English education system differed. Based on a deductive data analysis informed by Schedler's (1999), Lindberg's (2013) and Dubnick's (2014) conceptual distinctions, this section foregrounds recommendations for reconfiguring accountability that go beyond improving its information dimension (Schedler 1999). *Policy actors* recommended an improved cycle of inspections for all schools; financial resources for more comprehensive Ofsted inspection visits, and more detailed data on students' and parents' views, with vertical accountability relationships seen as essential, within an overarching aim of maintaining a (slightly improved) status quo. By contrast, *policy influencers* and *school practitioners* recommended a more radical reconfiguring, characterised by a departure from the government approach (Table 1) to underpin accountability by horizontal relationships (Lindberg 2013), the answerability dimension (Schedler 1999), and a democratic promise of accountability (Dubnick 2014). Their recommendations reflected a more complex epistemology, grounded in an understanding of the situated nature of knowledge, and accompanied by calls to enrich the reductive epistemology of calculation and control with contextualised knowledge for a full 'story' (Orion) and 'real account' of schools (Gordon).

Reconfiguring control

Participants who acknowledged the importance of an 'external eye' (Ned) emphasised that this did not necessarily denote Ofsted. Torr recommended a 'radical change' to the current 'hierarchical' structure, suggesting that teachers should be accountable to their students, not to their managers. Conflating accountability with responsibility similar to Earley and Weindling (2004), Ned explained:

> I am accountable to the children and the parents who I serve. So, if a child is unhappy, they come and see me. If a parent is unhappy, they come and see me. And they tell me what the problem is, and together we try and work through to get a solution … (Ned)

In terms of external accountability, Ned advocated a combination of accountability through peer review and to Regional Schools Commissioners in individual schools 'where the peer review process exposes fundamental weaknesses'. Regional Schools Commissioners were first appointed in 2014 to preside over academies in eight geographical regions across England, with their role later extended to an oversight of all schools in their region (House of Commons 2016). The creation of this role within the DfE coincided with the phasing out of Local Authority control over education.

Kelvin and Gordon advocated for area-based inspection, inspecting geographically close partner schools as a whole rather than individually. For Cooper, to secure improvement, accountability should be to school governors and the local community:

> . . . schools operate in a particular context. So, if I look at the schools around the corner here, I have to think, 'Well what would be its accountability framework?' Then, it might be different from a school 20 miles away. So, it should be locally brokered, it should be by agreement. (Cooper)

If Ofsted were to retain their position in the current accountability mechanism, then their role should be 'modified', by acting as a 'professional adviser' (Cooper) or 'professional colleague' (Gordon):

> Ofsted comes in to find out what they think is wrong. So, what the teachers in England say is: 'Inspection is done to us. It is not done with us.' I would be in favour of a move towards a partnership between the teaching profession and Ofsted to do this together. (Gordon)

Importantly, Cooper argued that there should be a clear distinction between accountability and inspection, because 'inspection is not accountability'.

From vertical to horizontal accountability

Gordon described a reconfigured approach as a combination of 'horizontal' and 'vertical' accountability:

> With vertical accountability, we mean something like Ofsted where people from above come in and tell you what to do or tell you what's wrong. With horizontal accountability, this is self-evaluation, this is peer evaluation. Other headteachers from other schools coming in and helping you in a very rigorous way to improve, telling you what's right and what's wrong . . . I also think that we need to have local inspectors, not just national inspectors. Local inspectors who are more likely to know the context within a particular school, who know the teachers on a regular basis, who are in and out of the school much more often . . . You cannot have that in a national system. I mean, there're 25,000 schools in England. You cannot be an expert in all of them simultaneously with a force of 1,500 inspectors. . . (Gordon)

A recommendation to reconfigure accountability as a 'shared endeavour' (Nora) entailed 're-setting the relationship between Ofsted and the teaching profession through two-way conversations with schools and teachers' (Dennis). Accountability went beyond 'being purely school based' (Kelvin), with all stakeholders informed about the strengths and weaknesses of policies that impact schools by a 'properly independent inspectorate' (Orion). For Orion, the impact of education policies and funding for schools should be taken into account to transcend the current 'unfair' approach whereby only schools 'get blamed' for the negative impact of government policies.

