

DIRECTORATE FOR EDUCATION AND SKILLS

**MEANS, ENDS AND MEANING IN ACCOUNTABILITY FOR STRATEGIC
EDUCATION GOVERNANCE**

OECD Education Working Paper No. 204

Glenn Fahey (Centre for Independent Studies) and Florian Köster (OECD)

This working paper was prepared by Glenn Fahey and Florian Köster at the OECD Directorate for Education and Skills. It lays the groundwork for developing tools that help education systems systematically improve their accountability arrangements. It is part of the Strategic Education Governance project at the OECD Centre for Educational Research and Innovation (CERI).

The Strategic Education Governance (SEG) project supports countries in identifying the best ways to achieve national objectives for education systems in a context of multi-level governance structures and complex environments. It identifies and promotes effective governance processes in the realms of accountability, capacity, knowledge governance, stakeholder involvement, strategic thinking and a whole-of-system perspective.

Glenn Fahey (gfahey@cis.org.au) and Florian Köster (florian.koester@oecd.org)

JT03449814

OECD EDUCATION WORKING PAPERS SERIES

OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed herein are those of the author(s).

Working Papers describe preliminary results or research in progress by the author(s) and are published to stimulate discussion on a broad range of issues on which the OECD works. Comments on Working Papers are welcome, and may be sent to the Directorate for Education and Skills, OECD, 2 rue André-Pascal, 75775 Paris Cedex 16, France.

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgement of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to rights@oecd.org.

Comment on the series is welcome, and should be sent to edu.contact@oecd.org.

This working paper has been authorised by Andreas Schleicher, Director of the Directorate for Education and Skills, OECD.

www.oecd.org/edu/workingpapers

© OECD 2019

Acknowledgements

The authors thank Professor Melanie C.M. Ehren (Vrije Universiteit Amsterdam) and Professor Thomas Schillemans and his team at Utrecht University School of Governance for valuable comments on the draft of the paper.

Many thanks to Claire Shewbridge (OECD) and Marinus 'Rien' Rouw (Ministry of Education, Culture and Science, Netherlands) for their valuable contributions and guidance throughout this work, and to Leonora Lynch-Stein (OECD) for editorial support.

Abstract

Recent decades have revealed a gap between promises and realities of accountability in education governance, as well as further afield. Despite efforts identifying and analysing cautionary tales of accountability interventions, a systematic approach to support progressive improvements for managing accountability in complex education systems is yet to be widely adopted. This review of the interdisciplinary literature leverages the complexity paradigm to address the extant lack of clarity in theory and practice surrounding accountability. Central to this are recent research efforts that embrace accountability's foundations as a relational exchange of accounts. The approach enables distancing accountability from misconceptions arising from its patchy track record in practice and the rhetorical usage of the term. The review discusses implications for understanding and managing accountability within contemporary strategic governance. In such an environment, this means rethinking the management of accountability stresses by practitioners as a new normal to be grappled with – and where possible attenuated – rather than the object of tractable solutions, as has been mistakenly conceived in the past.

Résumé

Les dernières décennies ont révélé un écart entre les promesses et réalités de la responsabilité de la gouvernance dans l'éducation, mais également au-delà. Malgré les efforts pour identifier et analyser les avertissements des interventions de responsabilité, l'approche systématique pour supporter les améliorations progressives de gestion de responsabilité dans les systèmes éducatifs complexes reste toutefois encore loin d'être adoptée largement. Cette revue de littérature interdisciplinaire fait appel au paradigme de complexité afin de résoudre le manque existant de clarté autour de la responsabilité, dans la théorie comme dans la pratique. Au cœur de la question se trouve les récents efforts de recherches qui définissent les fondations de la responsabilité comme un échange relationnel de comptes. Cette approche permet d'éloigner la responsabilité de tous les malentendus soulevés par son bilan inégal dans la pratique et l'utilisation rhétorique du terme. Cette revue discute des implications pour comprendre et gérer la responsabilité dans la gouvernance stratégique contemporaine. Dans un tel environnement, cela implique de repenser la gestion de la responsabilité, souvent faussement considérée dans le passé, et perçue par les professionnels comme la nouvelle norme à combattre – ou si possible atténuer – plutôt que comme l'objet de solutions concrètes.

Table of Contents

Acknowledgements	3
Abstract	4
Résumé	4
1. Introduction	7
2. Complexity foundations for accountability	9
2.1. Accountability within complex policy settings.....	9
2.2. Accountability and the relations between the parts of complex systems	11
2.3. Accountability within complex adaptive systems.....	13
2.4. Implications of complexity for conceptual development of accountability.....	14
3. The means of accountability	16
3.1. Instruments regulating behaviour of actors.....	16
3.2. Participants to an accountability exchange	21
3.3. Fundamental dimensions of accountability exchanges.....	25
3.4. Experienced accountability – gaps, deficits and overloads.....	31
4. The ends of accountability	37
4.1. Promise of integrity	37
4.2. Promise of democratic legitimacy	38
4.3. Promise of justice.....	39
4.4. Promise of control.....	40
4.5. Promise of appropriate behaviour	44
4.6. Promise of performance.....	45
4.7. Promises and harsh realities.....	48
5. Learning and meaning in accountability.....	53
5.1. Leveraging benefits from both process and outcome accountability.....	53
5.2. Meaningful accountability: effective exchanges and learning.....	58
6. Conclusion and outlook.....	65
References	67

Tables

Table 1. Implications of complexity for accountability	15
Table 2. Mechanisms to regulate behaviour in public management	17
Table 3. Hybrid instruments to regulate behaviour	21
Table 4. Dimensions of accountability at each stage of the accountability process.....	26
Table 5. Suboptimal promises and realities.....	49
Table 6. Potential compromised promises based on two-way combinations	52
Table 7. Implications of accountability stresses on conduct, exchanges and possible learning.....	60
Table 8. Conditions for learning under accountability	63

Figure

Figure 1. Flow and direction of accountability between accountability exchange participants 23

Boxes

Box 1. Interdisciplinary approaches to accountability 8
Box 2. The Friedrich-Finer debate 28

1. Introduction

Accountability is an essential component of strategic governance in complex education systems. Designing and managing effective accountability processes is a perennial challenge at all levels of education governance – from system-level policy makers and departmental authorities, to teachers, school leaders, parents and school boards. Scholars, too, have long grappled with the contentious conceptual and practical considerations regarding accountability (Bovens, 2007^[1]; Lindberg, 2013^[2]; Bovens, 2010^[3]).

Many contentions revealed in the literature stem from imprecision in specifying *what we mean* by ‘accountability’ and exactly *what we want out of it*. The term ‘accountability’ can be entangled with wider political and stakeholder contestation (Romzek, 2000^[4]), which, in education circles, has sometimes earned it an unenviable status as a ‘dirty word’, coupled with a common, albeit mistaken, association with punitive consequences (Suspitsyna, 2010^[5]). Scholars point to ‘conventional wisdoms’ in accountability, but many of these are considered inappropriate as they are unable to reflect that, in practice, accountability is extremely complex (Willems and van Dooren, 2011^[6]). As governance has become more complex, so too have the corresponding accountability arrangements (Pierre and Peters, 2005^[7]; Schneider, 2012^[8]), such that “if accountability was ever a straightforward concept that time has now passed” (Lewis, O’Flynn and Sullivan, 2014, p. 401^[9]).

Against this background, practical recommendations for governance purposes, to date, have tended to be limited to a list of ‘do nots’. Motivated to identify ways that accountability can facilitate opportunities for learning and continuous improvement within education systems, this review aims to uncover constructive avenues towards accountability from the interdisciplinary literature (see Box 1. Interdisciplinary approaches to accountability).

Following this introduction, Section 2 establishes the complexity foundations as suitable inroad to the study of accountability in modern education systems. This exposes some of the deficiencies of current approaches towards managing accountability. In particular, complexity theory underlines the importance of the interactions among components in a complex system. Accordingly, this suggests the need to analyse accountability exchanges and the respective parties to the exchange – namely those rendering account, known as *actors*, and those holding to account, known as *forums*.

Section 3 examines the means of accountability. It outlines three key stages of accountability exchanges between actors and forums: account rendering; account assessment; and account judgement. It establishes a classification of the range of instruments that are commonly employed in regulating behaviour, before proceeding to consider the participants in accountability exchanges, and then how accountability is manifested, or experienced, through instruments. Finally, the implications of experiencing: none, too much and too little accountability out of exchanges are discussed.

Section 4 discusses the ends of accountability relationships with respect to its associated promises and realities. It examines how accountability may promise integrity, democratic legitimacy, justice, control, appropriate behaviour and performance – along with potential stresses each may bring or be subjected to.

Section 5 synthesises the lessons informed by the literature, particularly with respect to the key findings from social psychology on accountability that contribute towards a meaningful

conceptual understanding and towards an application for practice. This also includes an examination of the expected benefits from process accountability and from outcome accountability and how to avoid trade-offs in designing and implementing these.

Section 6 concludes the review with a discussion of the implications for managing accountability. The section provides an outlook for future research and derives preliminary practical insights for managing accountability to establish meaningful exchanges.

Box 1. Interdisciplinary approaches to accountability

Within public administration and management, accountability is generally considered a ‘magical concept’ (Pollitt and Hupe, 2011^[10]) that is somewhat ‘hard to oppose’. Indeed, despite a proliferation of accountability failures identified in this literature – mostly observations of ‘overloads’ and ‘deficits’ of accountability in practice – this does not tamper an optimism for its virtues in terms of curbing the conduct of potentially wayward public officials (Bovens, 2010^[3]).

Within the fields of management and organisational studies, accountability is generally oriented towards performance management (say, employee-employer relations and appraisal rating processes) and monitoring (say, in how individuals and organisations perform under measurement regimes), through to related matters like corporate governance. This literature points to the potential for accountability to bring advantages to individuals and organisations (such as improved individual, team and organisational performance) as well as potential adverse effects (such as in the form of job stress and tensions, impacting upon job satisfaction).

In political studies (including international relations), the prevailing concern surrounds how the public regulates the conduct, and exposes possible impropriety, of elected (and unelected) representatives – particularly with respect to concerns of so-called ‘democratic deficits’. In accounting literature, accountability is interested in promoting accurate recording and reporting of information, and setting constructive auditing regimes – that is, with an emphasis on matters of financial probity. Legal scholarship focuses on due process and the integrity of procedures used to reach judgements – emphasising balanced assessments of evidence, application of due scrutiny and the legitimacy of determinations.

Finally, accountability in behavioural sciences, particularly in social psychology (generally informed by observations under experimental settings), follows an analytical approach that is more generalised and divorced from particular ‘real-world’ events, policy instruments and the like. The discipline has devoted much attention to the core of accountability in a generalised sense (Frink and Klimoski, 1998^[11]) – namely the relationship between those exchanging accounts. It draws on the ‘social contingency model of judgement and choice’ (Tetlock, 1992^[12]). This attention has produced significant conceptual and analytical insights, despite only in recent years being integrated within public administration (Grimmelikhuijsen et al., 2017^[13]; Schillemans, 2016^[14]).

2. Complexity foundations for accountability

Education governance has regularly been described as complex (Snyder, 2013^[15]; Burns and Köster, 2016^[16]; Mason, 2008^[17]). This is because education systems cannot be considered as isolated (devoid of context), of limited connectivity or interdependency (devoid of collaborative governance and decentralisation) or static (devoid of interdependence over time). In turn, the complexity perspective has received attention within the literature and in practice over recent years (Burns, Köster and Fuster, 2016^[18]), including systems approaches to address public sector challenges more generally (OECD, 2017^[19]).