From epistemology of calculation to more sensitive accountability systems

Recommendations for a shift from the current reliance on vertical, external accountability were accompanied by a recognition of more complex epistemologies. In this regard, *policy influencers*' and *school practitioners*' views converged on the vital importance of the social, financial, and political contexts within which schools operate, leading to creating 'more sensitive systems of accountability' that are 'more attuned' to their own complexities (Orion). Gordon argued that a contextualised understanding of accountability could help 'treat schools in challenging areas equitably' whilst Felicia believed it would allow schools to concentrate on their own needs and improvement strategies. Orion emphasised that contextual variables are important in understanding schools' performance, both success and failure.

Arguments for more complex *accountability measures*, beyond student performance in high-stakes tests, were accompanied by recommendations to focus on diverse educational outcomes, purposes and values. For example, for Dennis accountability should 'look at the whole life of the school and the qualities and attitudes and skills developed'. Orion emphasised the quality of educational relationships and resources, children's wellbeing, mental health indicators, exclusions, and other factors. Kelvin emphasised 'wider aspects of education' which offer every child 'a responsibility experience, work experience, sporting experience, team experience'. For Ned, accountability should be looking at 'the long-term success of young people', 'personal development' and 'growth' of students, specifically, happiness, feeling confident, feeling positive about the future and the relationships with others, being articulate and self-assured, 'feeling that they have a future and a place in our world'. Going beyond its information dimension (Schedler 1999) thus enabled the accountability system to protect different values.

These recommendations arose from critiques of Ofsted's methodology, which reduces a school's performance to a single-word judgement: 'outstanding'; 'good'; 'requires improvement', and 'inadequate'. The key recommendation pertained to a stronger emphasis on qualitative data, because 'the qualitative stories behind the quantitative data' explain 'far more' than quantitative data (Felicia). Dennis emphasised that any inspection findings are 'subjective, contestable, and never definitive' and recommended a departure from the current approach as a 'judgemental, arbitrary and a fault-finding accounting process' towards inspection as 'developmental', 'educational' and 'principled'. A developmental approach would enable schools to learn from their 'mistakes' without 'being afraid' (Nora) and afford teachers an opportunity to explain and justify the reasons 'when something isn't as expected' (Torr). Orion recommended developmental forms of inspection and evaluation of schools triggered when the 'alarm bells go' rather than as a matter of routine:

> . . . you want inspection to be protective of the rights of the child as it were, you want inspection to stand up for the quality of education for the child or the relative like wellbeing, safety, access to high quality education and to only take place when those alarm bells go . . . The external evaluation should take account of the full range of quantitative evidence available on the school and should collect direct evidence of contextual variables, school processes and the milieu of the school, its atmosphere and the sociability and things of that kind . . . it should pay attention to stakeholders . . . to the political and contextual variables that are important in understanding why the school performs as it does. (Orion)

Orion's reference to the importance of understanding the 'why' of school performance also recognises the importance of the justification dimension of accountability (Schedler 1999).

Importantly, experts highlighted that accountability should not be equated with Ofsted inspections and a national inspectorate should address the basic governmental need to check that taxpayers' money is well spent by schools, as well as ensure the safety of students. This represented a shift from performative accountability (Perryman and Calvert 2020), seen by these experts as problematic due to narrowing down broader educational purposes and creating 'perverse' incentives (Lingard and Sellar 2013), to align accountability more to its conceptual core (Lindberg 2013) as schools meeting their obligation to inform and explain/justify their decisions about a range of actions: from efficiently spending taxpayers money and putting the learning and wellbeing of students at the heart of decision-making to becoming more democratic as institutions. A democratic focus on horizontal relationships within the school and the local community was emphasised by Nora. Cooper pointed to the 'fundamental mistake' of the current system, dominated by league tables and competition, of thinking that schools are 'separate', even though:

> . . . in a system of public education . . . we shouldn't be interested in school improvement for a school. We should be interested in school improvement for all schools . . . For the quality of public education as a whole, that's the only thing we should be interested in, not of some schools at the expense of other schools. . . (Cooper)

Accountability in a system of public education was thus envisaged as 'public accountability', in the sense of schools being accountable to the public for decisions and actions which serve the interests of children, parents and society.