Rather than a strictly unified set of theoretical axioms, the complexity paradigm represents a broad branch of science which provides a general approach that seeks to better explain a range of contemporary real world phenomena (Mitchell, 2009^[20]; Mason, 2008^[17]; Horgan, 1995^[21]), particularly through offering alternative explanations to those of traditional scientific paradigms (Mitleton-Kelly, 2003^[22]; Boulton, 2015^[23]). This has led researchers to identify and transfer a series of important lessons from complexity science into interdisciplinary fields of study¹ and practice (Maguire and McKelvey, 1999^[24]; Teisman and Van Buuren, 2007^[25]; Byrne, 1998^[26]), including in educational accountability as a means of overcoming extant challenges in understanding (O’Day, 2002^[27]). To this end, this section considers the implications that a complexity frame of reference offers to the study of accountability in guiding the subsequent aspects of this review. The implications derived from complexity when it comes to proceeding towards conceptual development of accountability are summarised in Table 1 to conclude the section.

2.1. Accountability within complex policy settings

The nature of policy problems faced in contemporary systems is generally considered to be more complex than those faced ‘in simpler times’ (Head and Alford, 2015^[28]; Weber and Khademian, 2008^[29]; Head, 2008^[30]). In particular, two relevant developments in recent decades that result from the complexity of contemporary policy settings are most applicable to the study of accountability (Klijn, 2008^[31]). The first is that complex systems are characterised by a wide range of actors, whom are interdependent and self-organising (Kauffman, 1993^[32]). Second, complex systems are characterised by more and more vocal stakeholders (forums) (Hooge, Burns and Wilkoszewski, 2012^[33]).

2.1.1. Diversity and interdependence of actors

The nature of policy problems faced today do not respect ministerial portfolio boundaries, departmental mandates, territorial boundaries or even the work of one level of governance (Burns, Köster and Fuster, 2016^[18]; Hooghe and Marks, 2003^[34]). As a result of this, there is an increased number of actors (multiplexity), who each need each other (interdependency), and their activity increasingly takes place *across* rather than *within* organisations (interactivity) (Osborne, 2006^[35]; Head, 2008^[30]).

¹ While far from new to the fields of public administration and management (Klijn, 2008^[31]), the practical applications for the discipline remain under-utilised (Eppel and Rhodes, 2018^[314]).

Multiplexity² results in ‘many hands’ working together in conducting governance activity, which can make assessment and attribution for accountability less clear, as responsibilities may be shared amongst these ‘many hands’ (Thompson, 1980_[36]; Romzek and Dubnick, 1998_[37]) – see further in Section 3.2.2. Namely, with joint provision and collaborative activity, many actors may each possess respective pieces of the accountability puzzle. A key challenge is ensuring that accountability expectations that are shared are sufficiently felt by relevant parties and are not either diluted (with nobody held to account) or excessively redundant (with everybody held to account) (Mulgan, 2014_[38]; Halachmi, 2014_[39]) – see further in Section 3.4.1.

Interdependency results in accountability relationships that are less vertical and instead more horizontal (Torfing et al., 2012_[40]). Actors are mutually dependent on each other ‘holding their end of the bargain’ in order to fulfil their substantive purposes, service provision expectations, or meet departmental demands (Romzek, LeRoux and Blackmar, 2012_[41]). A key challenge in this situation is to address ‘entanglement’ as actors may ‘wear a range of hats’ in the process (Romzek, 2011_[42]) – see further in Section 3.4.1.

Interactivity results in accountability arrangements that are relatively less formal because they may work around rather than within bureaucratic means (Romzek, 2014_[43]; Romzek et al., 2014_[44]). When accountability can be contained within organisational structures, instruments can follow managerial or supervision lines, but this is not always the case when accountability takes place across organisations. This can make accountability less transparent and may function through more implicit rather than explicit means (Romzek, LeRoux and Blackmar, 2012_[41]; Rhodes, 1997_[45]) – see further in Section 3.3.4.

2.1.2. More and more vocal forums

Complexity manifests in increased local voice and responsiveness for decision making, along with evolving perspectives and priorities for education (Burns, Köster and Fuster, 2016_[18]; Hui and Bonan, 1991_[46]). This is exacerbated by a relatively high propensity for stakeholder contestation given that there can be widely held experiential knowledge of education among interested parties (Hooge, 2016_[47]). There also tends to be a particularly prominent ‘sharedness’ towards accountability in education systems amongst an especially wide group of stakeholders (Linn, 2003_[48]). Moreover, the ideas of what education can and should achieve have transformed over time, accompanied by, at times, stark changes in the envisioned roles and responsibilities of those working in education.³

A multiplicity of stakeholders places greater demand for negotiation and presents a diversity of interests to navigate (Burns and Köster, 2016_[49]). Accordingly, with these ‘many eyes’⁴ (Bovens, 2007_[11]) comes more varied, multiple expectations (Koppell, 2005_[50]; Klingner, Nalbandian and Romzek, 2002_[51]), and being held accountable for achieving broader collective outcomes in service towards the public good (Moore, 2006_[52]; Stoker, 2006_[53]; Gains and Stoker, 2009_[54]).

² For more on ‘multiplexity’ see Liu et al. (2019_[312]), Kapucu and Hu (2014_[310]), Lee and Lee (2015_[311]).

³ A number of historical accounts discuss these changes and discourses (Hargreaves et al., 2009_[126]; Hextall et al., 2007_[127]).

⁴ See also, with reference to ‘many masters’ Romzek and Dubnick (1998_[37]).

Embedded in such context, observers note a tendency in the field of education towards entering accountability-related discussions in a combative and reactive stance (Sahlberg, 2007_[55]; Sahlberg, 2010_[56]). In turn, accountability tensions in education are especially prone to serve as ‘battlegrounds’ for more fundamental stakeholder positions (Reichersdorfer, Christensen and Vrangbæk, 2013_[57]). These well-trodden battlegrounds can produce tensions amongst stakeholders periodically swelling in line with adjustments to respective interests, for example along the lines of reconciling centralised or localised curricula or balancing teacher professional discretion with responsiveness to parental expectations (Jaafar and Anderson, 2007_[58]; Cochran-Smith, Piazza and Power, 2013_[59]).

2.2. Accountability and the relations between the parts of complex systems

A key element of the complexity sciences is that, due to the many connexions that exist in complex systems, it is more meaningful to consider the nature of their *interactions* than focus on the elements of the system themselves (Epstein and Axtell, 1996_[60]). Put differently, the relations between the parts warrant attention more so than the parts *per se*. This is particularly relevant within the education discipline because scholarship tends to employ an ‘instrumental’ approach, analysing interventions almost solely with respect to the key instruments employed in practice and deducing implications accordingly⁵.

However, this focus on instruments (see Section 3.1), with no or limited attention to the relationships underpinning accountability exchanges, can limit more constructive learning about what makes accountability function more effectively. To this end, complexity brings attention to subjectivity of accountability experiences (both because of complex human behaviours and of an irreducible heterogeneity of units); sensitivity to an individual’s history; and to local contextual factors influencing relationships.

2.2.1. Subjectivity of accountability experiences

The complexity paradigm informs that, in a complex world, simple rules do not fully explain the behaviour of actors (Epstein and Axtell, 1996_[60]), suggesting that accountability can be better understood by digging deeper in terms of how decisions are made and behaviours formed by individuals – see, in particular, Section 5. To this end, cognitive studies have observed that behaviours and decision making in practice rarely follow stylised assumptions when it comes to accountability (Lerner and Tetlock, 1999_[61]; Vibert, 2014_[62]; Tetlock, 1992_[12]; Schillemans, 2016_[14]). Instead, social psychologists point towards the tendency for behaviours and decision making to be subject to: heuristics (decision making methods based on cognitive shortcuts) and biases or in particular emotions (preferences held for irrational reasons such as a preference towards the familiar) and self-deception (such as ex-post rationalisation of pre-existing beliefs). In turn, the ‘phenomenological’ perspective on accountability that is a feature within the social psychology discipline (Tetlock, 1992_[12]) is insightful, given its emphasis placed upon the ‘individual, subjective, and internal’ nature of accountability.

⁵ For instance, school inspection and evaluation (Perryman, 2009_[296]; Ehren and Visscher, 2006_[297]); performance appraisal of teaching staff (Shepard, 2000_[298]) and education leadership (Marks and Nance, 2007_[299]); collective professionalism as a moderator for pedagogical governance (Green, 2011_[300]; Eraut, 1994_[301]); transparency of school financing and demonstration of ‘value for money’ (Mante and O’Brien, 2002_[302]; Hoxby, 2003_[303]); and ‘high stakes’ instruments for performance monitoring of schooling (Sirotnik, 2004_[294]).

Similarly, complexity emphasises that individuals are not uniform (Epstein and Axtell, 1996_[60]; Holland, 1995_[63]; Stacey, 1995_[64]), and, so, experiences of participants within accountability exchanges are likely to be subjective because they are felt differently between them (heterogeneous rather than homogeneous units) (Hall et al., 2006_[65]; Hall, Frink and Buckley, 2017_[66]) – see further in Section 3.4. Because experiences are not uniform this suggests that employing accountability instruments may not have uniform effects in stable and predictable ways (Gunderson, 2002_[67]), but instead points to the need to look at how and why individual experiences may differ (Holland, 1995_[63]).

Complexity theory suggests participants associate varying meanings from flows of information (so-called ‘substantive complexity’), undertake strategic choices (so-called ‘strategic complexity’), and form divergent understandings of expectations (so-called ‘institutional complexity’) (Klijn and Koppenjan, 2014_[68]). In turn, the implication is to focus on the connexions and interactions of participants in accountability exchanges, rather than focussing on the implementation of assorted instruments *per se*.

2.2.2. History

Conventional thinking tends to assume linear cause-effect relationships (Stacey, 1995_[64]). Complexity, on the other hand, assumes ‘bi-directionality’ because effects – or, projections of effects – can influence the cause (Styhre, 2002_[69]; Holland, 1995_[63]). Put simply, actions are not independent of the past or future (Stacey, 1995_[64]).

This recognises that participants in accountability relationships undertake strategic choices, since they possess anticipation and memory (Kearns, 1994_[70]). That is, decisions and behaviours today are a function of expectations of the future and reflections on the past. Moreover, the stages of the accountability ‘process’ (Bovens, 2007_[11]) – account rendering, assessment, and judgement – are interdependent, because, for instance, accounts rendered are a function of expected judgements and so on. To this end, the notion of an ‘accountability cycle’ has been proposed (Schillemans, Van Twist and Vanhommerig, 2013_[71]; Schillemans and Smulders, 2016_[72]). Among other reasons, this makes the role of timing in which accountability expectations are set important in determining the result of any exchange (Tetlock, Skitka and Boettger, 1989_[73]) – see further in Section 4.5.

2.2.3. Local context

Complexity promotes adopting a holistic perspective that considers context first, and then, through this lens, interrogates the functioning of accountability arrangements within the wider system – rather than the other way around. To this end, managing accountability requires sensitivity to context regarding what will most appropriately ‘fit’ with relevant policy interventions, contexts, institutional environments, or experiences (Romzek, 2000_[4]) – as these can be enabling or constraining influences on the relationship. The implication for accountability in governance is that the exchanges and participants involved in accountability cannot be taken in isolation of context (Lawrence and Lorsch, 1967_[74]).