Our key focus in this section has been on more radical recommendations, due to the problems with the high-stakes, performative forms of accountability reported by the research participants and the literature (Brady and Wilson 2021; Lingard and Sellar 2013; Olivant 2015; Perryman and Calvert 2020; Webb 2005). The Government's approach to school accountability (Table 1) rests on external vertical accountability, focused mainly on the information dimension of answerability and deployed through enforcement (Schedler 1999). In this approach, accountability is assumed to be a driver of school improvement, measured through high-stakes test results and Ofsted grades. Based on experts' recommendations, a reconfigured approach rests on horizontal relationships that enable the justification dimension of answerability (Schedler 1999), as a foundation of accountability that privileges full transparent accounts over a 'partial story' (Orion) and 'locally brokered' measures of educational quality over centralised, reductive *accountability measures* (Cooper). As such, reconfigured accountability may accomplish its democratic promise (Dubnick 2014) which implies that particular arrangements within an accountability system are likely to result in a greater degree of democratic governance of schools. In this regard, Dubnick (2014, 32) points out that accountability as a 'distinct form of behavior and social relationships' can be as much a manifestation of 'authoritarian hierarchies' as that of 'democratic regimes'. In a system of public education in a democratic country, a 'reformed accountability system' (DfE 2016) would be developed through democratic, i.e. informed, horizontal and inclusive, debate about how schools should be governed, focused on the 'quality of public education as a whole'

rather than on raising the stakes to improve the 'single' school, at the expense of other schools (Cooper).

Discussion

Participants' evaluations of the deployment of accountability in England revealed that framing the purpose of accountability in terms of school improvement is complex and contested. Accountability may have improved educational standards in 'some instances' but its 'aggregate impact' on schools in England has been negative (Orion), as evidenced by numerous negative consequences of high-stakes, performative accountability reported both by the participants and empirical studies (Brady and Wilson 2021; Glassow 2023; Mandel and Pendola 2021; Olivant 2015; Perryman and Calvert 2020). As Cooper observed, whilst accountability is a 'good idea', 'it's not the idea to produce school improvement'. The claim that accountability drives school improvement involves a conflation of means and ends: the stated goal of accountability is to improve schools but accountability is also an end in itself. As Maroy and Voisin (2017) point out, as a centrepiece of education reforms in many countries over the last three decades, accountability has become both an 'instrument' and a 'goal'.

However, to effectively deploy accountability for its democratic purpose of fostering answerability in the delivery of public services (Dubnick 2014; Schedler 1999), the means-ends logic needs to be disentangled. The dimensions of accountability presented in Table 1 may support a more robust scrutiny of its deployment in education. Working with the conceptual core of accountability may also address semantic 'confusion' (Lindberg 2013), such as conflating accountability with Ofsted inspections reported above. To complement conceptual work on 'accountability', defining 'school improvement' from diverse stakeholders' perspectives is also important. School improvement is predicated on complex factors and should be focused on outcomes beyond student performance, such as students' 'qualities, attitudes and skills' (Dennis), as well as their sense of belonging and a 'feeling that they have a future and a place in our world' (Ned).

It is clear from the data presented above that critical evaluations of the deployment of accountability in England provide a strong rationale for its radical reconfiguration. However, 'radical' change was recommended in the specific English context of a high-stakes, performative accountability system. Some forms of accountability recommended by the participants have already been deployed in other countries and are therefore 'radical' only in the English context. For example, a combination of vertical and horizontal accountability suggested by Gordon is deployed in Ontario and Scotland (Maroy and Voisin 2017). Similarly, Cooper's and Orion's advocacy for 'soft' accountability resonates with the system deployed in Belgium and France (Maroy and Voisin 2017). The implication of such alternative models for policymakers in England is the need to learn from other education systems. It is also important to consider how the same agenda of New Public Management has been taken up differently in different education systems (Wilkins et al. 2019).

If the 'aggregate' impact (Orion) of accountability on improving the system, as opposed to improving outcomes for the 'single school' (Cooper) is questionable, then it is also important to consider whether its continued deployment has a different rationale,

as a technology of government useful in controlling schools through enforcement (Schedler 1999). As discussed above, the 'rationalities' and 'technologies of government' use a style of thinking which renders reality amenable to calculation, leading to the privileging of reductive epistemologies of calculation and control (Miller and Rose 2008). In this regard, Ofsted privilege quantitative metrics of student performance to evaluate schools' performance and use sanctions on schools judged as 'inadequate', e.g. further frequent inspection, conversion of schools to academies and replacing school leadership, to name a few (Perryman 2009; West, Mattei, and Roberts 2011). However, without the full information and justification-giving function of accountability-as-answerability, accountability-as-enforcement obscures both the complex contextual information and decision-making processes at school and policy levels which affect school improvement.