Moreover, complexity provides a cautionary warning against replication of instruments indiscriminately across contexts, given that interventions can be contingent on a unique confluence of factors which may, further, be unanticipated and unforeseeable based on initial conditions (Stacey, 1995_[64]). Reflecting that individuals are creatures of their context, and, thus, that this shapes accountability exchanges, research has highlighted, among other things, differential outcomes of accountability interventions according to organisational ‘fit’ (Lanivich et al., 2010_[75]), the appropriateness with respect to socio-cultural contexts (Gelfand, Lim and Raver, 2004_[76]), and to pre-existing beliefs of

participants involved (Tetlock et al., 2013^[77]). This is because different cultures, different organisations (say, team-based cultures and the like) and those holding different ideologies are predisposed to interpret accountability in varied ways that defy easy classification.

2.3. Accountability within complex adaptive systems

Another of the fundamental principles of the complexity paradigm is the dynamic nature of systems – based on observations from fields such as chaos theory (Prigogine and Stengers, 1984^[78]) – and accordingly, the evolutionary nature of so-called ‘complex adaptive systems’ (among others developed from ecology). Rather than operating under linearity, such systems function in a more cyclical fashion. In particular, the ordinary course of business is the experience of both positive (amplifying shocks) and negative feedback (dampening shocks) loops. When it comes to accountability relationships, the implications of the complexity paradigm are most pertinent in terms of: the related notions of disproportionality and unpredictability (essentially, operating under positive feedback); and evolutionary adaptivity, since this provides insight to the operation of systems under perpetual instability (essentially, operating under negative feedback).

2.3.1. Disproportionality and unpredictability

Complexity rejects that for every action, there must be an equal and opposite reaction (proportionality) (Stacey, 1995^[64]). This is because accountability interventions can produce potentially disproportionate and unpredictable responses to marginal adjustments [the so-called ‘butterfly effect’ (Dooley and Van de Ven, 1999^[79]; Lorenz, 1972^[80]; Gunderson, 2002^[67]], which can be reinforced through vicious and virtuous (positive feedback) cycles (van Twist et al., 2013^[81]). To this end, a notionally marginal intervention at one time could produce only a minimal effect, and at another time it could produce an exponential one (based upon what are referred to as the ‘sensitivity of initial conditions’) (Duit and Galaz, 2008^[82]; Granovetter, 1978^[83]). Shocks can also be intensified in practice because individuals can tend to get accustomed to accountability settings, which can make them sensitive to even minor adjustments in arrangements.

In accountability, this follows the observation in practice of periodic, pendulum-like swings over time (Romzek, 2000^[4]), as well as possible ‘tipping points’, ‘threshold effects’ or ‘critical states’, following reactive adjustment and readjustment to address pressures (Duit and Galaz, 2008^[82]). Within the education field, tipping points can be artificially constructed based upon the perceptions of accountability (and, particularly, its role in provision of quality education) that are engendered in cohorts of educators – since they are oriented towards particular settings based upon the arrangements that they have become accustomed to (say, through their training or experience).

Tipping points of this kind can be difficult to recognise because new accountability arrangements are often adopted over time – “without regard to elegance of design or redundancy” (Romzek, 2000, p. 23^[4]) – especially being stacked on top of another, in a thick, layered, and overlapping set of demands (Halachmi, 2014^[39]) – see further in Section 3.4.2. In turn, this could result in over-shooting – since forces can contribute to overcorrection through positive feedback loops. For instance, the management literature on accountability often examines the application of managerial pressures placed on employees and the resulting disproportionate and unpredictable responses in employee stress that are unanticipated and unforeseen from initial conditions (Hochwarter et al., 2005^[84]).

2.3.2. Dynamics of complex adaptive systems

Complexity rejects the assumption of inherent stability of systems, (Stacey, 1995_[64]), instead examining ‘complex adaptive systems’ that are constantly adapting through evolutionary change (Prigogine and Stengers, 1984_[78]; Macintosh and Maclean, 1999_[85]; Gunderson, 2002_[67]). In turn, arriving at and sustaining balanced, equilibrium-like accountability arrangements is the exception rather than the norm, as underlying environments are rarely (and only transitorily) stable (Macintosh and Maclean, 1999_[85]). To this end, accountability research has noted that, rather than constantly arriving at ‘ideal configurations’, it is more a matter of managing trade-offs on an ongoing basis (Tetlock and Mellers, 2011_[86]; Patil and Tetlock, 2014_[87]).

However, it is important to note that, despite complexity, this does not imply that there is no ‘order to the madness’ in complex adaptive systems. Namely, this is because systems are considered to operate best within a degree of ‘bounded instability’ (Stacey, 1995_[64]) – as described in the strategic management literature: ‘surfing’ at the ‘edge of chaos’ (Pascale, 1999_[88]; Carroll and Burton, 2000_[89]). This reflects the implications in the organisational studies literature that envisions a ‘healthy’ degree of tension (such as competition or workplace pressure) as necessary to promote continuous, evolutionary improvement in business processes⁶ (Styhre, 2002_[69]). Indeed, adaptive learning (Duit and Galaz, 2008_[82]) in accountability is considered to occur through the exchange of accounts feeding into subsequent actions, in a cyclical feedback loop (Schillemans and Smulders, 2016_[72]).

Given the dynamic nature of systems, this sees ‘adaptive management’ as the means to embrace complexity in a responsive manner, through perpetual negative feedback loops (Boulton, 2015_[23]) – namely, adapting responsively under pressure of being displaced from balanced points. This highlights the role for managing accountability arrangements by seeking to attenuate possible pressures that place stress on a given accountability exchange. Because of the focus on attenuation of stresses that disrupt the system’s balance, this suggests that, rather than being object of ‘quick fix’ solutions, managing accountability stresses tend not to have tractable solutions. As Romzek (2014, p. 29_[43]) puts it, “accountability is not a puzzle or a formula to be solved” but rather “is a complex dynamic to be managed at best”.

2.4. Implications of complexity for conceptual development of accountability

By synthesising interdisciplinary research on complexity and accountability throughout this section, this points to three main insights that guide the analytical approach throughout the remainder of this paper:

- Efforts to improve accountability systems need to focus on the accountability *relationships* themselves beyond individual accountability *instruments*.
- Stresses on accountability relationships tend not to have tractable solutions. Rather than attempting to ‘solve’ stresses, efforts should seek to *accommodate* complexity of accountability relationships to mitigate stresses.

⁶ Related concepts include Schumpeterian notions of ‘creative destruction’, transformational states and renewal (Chakravarthy and Doz, 1992_[313]; Pascale, 1999_[88]).

- Accountability relationships are grounded in each participant's subjectivity, history and local context.

The implications of complexity (Table 1) motivate developing a suitably generalised analytical approach. To this end, this encourages a return to foundations of accountability exchanges in a manner that is non-specific to particular instruments of education governance. In turn, importing the relational approach that has become popular within the public accountability literature (Bovens, 2007^[1]) provides the opportunity to introduce these micro foundations accordingly. This is because it focusses instead on how accountability is experienced rather than on how behaviour is controlled.

Table 1. Implications of complexity for accountability

Dimension	Implications
Diversity and interdependence of actors	<ul style="list-style-type: none"> • Actors work with others across organisations, since complex policy problems can rarely be addressed single-handedly. • There are more actors involved, which are also more diverse. This means that accountability instruments tend to be more horizontal (often also reciprocal) and informal.
More and more vocal forums	<ul style="list-style-type: none"> • Stakeholders and their demands are potentially highly varied, with competing expectations that are difficult to anticipate and meet. This means that accountability instruments are arbitrated against expectations that are difficult to discern.
Subjectivity of accountability experiences	<ul style="list-style-type: none"> • The accountability exchanges are shaped in the relationship between two parties – the actor <i>and</i> the forum. The analytical focus rests on the relationship between them. Instruments matter in their role to shape relationships and how accountability is experienced. • The behaviours of both actors and forums make decisions following complex rules. The analytical focus integrates the human element of relationships, recognising that inconsistencies are the rule rather than the exception. • Experiences of accountability exchanges vary greatly, so that considerations need to be sensitive to the participants involved, their context and history.
History	<ul style="list-style-type: none"> • Decisions and behaviours today are a function of expectations of the future and reflections on the past so that accounts rendered, assessed and consequences attached are interdependent. As actors can anticipate account rendering, this influences not only their conduct but the exchange of accounts.
Local context	<ul style="list-style-type: none"> • Instruments cannot be treated in isolation of context and require a holistic approach, taking as point of departure the context and what is appropriate in a given context.
Potential for disproportionality and unpredictability	<ul style="list-style-type: none"> • Interventions to accountability arrangements have the potential to produce disproportionate responses by under- or over-shooting. This demands careful consideration of initial conditions ahead of interventions and room for adjustments throughout.
Dynamics	<ul style="list-style-type: none"> • Accountability arrangements are rarely in harmonious balance. The norm is for the exchange of accounts to be under pressure of stresses, which need to be attenuated. • Actors learn and evolve over time, and this means that accountability instruments and expectations must 'move with the times'.

3. The means of accountability

Contributing to the challenge faced by accountability in theory and practice is a lack of common language and general definitions (Mulgan, 2000^[90]). Arriving at common definitions is far from straightforward because the remit of exchanges under the accountability banner (the ‘means of accountability’) has greatly expanded (Romzek, 2014^[43]).

Nonetheless, this section aims to introduce a common language applicable across governance contexts and with special attention to education systems. To establish a general analytical approach, this review focusses on the notion of ‘accountability exchanges’. An accountability exchange is defined as any “relationship between an actor and a forum, in which the actor has an obligation to *explain* and *justify* his or her conduct, the forum can *pose questions* and pass *judgement*, and the actor may face *consequences*” (Bovens, 2007, p. 447^[1]).

First, it describes broad families of instruments that may be used in accountability exchanges (Section 3.1). Then it defines the respective parties to the exchange – namely those rendering account (actors), and those holding to account (forums) (Section 3.2). It then presents the fundamental dimensions of accountability exchanges⁷ (Section 3.3) at three distinct phases – albeit interdependent (Schillemans and Smulders, 2016^[72]): account rendering; assessment of accounts (evaluation of actors’ conduct and justifications); and account judgement and application of consequences (Bovens, 2007^[1]). Finally, it introduces the important notion of the subjective (‘felt’) nature of accountability exchanges (Section 3.4), by way of explaining the various cases of accountability ‘deficits’ and ‘overloads’.

3.1. Instruments regulating behaviour of actors

The predominant analytical focus to date has tended to concentrate on the array of accountability instruments at disposal of practitioners – to varying degrees with an interest in control of behaviour through some mechanism or other (and, accordingly, neglecting the nature of accountability exchanges). This is unsurprising since the design and implementation of interventions represents a central concern for practitioners and scholars alike. As accountability exchanges are articulated through an instrument of one kind or another, appropriate mapping of the range of instruments that are employed is instructive.

3.1.1. Basic families of instruments regulating behaviour

Much of the accountability discussion in education tends to focus on specific accountability instruments, such as school inspection, standardised testing regimes, teacher performance monitoring, and the like. Though the potential array of instruments cannot be fully categorised exhaustively, four basic means to regulate behaviour are common: oversight, mutuality, competition and contrived randomness (Hood, 1998^[91]) (see Table 2).