And yet, the overwhelming consensus amongst the participants on the necessity of accountability, captured by Gordon's point that 'in the modern world . . . everyone must be accountable', resonates with Dubnick's (2014) analysis of accountability as not just managerial mechanisms but also a widespread 'cultural phenomenon'. Accountability as a cultural phenomenon is a 'moral force which can be – and often is – used to promote and foster the application of compliance mechanisms and instruments of change' (Dubnick 2014, 33). This moral aspect underpins the UK Government's *Seven Principles of Public Life*, with accountability as a principle guiding the delivery of public services, together with: selflessness, integrity, objectivity, openness, honesty and leadership (gov.uk 1995). In the empirical data, a polarity emerged between a view supportive of the moral 'push' of accountability that ensures that schools do a good job and a view that the moral 'pull' to deliver quality education comes from within. The former was captured by Kent's emphatic statement that, 'before accountability', there was 'nothing to make sure schools were doing the job'. The latter was articulated by Cooper, who pointed to a wrong assumption underpinning accountability in England that 'professionals don't care', whereas most teachers 'want to do a good job for children' and are 'already self-improving'. This polarity highlights a fundamental question of teacher professionalism, echoing Maroy and Voisin's (2017) analytical distinction between the nature of social actors as either utilitarian and in need to be 'managed' or reflexive and socially situated within the relationships of 'soft' accountability.

Conclusions

From its traditional use as method of reporting financial transactions, 'accounting' became 'accountability' and spread to diverse applied contexts to realise different 'promises', performative and administrative, as well as democratic (Dubnick 2014), through a range of vertical or horizontal relations and internal or external sources of control (Lindberg 2013). As these distinctions indicate, school accountability involves a complex web of relationships, assumptions, rationalities and epistemologies which underpin how schools can, or should, be held to account for education they offer to their students. And yet, when deployed by policymakers in England, accountability remains ill-defined and its mechanisms and measures are presented as given and beyond scrutiny. Whilst adjectives to qualify accountability as 'intelligent', 'robust' and 'clear' proliferate, policy texts fail to fully justify the rationalities and epistemologies underpinning the government framing of school accountability as a driver of school improvement. This study has,

therefore, sought educational experts' evaluations of the deployment of school accountability, as well as their recommendations for improving the accountability system in England.

Despite the consensus about the necessity of accountability, empirical data from interviews with educational experts revealed the deployment of school accountability in the education system in England as contested and complex. The data pointed to a polarisation of answers to questions whether accountability is a driver of school improvement and how the current accountability system could be improved. Experts working in the roles of *policy actors* aligned their views to the Government's approach and suggested minor improvements to the information dimension of accountability (Schedler 1999). Experts who critically scrutinised the current system emphasised that equating accountability with Ofsted inspection is problematic, as is the mechanism of external, vertical accountability (Lindberg 2013) and its underlying assumption that schools need an 'external eye' to improve. They were also critical of the deployment of accountability as a driver of school improvement. In their view, the purpose of accountability is for schools as public institutions to account to students, parents, governmental agencies and the wider society, in alignment with Dubnick's (2014) point about fostering the answerability and responsiveness of schools within the *promise of democracy* narrative.

Another problem reported by these experts was *accountability measures*, which create 'perverse' incentives for schools to prioritise performance outcomes at the cost of broader educational purposes such as student wellbeing, personal development and social skills. Whereas accountability may have contributed to improvement in some schools, its 'aggregate' impact was assessed as 'negative'. Due to its preoccupation with enforcement (Schedler 1999) and sanctions for schools which do not measure up, instead of engendering transparency, the current accountability system renders 'real' accounts of schools' successes and failures opaque. It also obscures the policymakers' role in school improvement, 'unfairly' apportioning blame for the failure of policies to schools alone.

Diverse recommendations were cited for a reconfiguration of *accountability mechanisms*, pertaining mainly to a departure from the hierarchical structure and external accountability in favour of 'locally brokered' accountability, with an independent inspectorate taking on the role of 'professional advisers', thus prioritising answerability (Schedler 1999) when holding schools to account. Recommendations for a reconfiguration of *accountability measures* referred to a broader range of outcomes and the democratic presence of a school within its local community, meeting its obligation to justify its decisions about a range of actions, from efficiently spending taxpayers' money to putting the wellbeing and learning of students at the heart of its endeavours. These recommendations were made from an understanding that schools are unique and operate in different social, political, and financial contexts. A contextualised understanding of school accountability would help schools to use their autonomy to decide for themselves what should be improved, even if this might entail rejecting the current definition of 'improvement' based on better test results and higher Ofsted grades.

The conceptual and empirical analysis presented in this article points both to alternative mechanisms for holding schools to account and alternative accounts that schools could offer within a reconfigured accountability system. Transparent, 'real' accounts of schools in a democracy are predicated on a shift from the paradigm of

calculation and control towards situated knowledge, co-constructed in specific local contexts. They foreground social justice, seek to address systemic inequities and promote equal access to quality education rather than a market approach in which parents-as-consumers choose the best school for their children. If accountability is a necessity, then to give currency to the ideas of 'intelligent' and 'fair' accountability (DfE 2016), policymakers and school practitioners need to acknowledge the limits to the 'rationalities' and 'technologies of government' (Miller and Rose 2008) deployed in the current accountability system. School practitioners may need to become more critical about the *measures* and *mechanisms* through which they are held to account, based on conceptual insights and distinctions offered by scholars such as Schedler (1999), Lindberg (2013), Dubnick (2014) and Maroy and Voisin (2017). Policymakers need to make decisions from a reflective stance in which the 'rationalities' and 'technologies of government' are acknowledged and curtailed and be more open to being held to account themselves, in line with the *Seven Principles of Public Life* (gov.uk 1995), from an understanding that, in its core, accountability refers to methods of limiting (their own) power (Lindberg 2013). Debates involving policymakers, practitioners, researchers and other educational experts would also be in order, about who gets to define what 'school accountability' means and how it could be deployed in education as a more effective approach, sensitive to the mechanisms, measures, people and contexts. Orion's point about a genie out of the bottle captures the above considerations:

> . . . we can't go back to a situation where schools are not really accountable, or accountability doesn't matter. That particular genie is out of the bottle. . . So, it used to be the case that by and large the work of doctors was not questioned. Now that isn't true. No, I can't see us going back to a situation where we don't question doctors. So, can we see ourselves in a situation where accountability becomes less important? I can't see it myself . . . So, I think, what we need is more sensitive systems of accountability that are more attuned to the costs and benefits, and the strengths and weaknesses, and the threats and opportunities that accountability mechanisms come with. (Orion)

On the account of educational experts who participated in this research, a reconfigured, sensitive accountability system could go beyond its present use as a technology of government (Miller and Rose 2008) to hold both schools and policymakers to account for a wider range of outcomes and to enable more democratic approaches to improving education in diverse local communities.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work was supported by the Milli Eğitim Bakanliği. The first author's PhD research was funded by the Turkish Ministry of National Education. This co-authored article has not been directly funded by any organisation.

Notes on contributors

Sevda Ozsezer-Kurnuc is a policy maker at the Ministry of National Education Inspectorate Board, and an affiliated research associate at the University of Glasgow, where she previously worked as a postdoctoral researcher. She specialises in accountability, inspection, and evaluation policies for schools and teachers.

Agnieszka Bates is Head of the School of Education at Bath Spa University. Her research focuses education policy and practice, informed by a range of theoretical perspectives.

Ethical considerations

Data was collected after the approval of University of East Anglia Ethics Committee.

References

Altrichter, H., and D. Kemethofer. 2015. "Does Accountability Pressure Through School Inspections Promote School Improvement?" *School Effectiveness and School Improvement* 26 (1): 32–56. https://doi.org/10.1080/09243453.2014.927369.

Ball, S. J., M. Maguire, and A. Braun. 2012. *How Schools Do Policy: Policy Enactments in Secondary Schools*. 1st ed. London and New York: Routledge. https://eric.ed.gov/?id=ED527845.

Banghart, K. 2021. "Colorado's Network for Local Accountability." *State Education Standard* 21 (1): 25–28. https://eric.ed.gov/?id=EJ1286833.

Barber, M. 2005. "The Virtue of Accountability: System Redesign, Inspection, and Incentives in the Era of Informed Professionalism." *Journal of Education* 185 (1): 7–38. https://doi.org/10.1177/002205740518500102.

Bates, A. 2016. *Transforming Education: Meanings, Myths and Complexity*. Abingdon: Routledge. https://doi.org/10.4324/9781315687346.

Bates, A., T.-H. Choi, and Y. Kim. 2021. "Outsourcing Education Services in South Korea, England and Hong Kong: A Discursive Institutionalist Analysis." *Compare: A Journal of Comparative and International Education* 51 (2): 259–277. https://doi.org/10.1080/03057925.2019.1614431.