The first three of these are particularly recognisable within education systems: oversight instruments are commonly employed in the form, say, of school evaluations or school

⁷ This follows the logic behind Lewis et al’s (2014^[9]) approach of assessing accountability along the lines of: ‘to whom, in relation to what, and why?’

safety certification; mutuality instruments are employed in teachers' professional standards or mentoring efforts; and competition instruments are employed through, say, school choice interventions and voucher schemes.

Moreover, stakeholder positions tend to coalesce around the basic instrument categories. This is because constituencies are most concentrated around each of these respectively – with stakeholders particularly exhibiting preferences in education across oversight-mutuality-competition lines. Namely, oversight is favoured, for instance, by stakeholders seeking to keep ‘a watchful eye’ over the conduct of schools. Mutuality is favoured, for instance, by stakeholders that feel that teaching staff can more constructively monitor behaviour from within the profession rather than from outside. Competition is favoured, on the other hand, by stakeholders that prefer the virtue of choice (leaning on consumer sovereignty arguments) in regulating conduct of schools. Because competing interests can be articulated along these lines, this can motivate practitioners to seek out hybrid instruments to encourage the balancing of respective stakeholder interests – see Section 3.1.2.

Table 2. Mechanisms to regulate behaviour in public management

Instrument type	Explanation	Examples
Oversight	<ul style="list-style-type: none"> Aims to promote desirable behaviour according to expectations of relevant forum. Characterised by monitoring and direction of actors by: inspection, audit, arbitration, authorisation, and certification. 	<ul style="list-style-type: none"> School inspection School safety certification
Mutuality	<ul style="list-style-type: none"> Applies peer group horizontal influence rather than to an external forum – such as permissions, standards, ethics, and discipline formed by the profession. 	<ul style="list-style-type: none"> Professional standards Mentoring
Competition	<ul style="list-style-type: none"> Aims to foster rivalry through harnessing competitive processes, without recourse to oversight. Characterised by instruments seeking to prevent corner-cutting. 	<ul style="list-style-type: none"> School choice interventions Voucher schemes Honourific rewards for exceptional achievements
Contrived randomness	<ul style="list-style-type: none"> The deliberate use of chance in organisational design to constrain opportunities for members to promote private or sectional advantage. 	<ul style="list-style-type: none"> Rotation of senior staff Snap inspections

Source: Adapted from Hood (1998^[92]), “Taking Stock: The State of the Art of the State”, in *The Art of the State: Culture, Rhetoric, and Public Management*, Oxford University Press.

Making use of randomness in instruments

The logic of contrived randomness in the design of accountability instruments (Hood, 2004^[93]) is somewhat less familiar in education nor more generally as a desirable mechanism for accountability (Bovens and Schillemans, 2014^[94]) – despite being a feature of the regulatory governance literature and practice (Majone, 2011^[95]). Indeed, where randomness is employed in education this is often motivated by one of two reasons. The first are resource rationalisation intentions, say to achieve cost reductions – for instance, conducting random inspections of samples of schools rather than a full census of schools. The second is to limit the exertion of authority – say, by conducting a limited number of random inspections rather than inspecting all units. Moreover, Hood (1998^[92]; 1998^[96]) points out that there are also beneficial behavioural implications possible when introducing randomness in instruments, particularly in two ways.

First, at an actor-forum level, scrutiny procedures to reduce the prospects for actors to ‘get away with’ wrongdoing can, and are, regularly subjected to randomisation – for instance,

randomised testing of drivers, security searches, public transportation ticket inspections, or regulatory compliance checks (or indeed, the widely employed organisational practice of ‘mystery shoppers’). Randomisation can be used as a general rule to relieve selection issues that could otherwise result in selective filtering, such as applying favouritism, discrimination, or profiling. Second, at the actor-actor level, the internal structuring of organisations can be randomised to discourage wrongdoing by creating uncertainty of potential payoffs. For instance, governing authorities may be regularly rotated, collegial networks may be randomised, and departments may be segmented to prevent potential conflicts of interest, fraud or collusion.

Making use of redundancy in instruments

Deliberate redundancy in accountability relates to the intentional duplicating and overlapping of efforts (Hoos, Pruijssers and Lander, 2017^[97]), such as with extraneous peer evaluations. While not costless (say, because of duplication of efforts or resources), redundancy in design can be functional (Landau, 1969^[98]; Schillemans and Bovens, 2011^[99]), drawing on the famous ‘Landau thesis’⁸ (Landau, 1969^[98]; Lerner, 1987^[100]) that observed, in particular, that the combination of defective units can compensate for their individual flaws. To this end, mutuality instruments applied to ‘defective’ actors – say, having two actors that may be individually prone to misconduct report on each other’s conduct – can reduce their likelihood of carrying out malpractice, at least, in the presence of each other. This logic is extended to some collective and shared approaches towards accountability as the overlapping pieces of the puzzle necessarily fill possible gaps (Schillemans and Bovens, 2011^[99]; Schillemans, 2010^[101]).

In addition, Landau’s findings also point to other potential benefits derived from redundancy – say, by multiple forums (or, say, in both peer and superior evaluations) providing seemingly duplicated functions. This can be employed to encourage a plurality of perspectives in feedback, as well as to act as an early warning signal of potentially defective forums – in turn, providing an ‘overseeing of overseers’.

3.1.2. Hybrid instruments for regulating behaviour

It is common practice for hybrids of the broad families of instruments to be employed (see Table 3, as a purist application of instruments can be vulnerable to well-documented difficulties (Peters, 2014^[102]; Hood, 1998^[92]; Hood, 1998^[103]). For instance:

- Oversight instruments may curb the discretion of actors, and may be infrequent (that is, behaviour may be controlled effectively, but only during a prescribed day of inspection, audit period, or the like);
- Mutuality instruments are perceived as being prone to excessive ‘closeness’ or, in the extreme, to facilitate corrupt conduct;
- Excessive commitment to competitive mechanisms can distort behaviour disproportionately according to associated rewards; and

⁸ See also Streeter (1992^[309]) and Lerner (1987^[100]), who find that redundancy can take one of three ‘virtuous’ forms: ‘enlightened waste’ (essentially, duplicated account rendering in the possibility that a forum may become defective – akin to what he calls ‘active redundancy’); ‘stressing a survivor’ (maintaining an excess capacity to draw upon should a forum become defective – what Streeter (1992^[309]) calls ‘standby redundancy’); or by ‘mobilising reserves’ (maintaining ‘shadow’ forums that can be activated if required).

- A complete commitment to randomness that leaves all to chance undermines agency.

Accordingly, a range of hybrid augmentations of instruments are applied in practice (Hood, 1998^[92]). The most common are discussed in the remainder of this subsection.

Peer-augmented hybrid instruments

‘Peer-group review’ pertains to externally governed oversight through peers (sometimes considered as ‘trial by peers’) in which an external authority selects members of a peer group or sets the rules governing such group. Peer-group review acts as a check against pure mutuality instruments in which a peer-group may lack impetus to hold its members to account, as well as against the potential of pure oversight instruments propagating a perception that ‘outsiders’ are passing judgement, without sufficient voice of ‘insiders’ (‘us versus them’).

‘Peer-group competition’ employs peer rivalry and competition as a check on conduct – relying upon ‘competition among friends’. Peer-group competition mitigates the tendency of pure competition to narrow objectives and curb co-operation, whereas it may also be used to integrate performance incentives into mutuality instruments. It can promote cooperative conduct towards, together, achieving desired objectives. For instance, members of a faculty may choose to prepare independent research proposals in competing teams. The faculty would then choose the best proposal to represent the entire faculty at a conference.

Randomisation-augmented instruments

‘Randomised competition’ employs randomised processes to apply to a competitive mechanism. Typically, instruments of randomised competition are characterised by offering a defined set of possible evaluative standards for actors to compete towards, and then employing a semi-random process in order to assign priorities across the possible evaluative standards that can be applied during assessment. This is intended as a check against conventional competition (where goals may be displaced in favour of one or few narrow standards), along with semi-randomisation (random assignment of priorities from a known list of possible standards), to introduce an element of *known uncertainty* to assessment of conduct. For instance, a performance measurement system which defines a range of possible metrics but assesses only a randomly selected subset of the full metrics would represent a case of randomised competition. This can sometimes be illustrated as akin to students knowing which content examinations may possibly cover – that is, knowledge the *possible* questions or parameters (as opposed to being able to predict the questions that *will* be assessed) – encouraging them to study the breadth of a course, rather than focussing narrowly.

Instruments of ‘randomised oversight’ introduce unpredictability to activities of oversight. For instance, in fields of regulatory compliance, such instruments are important as a means to avoid ‘capture’ and potential conflicts of interest, by regularly rotating assessors between industries. In addition, randomised oversight can be employed when evaluators or inspectors reduce the predictability of their activities. Unlike the case of randomised competition, the particulars of, say, an inspection, remain fully known, only the timing or particular inspector deployed is unpredictable.

Quasi-markets and randomisation of authority

A hybrid of competition and oversight, ‘quasi-markets’ employ competition to encourage actors to meet an external evaluative standard. Pure oversight may not offer sufficient impetus for actors to self-determine their conduct towards some desired end, or fail to incentivise performance above a set baseline standard. Quasi-markets also offer some degree of standard comparison to which competition amongst actors is anchored – unlike with pure competition instruments, which may not necessarily compel a particular outcome from actors. Quasi-markets are employed, for instance, in procurement of commercial contracts, since tenderers are requested to respond to common criteria while providing scope to demonstrate their intended scope, unique approaches, and as a check on pricing. Through similar means, the employment of ‘league tables’ and the like are commonly employed for a range of service delivery providers⁹ in many jurisdictions.

A hybrid of mutuality and contrived randomness, ‘randomisation of authority’ (‘demarchy’ (Burnheim, 1985_[104])) acts as a check on behaviour of actors in positions of authority – as today’s authority figure does not know with whom authority will lie in the future. This can mitigate against the abuse of power (because they may be subjected to the same abuse once without authority), and conflict of interest (by building the incentive to curry favour with a broad coalition rather than narrow, sectoral interests). Whilst in many situations, strictly randomised authority (such as pure selection of rulers and decision makers by lot) may be an abstraction, leadership in organisations that impose rotational rules for governance roles embody this general principle. Indeed, the very same principles can also be adopted within, say, local governing boards as a moderator against authority being wielded in an unfair and disproportional manner. This is because board members engage in greater efforts to curry favour with broader interests to ensure that their representation is fair and proportional in accordance with communities that may not have been incorporated – particularly those that may be less vocal – in absence of random authority.

3.1.3. From an instrumental to a relational approach to accountability

Table 3 reiterates the families of instruments and their hybrid forms discussed throughout Section 3.1. However, the utilisation of some instruments may not be appropriate within all governance settings, nor for education specifically. Some instruments that are commonly employed elsewhere may conflict with the broader – particularly, social justice – objectives in education. For instance, there may be few opportunities to replicate some elements of randomisation, such as a ‘mystery shopper’ in classrooms, by way of monitoring conduct. In addition, the use of quasi-market instruments in education can be relatively contentious and it may be difficult to contrive the abstraction with which market

⁹ Nonetheless, applications in education settings, particularly in the case of schools and universities, are widely discussed in the literature with respect to the potential for unintended consequences to result (Espeland and Sauder, 2016_[105]; Waslander, Pater and van der Weide, 2010_[106]; Rosenkvist, 2010_[323]; Koning and Van Der Wiel, 2010_[295]).

mechanisms might produce optimal social outcomes¹⁰, as has been well documented (Espeland and Sauder, 2016_[105]; Waslander, Pater and van der Weide, 2010_[106]).¹¹

Table 3. Hybrid instruments to regulate behaviour

	Competition	Mutuality	Contrived randomness
Oversight	Quasi market	Peer group review	Randomised oversight
Competition		Peer group competition	Randomised competition
Mutuality			Randomised authority

Note: The table omits duplicate combinations

Source: Adapted from Hood, C. (1998), “Taking Stock: The State of the Art of the State”, in *The Art of the State: Culture, Rhetoric, and Public Management*, Oxford University Press.