Bitan, K., A. Haep, and G. Steins. 2014. "School Inspections Still in Dispute–An Exploratory Study of School Principals' Perceptions of School Inspections." *International Journal of Leadership in Education* 18 (4): 418–439. https://doi.org/10.1080/13603124.2014.958199.

Brady, J., and E. Wilson. 2021. "Teacher Wellbeing in England: Teacher Responses to School-Level Initiatives." *Cambridge Journal of Education* 51 (1): 45–63. https://doi.org/10.1080/0305764X.2020.1775789.

Breakspear, S. 2012. "The Policy Impact of PISA: An Exploration of the Normative Effects of International Benchmarking in School System Performance." *OECD Education Working Papers, 71*. Paris: OECD Publishing. https://doi.org/10.1787/5k9fdfqffr28-en.

Brighouse, T., and M. Waters. 2022. *About Our Schools: Improving on Previous Best*. United Kingdom: Crown House Publishing Ltd.

Brill, F., H. Grayson, L. Kuhn, and S. O'Donnell. 2018. *What Impact Does Accountability Have on Curriculum, Standards and Engagement in Education? A Literature Review*. Slough: National Foundation for Educational Research.

Brundrett, M., and C. Rhodes. 2010. *Leadership for Quality and Accountability in Education*. USA and Canada: Routledge. https://doi.org/10.4324/9780203852934.

Chapman, C. 2000. "Improvement, Inspection and Self-Review." *Improving Schools* 3 (2): 57–63. https://doi.org/10.1177/136548020000300213.

Cohen, L., L. Manion, and K. Morrison. 2018. *Research Methods in Education*. 8th ed. London and New York: Routledge.

Cumming, J. J. 2012. "Educational Reform and Educational Accountability Legislation and Policy in the US, England and Australia." In *Valuing Students with Impairment: International*

Comparisons of Practice in Educational Accountability, edited by J. J. Cumming, 11–23. Dordrecht, Heidelberg, London, New York: Springer.
De Wolf, I. F., and F. J. G. Janssens. 2007. "Effects and Side Effects of Inspections and Accountability in Education: An Overview of Empirical Studies." *Oxford Review of Education* 33 (3): 379–396. https://doi.org/10.1080/03054980701366207.
DfE. 1997. "Excellence in Schools." London: Her Majesty's Stationery Office. https://www.education-uk.org/documents/official-papers/1997-wp-excellence-in-schools.html.
DfE. 2010. "The Importance of Teaching: The Schools White Paper 2010." Norwich: The Stationery Office. https://www.gov.uk/government/publications/the-importance-of-teaching-the-schools-white-paper-2010.
DfE. 2016. "Educational Excellence Everywhere." https://www.gov.uk/government/publications/educational-excellence-everywhere.
DfE. 2022. "Opportunity for All: Strong Schools with Great Teachers for Your Child." https://www.gov.uk/government/publications/opportunity-for-all-strong-schools-with-great-teachers-for-your-child.
DfES (Department for Education and Skills). 2005. "Higher Standards, Better Schools for All More Choice for Parents and Pupils." Norwich: The Stationery Office.
Dillard, J., and E. Vinnari. 2018. "Critical Dialogical Accountability: From Accounting-Based Accountability to Accountability-Based Accounting." *Critical Perspectives on Accounting* 62:16–38. https://doi.org/10.1016/j.cpa.2018.10.003.
Dubnick, M. J. 2014. "Accountability as a Cultural Keyword." In *The Oxford Handbook of Public Accountability*, edited by M. Bovens, R. Goodin, and T. Schillemans, 23–28. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199641253.001.0001.
Earley, P., and D. Weindling, eds. 2004. "Accountability and Ethically Responsible School Leadership." In *Understanding School Leadership*, 71–82. London, California and New Delhi: SAGE Publications, Limited.
Gilbert, C. 2012. *Towards a Self-Improving System: The Role of School Accountability*. Nottingham: National College for School Leadership. https://core.ac.uk/download/pdf/9588769.pdf.
Glassow, L. N. 2023. "Inequitable Teacher Turnover and Performance-Based Appraisal: A Global Trend?" *Journal of Education Policy* 39 (2): 276–302. https://doi.org/10.1080/02680939.2023.2189753.
gov.uk. 1995. "The Seven Principles of Public Life." https://www.gov.uk/government/publications/the-7-principles-of-public-life/the-7-principles-of-public-life–2 .
Harvey, W. S. 2011. "Strategies for Conducting Elite Interviews." *Qualitative Research* 11 (4): 431–441. https://doi.org/10.1177/1468794111404329.
Hofer, S. I., D. Holzberger, and K. Reiss. 2020. "Evaluating School Inspection Effectiveness: A Systematic Research Synthesis on 30 Years of International Research." *Studies in Educational Evaluation* 65 (June): 100864. https://doi.org/10.1016/j.stueduc.2020.100864.
House of Commons. 2016. The Role of Regional Schools Commissioners. *The Role of Regional Schools Commissioners Report*.
Leader, J. A., and B. L. Pazey. 2023. "More Than a Test Score: Toward a More Balanced School Accountability System." *School Leadership Review* 17 (2): 4. https://eric.ed.gov/?id=EJ1389943.
Levatino, A., A. Verger, M. Camphuijsen, A. Termes, and L. Parcerisa. 2024. "School Governance Through Performance-Based Accountability: A Comparative Analysis of Its Side Effects Across Different Regulatory Regimes." *Review of Research in Education* 48 (1): 248–286. https://doi.org/10.3102/0091732X241270672.
Levin-Rozalis, M., B. Rosenstein, and J. B. Cousins. 2009. "A Precarious Balance: Educational Evaluation Capacity Building in a Globalized Society." In *The Sage International Handbook of Educational Evaluation*, edited by K. E. Ryan and J. B. Cousins, 191–213. California: Sage.
Lewis, S., G. C. Savage, and J. Holloway. 2019. "Standards without Standardisation? Assembling Standards-Based Reforms in Australian and US Schooling." *Journal of Education Policy* 35 (6): 737–764. https://doi.org/10.1080/02680939.2019.1636140.

Lindberg, S. I. 2009. "Accountability: The Core Concept and Its Subtypes: Working Paper." London: Overseas Development Institute. https://assets.publishing.service.gov.uk/media/57a08b4740f0b652dd000bd6/APPP-WP1.pdf.

Lindberg, S. I. 2013. "Mapping Accountability: Core Concept and Subtypes." *International Review of Administrative Sciences* 79 (2): 202–226. https://doi.org/10.1177/0020852313477761.

Lingard, B., and S. Sellar. 2013. "'Catalyst data': Perverse Systemic Effects of Audit and Accountability in Australian Schooling." *Journal of Education Policy* 28 (5): 634–656. https://doi.org/10.1080/02680939.2012.758815.

Logan, C. 2023. "40 Years of Education Secretaries." *Ssat* (Blog). December 12, 2023. https://www.ssatuk.co.uk/blog/40-years-of-education-secretaries/ .

Maguire, M., S. Gewirtz, E. Towers, and E. Neumann. 2020. "Contextualising Policy Work: Policy Enactment and the Specificities of English Secondary Schools." *Research Papers in Education* 34 (4): 488–509. https://doi.org/10.1080/02671522.2019.1601758.

Mandel, Z., and A. Pendola. 2021. "Policy and Principal Turnover: The Impact of the Texas Special Education Cap." *Education Policy Analysis Archives* 29 (152): 1–32. https://doi.org/10.14507/epaa.29.5681.

Maroy, C., and A. Voisin. 2017. *Think Piece on Accountability*. UNESCO. https://shs.hal.science/halshs-01705982v1 .

Miles, M. B., A. Michael Huberman, and J. Saldaña. 2013. *Qualitative Data Analysis: A Methods Sourcebook*. 3rd ed. London: Sage.

Miller, P., and N. Rose. 2008. *Governing the Present: Administering Economic, Social and Personal Life*. Cambridge: Polity Press.

Milner, A. L., P. Mattei, and C. Ydesen. 2021. "Governing Education in Times of Crisis: State Interventions and School Accountabilities During the COVID-19 Pandemic." *European Educational Research Journal* 20 (4): 520–539. https://doi.org/10.1177/14749041211022198.

Milner, A. L., and C. Ydesen. 2024. "Producing Accountability with Autonomy in Denmark and England: The Analytical Potential of Assemblage Theory with Critical Realism in Comparative Education." *Nordic Journal of Studies in Educational Policy* 10 (3): 219–232. https://doi.org/10.1080/20020317.2024.2392892.