No instrument is inherently more or less appropriate, but is instead dependent on the context in which it is employed (recall Section 2.2). Here it is important to underline that instruments fundamentally take effect by shaping the behavioural incentives¹² of *participants* in an accountability exchange. They are, thus, as successfully or unsuccessfully implemented based on how participants experience them in the context of their exchange (Ebrahim, 2003_[107]). Indeed, accountability expectations are interpreted by participants (Radin and Romzek, 1996_[108]), requiring a look to the fundamental properties in an accountability exchange and how participants respond to these accordingly.

As the point of departure for developing a generalised conceptual framework to accountability, the following sections discuss the fundamentals of a relational approach to accountability and map its dimensions: who participates in accountability exchanges, the kinds of accountability exerted, and how participants experience accountability.

3.2. Participants to an accountability exchange

The participants to an accountability exchange – actors and forums – are a natural starting point for decomposing and better understanding the range of accountability exchanges that exist in practice from a relational point of view. This is because the practice of accountability is established through interaction, including in the mutual expectations that are generated amongst those involved in the exchange (Behn, 2001_[109]).

3.2.1. Dyadic accountability exchange

The most stylised case of an accountability exchange is a dyadic relationship (between two parties), such as one of a manager and an employee (Roch and McNall, 2007_[110]). In the simplest of terms, this means that one party gives account (the ‘actor’) and one holds

¹⁰ This is because such interventions can be highly dependent on the mobilisation of forums – in many instances, the general public or public service clients – with respect to, for instance: the interest in, comprehension of, visibility, reliability and integrity of the accounts rendered (Peters, 2014_[102]; Salamon, 2001_[307]; Van de Walle and Bouckaert, 2003_[308]; Le Grand, 2007_[320]). Moreover, economic costs must be minimised – namely: information (search) costs; transactions (relocation) cost; as well as potential for market frictions (spurious decisions).

¹¹ For instance, the ‘naming and shaming’ of schools not meeting inspection or safety standards, and the proliferation of leagues tables, can produce undesirable ‘flight’ that result in strategically complicated policy outcomes (van Twist et al., 2013_[81]).

¹² See for example Section 3.1.1 on behavioural implications for randomness and redundancy.

account (the ‘forum’). The basic behavioural implication is that an account renderer (actor) is acting on behalf of, or based upon expectations of, the account holder (forum) – invoking corollaries to principals and agents (Strøm, 1997_[111]; Strøm, 2000_[112]) – see also Section 4.4.

By way of extension, two further classifications might be applied to a dyadic exchange. First, the positioning of those involved indicates the orientation of the exchange (vertical, horizontal, or diagonal). In the case of the manager and employee, because there is an implied positioning of ‘unequals’ (O’Donnell, 1998_[113]), the account holder (forum) is in a superior-subordinate engagement with an account giver (actor) – representing a vertical positioning. In education, this is often associated with hierarchical compliance demands and general quality control measures (Hooge, Burns and Wilkoszewski, 2012_[33]).

Alternatively, exchanges between broadly equal parties (O’Donnell, 1998_[113]) are of a horizontal positioning. This is present, for instance, in professional exchanges in which colleagues regularly engage in bi-directional exchanges of accountability, say, through a commitment to each hold one another to their ‘end of the bargain’ – though excludes more superficial exchanges that do not result in account rendering, assessment, and judgement. In education, this can be associated with exchanges over conduct amongst, say, teachers or professional staff based on “insight into their educational processes, decision making, implementation, and results” (Hooge, Burns and Wilkoszewski, 2012, p. 8_[33]).

In addition to the respective position of parties exchanging accounts is the direction of accountability flows. This is with respect to whether account giving (or holding) flows in one direction or if there is at least some degree through which accountability flows in both directions (reciprocal). Together with positioning, this provides a way to classify dyadic accountability exchanges along each dimension (Antonioni, 1994_[114]). For instance, in a vertical accountability exchange, say, a patron-client relationship, as illustrated by Lindberg (2013_[2]), there is indeed also the opportunity for clients to demand account of the patron – for instance, to hold the patron to account for delivering promised benefits.

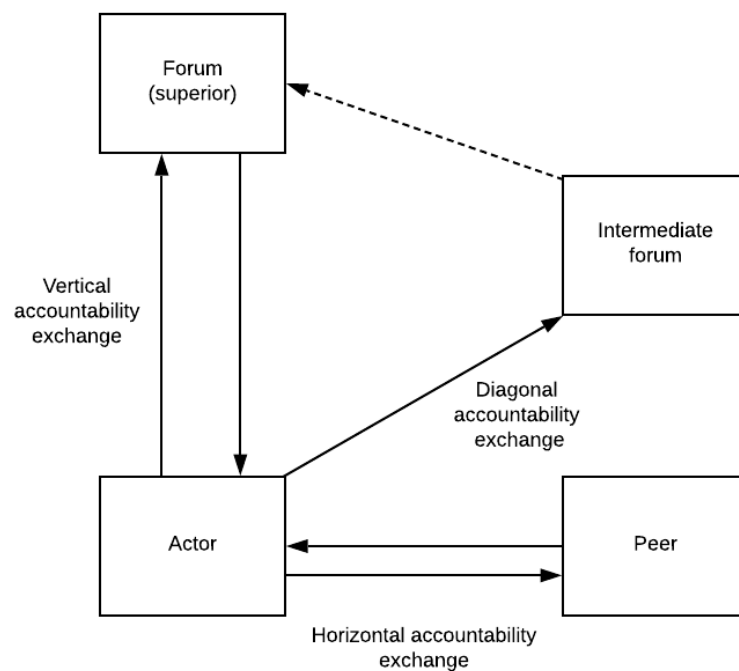
3.2.2. Beyond dyadic accountability exchange

In practice, contemporary public accountability systems often do not resemble a dyadic relationship at all (O’Connell, 2005_[115]; Page, 2004_[116]) – with third parties (indeed, feasibly many possible parties) involved. For instance, employees often enjoy at least some recourse by which, say an unprofessional manager, can be held to account (say, by being subjected to additional supervision from their respective supervisor). Indeed, so-called “360-degree reviews” represent (Behn, 2001_[109]), at least nominally, a series of bi-directional accountability exchanges within a unit of colleagues across multiple levels of hierarchy.

In some instances, third parties in an exchange may have a direct role in completing the flows of accountability – that is, there may be diagonally positioned parties (Bovens, 2007_[11]). This can arise because actors may render account, say, to one party but it is another party that ultimately holds the account – that is, accountability is realised essentially indirectly (because the account rendering has been detached from the assessment and consequences). For instance, the proliferation of ombudsmen and the like fulfil ‘diagonal accountability within many industries – essentially serving as an intermediary between the actor (say, a service provider that is the subject of a complaint) and the forum (say, a consumer complainant). This is because the ombudsmen may not possess powers to directly coerce the actor into compliance (impose consequences), though the forum may report to, say, parliament, ministers, and departmental authorities, who, in turn, can make

assessment and judgements to hold actors to account (albeit with consultation with ombudsmen). In this case, while an accountability exchange is indeed present, the actor and forum do not undertake the exchange directly, but through the intermediary. In the case of educational accountability, a common example is that of school inspectorates that advise ministers on non-compliance and quality concerns, however they do not directly have power to close or sanction a school themselves (Figure 1).

Figure 1. Flow and direction of accountability between accountability exchange participants



Within public service delivery, collaborations across agencies, departmental divisions and the like, rely on respective parties holding each other accountable along the way (Romzek et al., 2014^[44]; Romzek, 2011^[42]; Isett et al., 2011^[117]) – often without a clearly specified chain of authority, and with participants ‘wearing a range of hats’ in the process (see also Section 3.4.2). Indeed, the very act of identifying actors and forums in contemporary systems can be something of a complicated exercise. This can be associated with one of two possible challenges. The first arises because in many exchanges it can be difficult to determine who is ultimately accountable for particular conduct – known as the ‘many hands’ problem (Thompson, 1980^[36]). The second challenge, relatedly, arises because it can be difficult for actors to determine to whom they are ultimately accountable with respect to their conduct – known as the ‘many eyes’ problem (for instance, Bovens, 2007^[5]).

Many hands

The increased uptake of collaborative governance practices (Isett et al., 2011^[117]), in particular, can present difficulties in identifying, attributing, and distributing

accountabilities when there are multiple actors involved (since there are, increasingly, ‘many hands’ at work¹³).

Even within agencies with relatively straightforward functions, activity can conceivably involve multiple actors. For instance, policies may pass through many hands before they are implemented, or decisions could be made in committees and may be subject to various approval chains, before other actors bring these into effect. Governance activities often imply joint responsibilities and rely on co-production, meaning that actors are mutually dependent on each other ‘holding their end of the bargain’ in order to fulfil their substantive purposes, service provision expectations, or meet their own ministerial demands (Romzek, LeRoux and Blackmar, 2012^[41]).

A potential challenge for forums demanding account is in determining who, and to what extent, each actor should be required to render account. For instance, accountability can be applied collectively (all associated parties are equally accountable), individually (associated parties are accountable according to their contribution to the account-giving activity), or hierarchically (the most senior actor is accountable) (Bovens, 2009^[118]).

Shared accountability, though, can invoke associations with the dilution of accountability¹⁴ (Rhodes, 2008^[119]). Yet, while the presence of many hands can be obfuscating for forums, if managed effectively, many hands can potentially contribute to more breadth of account rendering and a shared sense of responsibility among actors. This is because many hands presents the opportunity to improve the division of labour in actors’ account-giving – whereby actors take responsibility for rendering account in behavioural facets in which they appropriately specialise, even though forums may judge conduct collectively.

Moreover, where a shared sense of responsibility is present, account rendering from some actors may complement accounts rendered by others, such that the whole is greater than the sum of its parts¹⁵. For instance, in schools, teachers, professional staff, and school leadership may effectively share responsibilities such that not only do they render account for their cumulative individual conduct, but that their actual conduct is influenced by an understanding of each other’s respective responsibilities.

Many eyes

The problem of ‘many eyes’ is apparent in many governance functions, especially in contemporary service delivery where organisational missions have extended beyond direct service provision to encompass wider societal missions (Acar, Guo and Yang, 2008^[120]; Gains and Stoker, 2009^[54]). This results in a much more diversified set of authorisers (possible forums) – including the public interest more broadly (Moore, 2006^[52]).

This can mean that there are multiple forums to which actors may render account and it can be difficult for actors to be conscious as to which forum they are expected to prioritise when rendering account, and for what (particularly, how to prioritise when these are

¹³ There is also the risk of ‘social loafing’ whereby ‘many hands’ generally dilutes individual efforts in a group setting (Latane, Williams and Harkins, 1979^[319]).

¹⁴ An emerging, but under-studied, domain of accountability is that of ‘group’ accountability (Wallace et al., 2011^[315]). This is, to date, most developed in organisational studies examining accountability of colleagues in team-based activities (Chen, Lam and Zhong, 2007^[316]; Fandt, 1991^[317]; Simons and Peterson, 2000^[318]).

¹⁵ See also the discussion of deliberate redundancy in accountability instruments – Section 3.1.1.

competing – see further in Section 3.4.2) (Klingner, Nalbandian and Romzek, 2002^[51]). Moreover, because forums generally demand different kinds of information and apply different criteria as to what constitutes responsible conduct, they can feasibly pass different judgements on the conduct of the actor (for instance, (Bovens, Schillemans and Hart, 2008^[121]).

For instance, schools are regularly held to account by the varying perspectives and priorities of stakeholders along the lines, in particular, of quality, effectiveness, and efficiency in schooling (for instance, (Hooge, 2016^[47]). Moreover, schools may be judged by forums simultaneously with regards to multiple accounts, with varied expectations for conduct, account rendering, and the like. For instance, schools might render account – not only in accordance with the ultimate function of educating young people (albeit to varied possible forums with varied expectations) – but also to a range of forums for such diverse criteria as the safety and security of schooling environments, cultivation of non-educational student outcomes, school sporting team achievements, or the school’s presence in community involvement.

3.3. Fundamental dimensions of accountability exchanges

While the principal focus for considering accountability has typically been that of instruments regulating behaviour, research has increasingly placed attention towards the fundamental characteristics that affect the experience of participants within accountability exchanges. Not only does this facilitate moving beyond strictly examining the role of accountability as a means of regulating behaviour, but also incorporating how the exchange itself is experienced. Namely, whether accountability is sourced internally or externally; is positively or negatively employed; instruments are mandatory or voluntary; or operate through formal or informal means.

Assessing exchanges in terms of these generic – rather than consequential or functional – terms can help in clarifying the experience of exchanges. Importantly, the various dimensions play out in distinct ways in each phase of the accountability process, that is, account rendering, assessment and judgement (Table 4).

Table 4. Dimensions of accountability at each stage of the accountability process

Dimensions	Account rendering	Account assessment	Account judgement
External	An external forum solicits accounts on the conduct of actors (such as requests/demands for information, oversight of activity, submission of financial records, etc.).	Some degree of universality in expectations is employed as external forums apply some standardisation (at least among comparable or peer units) in terms of how conduct will be assessed. That is, the external forum has some standard that is expected of relevant actors.	Rewards and sanctions are determined and imposed by outside adjudicators (such as adjustments to budgets, withdrawal of external support, further monitoring of activity, etc.).
Internal	Actors provide information on conduct to an insider forum or via self-reflection (such as self-evaluation, peer review).	Assessment of conduct is somewhat differentiated since actors interpret and self-form expectations according to local conditions.	Self-imposed reward and sanction (such as confidence, incentives), or reverted to an insider forum for consequences (such as disciplinary action or organisational reputation).
Negative	Forums, for example, investigate, scrutinise, inspect, and probe actors' conduct.	Forums assess possible breaches, misconduct, infringement, inaccuracies, discrepancies, etc.	Forums enact punishment, such as revocation of privileges, withdrawal of support, penalties, etc.
Positive	Forums seek out, for example, completion, adherence, attainments, and accomplishments in actors' conduct.	Forums set out to recognise the fulfilment, or exceeding, of expectations.	Forums dispense rewards, such as praise, financial compensation, credibility of reputation, etc.
Mandatory	Actors are compelled (non-negotiable) to render accounts, such as: forced disclosures, attendance at hearings, reporting requirements, compulsory inspections, etc.	Assessment of conduct is not negotiable (no opt-out). Evaluative standards against which conduct is assessed are compulsory for actors to comply with.	Forums impose binding rewards and sanctions on actors.
Voluntary	Actors choose to render account to a forum (such as FYIs, unsolicited disclosures etc.), or choose to render accounts in excess of demand (such as supplementary information).	Assessment of conduct is negotiable and evaluative standards may be more flexibly considered (can let minor indiscretions slide). Some accounts may go unassessed as they were not demanded.	The rewards and sanctions dispensed by forums are non-binding (such as recommended actions, optional forfeit of benefits, etc.).
Formal	Forums have codified processes for actors' reporting of conduct (such as legislative requirements, inspection or evaluation schedules, reporting procedures, etc.).	Forums have pre-defined, codified parameters for the assessment of conduct (such as in provisions or evaluative rules/standards) and procedures for carrying out assessment (such as routines, periodic reviews, audit cycles, etc.).	Consequences are prescribed with little or no room for forums to apply their discretion in imposing explicit consequences (such as mandatory minimum penalties, sentencing guidelines, employee promotion against appraisal criteria).
Informal	The processes for actors to report their conduct is not strictly codified (such as impromptu exchange, unstructured interactions, etc.).	Forums are not directed by strictly codified standards for assessing conduct. Instead, assessment may be more <i>ad hoc</i> , tailored to the actor, or contextualised outside strict interpretations of parameters and procedures.	Forums apply implicit rewards (such as compliments, favours, trust etc.) and sanctions (exclusion, diminished reputation etc.) without being strictly bound by codified application of consequences (such as customs or norms).

3.3.1. Internal and external accountability

As a first distinction, accountability demands can originate from a source external or internal to an actor (Lindberg, 2013^[2]; Romzek and Dubnick, 1987^[122]). External accountability occurs when an individual or an organisational actor¹⁶ is held directly accountable by a forum external to the organisation. A classic example of external accountability is the relationship between voters and their elected representatives. In education, this may refer to a teacher being held directly accountable for professional training requirements by a (school-external) education agency. At the organisational level, external accountability may refer to a school board holding a school (as organisation) accountable for spending its budget appropriately.

Internal accountabilities are self-formed within an organisation, group (such as members of a profession), or by individuals (Sinclair, 1995^[123]; Mansouri and Rowney, 2014^[124]). Where this differs with external accountability is that the actor interprets and transforms some external expectations into their own internalised accountability pressures. An example is an organisation's mission statement reflecting the organisation's aim to generate value for shareholders. In education, an example of internal accountability (derived from external accountability demands) are school and teacher self-evaluation efforts that may be appropriated as necessary for the given conditions (Matthews and Ehren, 2017^[125]).

Internal and external accountabilities may be differently distributed¹⁷. In the case of internal accountability, actors can interpret pressures differently – say, by considering which expectations are most appropriate for the organisation. For instance, school leaders may determine which accounts are most relevant to the school based on external expectations, and apply pressures accordingly. In turn, internal accountabilities apply best to actors who are required to be responsive to the demands that come their way, rather than having these determined (to varying extents) externally.

In the case of external accountability, because there can be concerns surrounding an information deficit (particularly where the forum is far removed from the actor), which may result in either undue pressures or 'low bars' in expectations of forums (since they may not have sufficient information to know what is appropriate for the actor from a distance). For instance, external accountabilities can attract criticism for setting expectations that do not sufficiently consider local considerations, or applying 'blanket' expectations across diverse actors.

The issue of balancing external or internal accountabilities has been subject to empirical research and extensive normative debate – particularly with respect to the validity of externally setting and fixing expectations across diverse actors and contexts. In public accountability it is particularly linked to the discretion afforded to public servants (Box 2), and in education, teachers' professional discretion has been a contentious issue for decades (Hargreaves et al., 2009^[126]; Hextall et al., 2007^[127]).

Within education, school leadership plays an important role at the intersection of internal and external accountabilities, by shouldering external account rendering and reconciling

¹⁶ Following Romzek and Dubnick's (1987^[122]) framework, earlier approaches to accountability took the view that the appropriate unit for analysis was the organisation level.

¹⁷ See Section 3.4 for the related accountability 'weight', as well as *degree of control* (Romzek and Dubnick, 1987^[122]).

these with school-level priorities (Knapp and Feldman, 2012_[128]). This is because the leadership should possess the relevant understanding of opportunities and constraints within their organisation and can thus be able to adapt and appropriate external demands accordingly.

Box 2. The Friedrich-Finer debate

The Friedrich-Finer debate of the 1940s refers to a classic discussion over the expectations for modern bureaucracy – namely with respect to the notions of appropriateness of bureaucratic discretion and control. Finer argued for the need for limited discretion and high degree of control and oversight of bureaucrats (whose self-interested behaviours ought to be curbed) by elected politicians in order to preserve political (external) accountability, and that the more punitive the potential sanctions, the more likely to deter recalcitrant behaviour of bureaucrats.

Friedrich’s argument, on the other hand, was that rather than imposing strict external control mechanisms, bureaucrats are, instead, capable of employing self-control through professional norms and expectations, namely in the form of internalised values and beliefs (what we largely refer to today as a ‘public service ethos’). In particular, he points to the need for administrative cultures that challenge political directives (i.e. mutual checks and balances) as opposed to strict rule-following bureaucracies, because the latter could potentially conflict with appropriate behaviours (such as resulting in illegal actions, unethical conduct, or misappropriation of resources in the service of the government of the day).

As Friedrich opined, as governance settings become more complex, bureaucrats would be called upon to utilise their own professional judgement and competence (that is, drawing on their internal accountability) – sometimes associated with the Aristotelian notion of ‘practical wisdom’ – which requires them to make judgements in reaching effective actions in complex situations.

Sources: Carr, F. (1999), “The public service ethos: Decline and renewal?”, *Public Policy and Administration*, <http://dx.doi.org/10.1177/095207679901400401>. On ‘practical wisdom’, Theisens, H. (2016), “Hierarchies, networks and improvisation in education governance”, in Burns, T. and F. Köster (eds.), *Governing Education in a Complex World*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264255364-5-en>.

3.3.2. Positive and negative accountability

Whilst accountability in itself is an essentially neutral concept (for instance, an annual corporate report provides organisational leadership with an avenue to present and describe the organisation’s periodic activity), accountability instruments tend to favour a positive or negative orientation (Behn, 2001_[109]). In regulation, for example, we refer to ‘carrots’ and ‘sticks’.

Positive accountability instruments can be tangible (say, in the form of physical rewards) or intangible (say, in the form of privileges like granting decision-making autonomy). Similarly, tangible negative instruments could be, for instance, sentences for misconduct, while intangible negative instruments could include organisational norms of exclusion and the like.

Nonetheless, actors regularly experience a combination of both negative and positive accountabilities simultaneously – for instance, managers in the corporate sector are

simultaneously submitted to scrutiny and threat of sanction, say, from shareholders, while also being rewarded with incentives like bonuses. In education settings, rewards and sanction are regularly applied, say in terms of school finances in some jurisdictions. For instance, in some systems performance improvements can be rewarded (penalised) with, say, additional (withdrawal or withholding of) discretionary school funding.

The appropriateness of applying positively and negatively oriented instruments depends on actors' preferences and the nature of conduct being considered. To this end, efforts to encourage inspirational or ambitious behaviours of their employees and the like may be more appropriately incentivised by rewards for doing so rather than threats of penalty for failing to. For instance, if it is intended that teaching staff motivate pupils towards attending, participating, and engaging in learning, a heavy reliance on negatively-oriented accountability (say, prospective punitive penalties for variations to instructional guidelines) may undermine the objectives of uplifting and motivating young people to learn in classrooms.

In any case, there can be strong preferences towards negatively oriented accountability within the broader public, particularly with matters of public concern and visibility, prompted with 'off with their heads' notions. For instance, allegations of misconduct may be heard in public forums within judiciary or parliamentary institutions, such as hearings or disciplinary committees and the like – indeed the publicness of denunciation in itself can serve as an important instrument for deterrence.