Moss, G. 2022. "Researching the Prospects for Change That COVID Disruption Has Brought to High Stakes Testing and Accountability Systems." *Education Policy Analysis Archives* 30 (139): 1–24. https://doi.org/10.14507/epaa.30.6320.

Norris, N., and S. Kushner. 2007. "The New Public Management and Evaluation." In *Dilemmas of Engagement: Evaluation and the New Public Management*, edited by S. Kushner and N. Norris, 1–16. Leeds: Elsevier. https://doi.org/10.1016/S1474-7863(07)10013-2.

Olivant, K. F. 2015. "I Am Not a Format": Teachers' Experiences with Fostering Creativity in the Era of Accountability." *Journal of Research in Childhood Education* 29 (1): 115–129. https://doi.org/10.1080/02568543.2014.978920.

Parcerisa, L., M. Pagès, A. Termes, and J. Collet-Sabé. 2022. "Why Do Opt-Out Movements Succeed (Or Fail) in Low-Stakes Accountability Systems? A Case Study of the Network of Dissident Schools in Catalonia." *Education Policy Analysis Archives* 30 (133): 1–23. https://doi.org/10.14507/epaa.30.6330.

Perryman, J. 2009. "Inspection and the Fabrication of Professional and Performative Processes." *Journal of Education Policy* 24 (5): 611–631. https://doi.org/10.1080/02680930903125129.

Perryman, J., and G. Calvert. 2020. "What Motivates People to Teach, and Why Do They Leave? Accountability, Performativity and Teacher Retention." *British Journal of Educational Studies* 68 (1): 3–23. https://doi.org/10.1080/00071005.2019.1589417.

Purdue, D. 2005. "Performance Management for Community Empowerment Networks." *Public Money and Management* 25 (2): 123–130. https://doi.org/10.1111/j.1467-9302.2005.00462.x.

Rasmussen, P., and Y. Zou. 2014. "The Development of Educational Accountability in China and Denmark." *Education Policy Analysis Archives* 22 (121): 121. https://doi.org/10.14507/epaa.v22.1693.

Sahlberg, P. 2011. *Finnish Lessons: What Can the World Learn from Educational Change in Finland?*. New York and London: Teachers College. https://eric.ed.gov/?id=ED525697.

Savage, G. C. 2019. "What is Policy Assemblage?" *Territory, Politics, Governance* 8 (3): 319–335. https://doi.org/10.1080/21622671.2018.1559760.

Schedler, A. 1999. "Conceptualizing Accountability. The Self-Restraining State: Power and Accountability in New Democracies." In *The Self-Restraining State: Power and Accountability in New Democracies*, edited by A. Schedler, L. Diamond, and M. F. Plattner, 13–28. Boulder London: Lynne Rienner Publishers.

Schleicher, A. 2014. *Equity, Excellence and Inclusiveness in Education: Policy Lessons from Around the World*. Paris: OECD Publishing.

Trask, S., and B. Cowie. 2022. "Tight-Loose: Understanding Variability, Trade-Offs and Felt Accountability Across the Curriculum-Pedagogy-Assessment Dynamic." *The Curriculum Journal* 33 (4): 587–601. https://doi.org/10.1002/curj.163.

Van Audenhove, L., K. Donders, and M. Puppis. 2019. "Talking to People III: Expert Interviews and Elite Interviews." In *The Palgrave Handbook of Methods for Media Policy Research*, edited by H. Van den Bulck, M. Puppis, K. Donders, and L. Van Audenhove, 179–197. Switzerland: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-16065-4.

Verger, A., and L. Parcerisa. 2017. *Accountability and Education in the Post-2015 Scenario: International Trends, Enactment Dynamics and Socio-Educational Effects*. UNESCO.

Webb, T. P. 2005. "The Anatomy of Accountability." *Journal of Education Policy* 20 (2): 189–208. https://doi.org/10.1080/0268093052000341395.

West, A., P. Mattei, and J. Roberts. 2011. "Accountability and Sanctions in English Schools." *British Journal of Educational Studies* 59 (1): 41–62. https://doi.org/10.1080/00071005.2010.529416.

Wilkins, A., J. Collet-Sabé, B. Gobby, and J. Hangartner. 2019. "Translations of New Public Management: A Decentred Approach to School Governance in Four OECD Countries. Globalisation." *Globalisation, Societies & Education* 17 (2): 147–160. https://doi.org/10.1080/14767724.2019.1588102.