There are three general reasons that explain a relative tendency towards negatively oriented accountability instruments. First, forums may be motivated from a position of 'mistrust' to prevent an actor erring from expectations (Greiling, 2014^[129]; Mansbridge, 2009^[130]) – that is, accountability acts as a check with which to prevent some undesired behaviour (at least from the perspective of the forum). Second, accountability demands may be borne in response to an undesired event with which a repeat occurrence is to be averted. Deterrence, accordingly, is said to be best sustained when sanctioning is clear, possible, and likely to occur in response to a breach (Lewicki and Bunker, 1996^[131]). Third, and related to the first, is that the association of negative consequences are commonly employed as a forum's recourse, particularly when attempting to prevent undesired actions – namely under threat of penalty. Moreover, the general logic of 'loss aversion' appeals to the general expectation that actors' preferences are greater in terms of avoiding negative consequences, compared to earning positive consequences (Tversky and Kahneman, 1991^[132]).

Nonetheless, the use of positively oriented instruments has become especially relevant in contemporary governance (particularly in common with supplying voluntary accounts – see Section 3.3.3) because it is increasingly difficult to 'surveil' and detect misconduct within complex webs of account renderers under threat of sanctions alone. In part, this is because, for instance, the nature of the burden of detection of misconduct has become more sophisticated (say, big data methods), at the same time that concealing misconduct has become more easy and cost-effective (say, through digital encryption). At the same time, potential sanctions may be a relatively blunt instrument when the likelihood of detection is relatively low. In turn, the design of accountability may likely be made more effective by some integration of both positive and negative orientation (potentially in combination with other dimensions).

3.3.3. Mandatory and voluntary accountability

Accountability exchanges must compel actors and forums to exchange accounts by the most appropriate means available. This often makes it necessary to use mandatory

accountability instruments, as in many circumstances, actors may not willingly render accounts and be subjected to assessment and possible consequences – especially when there is a possibility for negative consequences to be attached. Moreover, because mandatory accountability – such as that enforced by legislation, convention, or committed under contractual arrangement – can reduce uncertainty of exchanges, this does provide some assurance to the public or other interested stakeholders.

However, not all accounts rendered are mandatory, as actors can also be account-seeking (rendering accounts unsolicited by forums) and can choose to oversupply accounts (rendering additional accounts than is demanded by forums¹⁸). For instance, actors may render supplementary information to offer forums additional material on which they can base their assessment. In addition, actors may also voluntarily enter accountability relationships without being compelled externally, and are motivated to do so for a range of reasons (Karsten, 2015_[133]; Koop, 2014_[134]). Particularly when it comes to management literature, this is associated with employees seeking out employer feedback, as well as fulfilling individual needs of purposefulness and self-worth, and providing additional meaning for actions that they have taken.

The significance of understanding the mechanisms to encourage account-seeking behaviour (voluntary accountability) is that this can reduce the dependence of mandatory imposition of accountability in governance systems in which there is preference for flexibility (and potentially lower monitoring costs) (Karsten, 2015_[133]). Indeed, in some settings, the flexibility provided when accountability is voluntary can be more appropriate, and some exchanges are generally more inclined towards voluntary account seeking – such as where self-regulatory norms have been adopted as an evaluative standard. For this reason, efforts to encourage account-seeking activities have become increasingly popular, though the capacity and uptake of account seeking can vary somewhat (Koop, 2014_[134]).

In any case, because some voluntary exchange provides actors with the ability to exit without consequence, this can attract the criticism that it is temporary in nature and can be limited below a threshold of ‘genuine’ accountability (Lindberg, 2013_[2]). For this reason, the extent to which voluntary accountability is indeed truly voluntary remains contentious (Lindberg, 2013_[2]). Namely, because accounts could be rendered voluntarily, for instance: in anticipation that it could be demanded in future anyway (anticipatory); in order to deter being demanded in future (pre-emptive); or in order to influence or bias account holders (confirmatory).

In education, schools face mandatory accountability instruments – often based on legislative and regulatory requirements – that require account rendering to relevant forums (say, for instance, reporting of budgets to provincial authorities). Schools might also voluntarily render account to, say, parents, through disclosures made in forums like school newsletters or websites.

3.3.4. Formal and informal accountability

In many instances, accountability is formalised in instruments that explicitly state who, how, and for what account is to be rendered (say, expectations of hearing procedures), assessed (say, performance criteria), and consequences attached (say, charters and the like which formalise rules and inform of sanctions). More specifically, formal accountability

¹⁸ Indeed, this can contribute towards excess information as generated from redundancy – per discussion in Section 3.1.1; see also Schillemans (2010_[101]); and Schillemans and Bovens (2011_[99]).

could involve explicit performance standards and reporting relationships, such as specification of hierarchies and procedures for external oversight (Romzek, 2000_[4]; Gailmard and Patty, 2012_[135]).

Nonetheless, much accountability can be informal, especially in relatively complex governance and service delivery (Romzek, 2011_[42]), sometimes in settings of low transparency, and with few opportunities for role expectations and performance obligations to be made explicit (Romzek, LeRoux and Blackmar, 2012_[41]). In these settings, the relatively informally derived, shared goals and interpersonal interactions can be more appropriate mechanisms than attempting to implement strictly formalised regimes. In place of formal exchanges, informal exchanges can thus include implicit shared norms, discretionary behaviours, unofficial monitoring and informal rewards and sanctions (Romzek, 2014_[43]; Romzek et al., 2014_[44]).

However, informality is not confined to complex settings and manifests in many kinds of exchanges, such as a means to overcome burdens or complications that can arise out of formality. For instance, in the monitoring of performance between managers and workers, reputational consequences can be an important sanction (Busuioc and Lodge, 2017_[136]), which, albeit informal, may provide scope for managerial influence given that formal performance appraisal instruments may be relatively infrequent and possibly inconvenient. In education, teachers formally render account to school leaders, who might set explicit expectations for their conduct. Alternatively, additional flexibility may be applied, for instance, by school inspectors as an informal reward for trusted school leaders.

The appropriateness of the level of formality can be dependent on, for instance, the preferences according to consistency and variability. This is because formality can be relatively inflexible (say, by being bound to specified rules, such as codified expectations) and potentially cumbersome (say, by being bound to specified processes, such as hearing procedures). At the same time, though, formality means that expectations can be more readily anticipated, and so there may be few surprises when formal expectations are set out in a routine or procedural fashion.

The appropriateness of formality may also depend on the nature of consequences for, or assessment of, conduct. For instance, when high stakes consequences are attached to exchanges, formality may be highly relevant for the purposes of transparency. In addition, political hearings and the like that rely on, say, constitutional expertise or perhaps legalistic expertise for commercial, organisational or management settings, which demands formality in order to adequately express expectations.

3.4. Experienced accountability – gaps, deficits and overloads

While the dimensions discussed describe how accountability exchanges are experienced, it has not considered yet *how much* accountability is experienced. To this end, much debate and research is concerned with how accountability is *felt* (Hall et al., 2006_[65]; Hall, Frink and Buckley, 2017_[66]) – recognising the subjectivity of experiences (recall from Section 2.2). Along these lines, experimental research has focussed on actors' experiences along a progressive scale of accountability¹⁹ to identify the thresholds and pressures in the

¹⁹ For instance, De Zoort et al (2006, p. 376_[304]) propose a four point scale: “anonymity (no accountability), review (general review of performance by a superior), justification (supervisor review of materiality judgements and accompanying specific judgement explanations), and feedback

weight of felt accountability that induce changes in behaviour (Hackenbrack and Nelson, 1996_[137]). For instance, education scholars have observed that interventions tend to induce behavioural changes unequally (Hanushek and Raymond, 2002_[138]). In particular, those whose pre-existing behaviours deviate substantially from expectations are relatively unaffected (since meeting expectations may be assumed to be insurmountable, or may be patently unchallenging). In turn, accountability is felt in large part by those whose behaviour is already relatively congruent to expectations.

The notion of ‘feeling’ accountable is particularly important with respect to actors’ commitment towards promised action (Fry, 1995_[139]). A high commitment to promised actions can allow for less demanding accountability exchanges, potentially increasing flexibility of substantive conduct and improved outcomes²⁰. However, feeling ‘too much’ accountability may overwhelm a given actor or act as a stressor, contributing to ‘strain reactions’, anxiety, and inaction (Hall et al., 2006_[65]; Hochwarter et al., 2007_[140]). Experiencing too little or too much accountability is described in the literature in terms of accountability deficits and overloads, respectively (Mulgan, 2014_[38]; Halachmi, 2014_[39]; Bovens et al., 2014_[141]).

3.4.1. *Accountability deficit*

Accountability deficits can arise in the event that an accountability relationship is not sufficiently in place to capture particular conduct, or where accountability demands are too widely dispersed, and in turn, insufficiently felt within an exchange.

Gaps in accountability

Whilst accountability relationships, by definition, have identified actors and forums in which an exchange takes place, it is not uncommon for gaps in accountability to exist in practice – namely where a suitable actor or forum is ‘missing’. For instance, a suitable forum may not be available or is not readily identifiable by actors – that is, actors may not know who, when, or how to render account for some conduct, even though they may be willing or able to do so (Mulgan, 2014_[38]).

Such episodes can arise where accountability instruments lag the conduct (or actors) for (from) which accounts ought to be rendered (Mashaw, 2014_[142]). A particular cause for this is because accountability expectations can lag the underlying action undertaken. For example, in recent years the explosive growth in complexity of financial instruments, or the rapid development of autonomous vehicles, has outpaced accompanying accountability arrangements. In such instances, actors can be unable to identify to whom, for what, and how to render account for their conduct – as pre-existing arrangements may not be fit for purpose – resulting in a vacuum of accountability.

The upshot of this is that within contemporary governance environments and associated complexity, accountability arrangements are prone to suffer from lagged and discontinuous accountability. As, in practice, accountability is typically reactive in nature, this means that matters can remain unaccounted for until such time as sufficient momentum is generated

(supervisor review of materiality judgements and accompanying specific judgement explanations, with specific supervisor performance feedback)”.

²⁰ A high commitment to promised action might further lead actors to engage in account-seeking behaviour by rendering account without being prompted by a given forum (see Section 3.3.3 on voluntary accountability), potentially leading to lower monitoring costs (Karsten, 2015_[133]).

for action – such as, meeting a threshold for a public advocacy movement to gain traction in demands for a public inquiry.

For this reason, arrangements can benefit from being designed in such a way to better capture changes in behaviour (Bovens and Schillemans, 2014^[94]), prompting some scholars to discuss the need to adopt more ‘dynamic’ and ‘strategic’ approaches to accountability relationships (Kearns, 1994^[70]).

In addition to this, there is practical utility for promoting account-seeking behaviour of actors so that when they undertake actions without a readily identifiable forum, they may seek out opportunities to identify a relevant forum to render accounts to (Karsten, 2015^[133]). Namely, because actors who are willing, able, and motivated to seek account for their conduct (recall from Section 3.3.3), may be most appropriate for generating accounts in instances where forums are ‘missing’, as there is no need to compel them to do so.²¹

Stretched and tangled accountability

The challenges posed by ‘many hands’ can manifest in practice with the experience of thin (stretched), as well as related ‘tangled’ (Romzek, 2011^[42]), accountabilities – and together relate to one side of the family of so-called ‘multiple accountabilities’ challenges in the literature²² (Romzek, 2000^[4]; Romzek and Dubnick, 1998^[37]).

The problem associated with stretched accountability – that is, accountability that is shared across a range of actors – is that the expectations felt (particularly with regard to account rendering efforts) from forums can be diluted (hence, thinly felt) amongst the ‘many hands’ of actors (Rhodes, 2008^[119]). The result is that while accountabilities are shared amongst actors, they may not be intensely (or sufficiently) felt by any, or the distribution of accountability felt may not fully reflect the substantive behaviour – namely because it is unclear for actors to determine their respective ‘slice of the action’ and distribute accountabilities accordingly. Ultimately, this results in a deficit of account giving to relevant forums – a problem that is particularly identifiable within networked and collaborative governance conditions (Mulgan, 2014^[38]).

By extension, this condition of stretched accountability may not necessarily allow a forum to hold any particular actor accountable for outcomes – namely because ‘many hands’ contribute to constituent parts of a whole. This can contribute to an environment of ‘blame games’ (Hood, 2002^[143]; Schillemans and Bovens, 2011^[99]), in which actors commit their account-giving efforts towards ‘buck passing’ activity rather than concerted efforts at meaningful account giving. This results in not only a suboptimal distortion of actors’ efforts, but also obscures the information from which forums make their assessment of conduct, and associated consequences (resulting in evaluative errors (Tetlock, 1992^[12])).

The challenge posed by stretched accountability has a conventional prescription – accountability not sufficiently felt should be addressed through greater clarity and distribution of accountabilities, either by forums or amongst actors (notwithstanding that under complex conditions this can prove exceedingly difficult – see Section 2.1). In practice, the approach to make accountability more present can be pursued by attaching provisions to compel actors to render account (say, through introducing a mandate).

²¹ This is comparable to efforts described in the governance literature that seek to promote self-regulatory norms of behaviour (Majone, 2011^[95]).

²² The other being that of conflicted and diverse expectations – see Section 3.4.2.

Related to the concern presented by thin accountability is the challenge posed by ‘tangled’ webs of accountability (Romzek, 2011_[42]). This is again especially true in the case of collaborative and networked governance arrangements where not only many accountability relationships exist, but are more than likely rather loose and ‘tangled’ – because account-giving can flow in various directions, with adoption of actor and forum roles interchangeably (Romzek et al., 2014_[44]). Within a complex governance environment, however, the prospect that webs can readily and meaningfully be ‘disentangled’ is unlikely – particularly because governance arrangements which are prone to this condition likely experience a high degree of informal accountabilities which could be infeasible to adequately map out in a way that would indefinitely clarify expectations (Considine, 2002_[144]; Klijn and Koppenjan, 2014_[68]).

3.4.2. *Accountability overload*

Just as too little accountability can result in deficits, it goes that too much accountability can result in overloads. This can arise either from being of an overlapping (level) or disparate (range) nature of accounts.

Thick, overlapping accountabilities

Irrespective of the particular instruments employed, account-giving exercises can be susceptible to an additive bias (Lewis and Triantafillou, 2012_[145]). Namely, this is because, over time, new accountability expectations tend to be cumulatively adopted, while expectations are less regularly relaxed or peeled off (Halachmi, 2014_[39]). Moreover, additivity tends to be disproportionately felt as a stressor for actors (for instance, in organisational settings it is well established that managers feel accountable against antiquated expectations long after they have been discontinued) (Hall et al., 2006_[65]).

As Romzek (2000, p. 23_[4]) puts it, in practice, new accountabilities are adopted over time – “without regard to elegance of design or redundancy” – especially in a reactive nature. For instance, an instrument intended to prevent a particular episode of misconduct being repeated may remain in place long after the environment that enabled such misconduct in the first place may have since evolved. In part, this is exacerbated by risk-aversion on behalf of forums in public accountability settings – say, in the case of oversight committees – to ‘wind back’ expectations, like checks and balances, even if such expectations have ceased to remain current, are superseded, or offer little practical use (Power, 1997_[146]).

In turn, accountability expectations can in practice be stacked on top of another, resulting in a thick, layered, and overlapping set of demands (Halachmi, 2014_[39]). This contributes to additional demands being placed on actors, for instance, as accounts may be required to be rendered repeatedly for the same conduct (say, reporting to multiple forums for the same action, or to the same forum in repeated episodes) (Schillemans, 2010_[101]), without necessarily producing any associated gain in information from such accountability exchange(s) (Ebrahim, 2005_[147]). The act of repeated account giving, far from serving to stimulate further reflective behaviour in actors, however, can contribute to disaffection with the account-rendering process (Jos and Tompkins, 2004_[148]), including relegation to being little more than a perfunctory exercise. Moreover, it can fail to improve the decision quality of forums, such as by reducing individual scrutiny in assessment (Hoos, Pruijssers and Lander, 2017_[97]) – perhaps because of an expectation that another forum ‘has it under control’.

In addition to this is the issue of ‘ossification’ (Smith, 1995_[149]; Thiel and Leeuw, 2002_[150]) of expectations – in which an expectation, say, in the form of an evaluative standard, has

become essentially obsolete, yet there is little or no appetite to remove the standard from the expectations set. The result of ossification is two-fold. First, actors engage in efforts unnecessarily in order to meet an out-dated or substantively irrelevant expectation – potentially drawing substantive efforts in their conduct to meet these expectations (which can be unnecessarily costly, burdensome, and inefficient), as well as diluting the relative efforts afforded to more appropriate account-giving activity.

Unplanned redundancy can also be exerted on the functioning of forums, such as through contributing towards ‘information overload’ (Greiling and Spraul, 2010_[151]). This is because actors, consciously or otherwise, can oversupply accounts. This can obscure account assessment through increased search costs of forums that must filter through redundant information, resulting in additional efforts as well as increased probability of evaluative errors (such as through failing to identify the most appropriate information). Moreover, excessive information rendered can further contribute towards evaluative error through the practice of ‘oversharing’ if the information rendered lacks sufficient context (Greiling and Spraul, 2010_[151]), or the account is not communicated credibly.

Largely to blame for unplanned redundancy (Schillemans, 2010_[101]) and obsolescence is the over-emphasis on ‘default’ exchanges of accountability (Bovens and Schillemans, 2014_[94]) – those exchanges that are carried out essentially by default as routine, rather than being carefully considered and calibrated to settings. This can also contribute to a predictability and routine-ness that can undermine the experience of accountability (such as through a conformity bias), by essentially being reduced to a low effort, repetitive exercise rather than an effortful and well-considered exchange of information (Staw and Boettger, 1990_[152]; Tetlock, Skitka and Boettger, 1989_[73]).

Conflicting and diverse expectations

A possible consequence of ‘many eyes’ (Section 3.2.2) is that expectations can be unclear, exceedingly diverse, and potentially conflicting – impacting upon not only the substantive conduct of actors but also impeding on the account-giving experience. Whilst, nonetheless, some degree of diversity of expectations might be entirely appropriate – namely to generate moderating behaviour of actors (reducing tendency to concentrate activity towards extremes) – it can contribute towards episodes of dysfunction (Radin and Romzek, 1996_[108]; Romzek and Ingraham, 2000_[153]).

When faced with diverse expectations from forums, actors are required to commit additional efforts to account-giving, as well as additional augmentation of their conduct, to meet the respective expectations accordingly (with additional efforts commensurate with the ‘distance’ between respective forum expectations). The need to meet a diverse set of accounts, however, can potentially result in distracting actors from their core function (or the core account for which they ought to render). This can be exacerbated when the stakes are exceptionally high and the core function is integral to the actor’s identity (Romzek, 2000_[4]) or it is unclear who the ultimate arbiter of standards is, or ought to be (Rhodes and Wanna, 2007_[154]).

In a famous case illustrated by Romzek and Dubnick (1987_[122]), the Challenger tragedy is partly attributed to the National Aeronautics and Space Administration’s (NASA) accountability demands being distracted away from its core strength – professional standards – by efforts to meet political and bureaucratic accountability demands. The corollaries for education systems are ever-present, as actors (say, schools) face many accountability demands (administrative and the like) which can draw efforts away from their core strength (provision of educational instruction).

A particular case of problematic conflicting expectations is that of the so-called ‘multiple accountability disorder’ (Koppell, 2005_[50]). This condition relates to the experience of actors faced with arrangements that seek to achieve non-complementary ends. The result is that actors may either be paralysed (related to the decision associated with so-called ‘analysis paralysis’ (Ethiraj and Levinthal, 2009_[155])) between choosing between the respective expectations (such as through decision evasion), or try to moderate conduct by steering a course between meeting expectations of each forum. In practice, this is characterised as a ‘disorder’ because by trying to please everyone (or to not upset anyone), neither forums’ expectations are effectively met from charting the balanced, but suboptimal, path.

Indeed, for strategically motivated actors, decision making can be associated with so-called ‘tunnel vision’ (Smith, 1995_[149]) – actors, faced with a range of predictable expectations of forums, consciously pursue the easiest (or most defensible) conduct rather than the optimal course of action. By extension, for forums, the corollary is to seek out the most easily enforceable (least controversial) expectations (Behn, 2001_[109]). In education systems, attempting to satisfy expectations of potentially non-complementary forums – say, school evaluators and parent associations – can present school administrators with this kind of disorder – say, for the former, meeting a nationally consistent standard, while the latter may expect a locally-appropriate standard to apply. This can be accentuated in high stakes contexts, as this can make activity towards reductive goals relatively more attractive.

Social psychologists have highlighted the role of ‘conformity’ in account rendering (Tetlock, Skitka and Boettger, 1989_[73]; Cialdini and Goldstein, 2004_[156]), whereby actors engage in low cognitive effort strategies to augment accounts according to perceptions of what forums’ expectations are known to be – a finding also affirmed in applied experiments, such as in accounting and auditing (Hackenbrack and Nelson, 1996_[137]). This can result, in turn, in ‘rehearsed’ or ‘performed’ account-giving (for instance, Dubnik, 2005_[5]), rather than engaging in the collection and analysis of relevant information for decision making (Lerner and Tetlock, 1999_[61]), including self-critical thinking (Tetlock, 1992_[12]). At the heart of the problem is that adverse conditions can result where forum expectations are fully predictable – and thereby creating opportunities for ‘strategic’, rather than meaningful, account rendering.

4. The ends of accountability

In addition to the many means, policy makers also use accountability to achieve many different ends (Willems and Van Dooren, 2012^[157]). This is made all the more confusing because while policy makers readily identify the role of accountability as ‘the problem’ when interventions do not go to plan, the nominated ‘solution’ also tends to be the establishment of more accountability (Yang, 2012^[158]).

However, by serving as both ‘cause’ and ‘cure’ of challenges (Dubnick and Frederickson, 2011^[159]), efforts to improve accountability run the risk of conflating the different ends that accountability is intended to serve – thus fuelling the varied understandings that people take towards ‘what accountability means to me’ (Dubnick and Justice, 2004^[160]; Dubnick, 2003^[161]). For instance, a commonly observed phenomenon in education systems is that information exchanged in accountability interventions (particularly in terms of performance-related data collection) may be employed towards many different ends, even if the arrangements were purpose-built, intended, or appropriate only for a specific purpose.

In any case, engaging in an accountability exchange is not a costless exercise (Schillemans and Bovens, 2011^[99]). Forums engage in efforts – and indeed financial costs are regularly incurred – to develop, impose and adjust instruments, monitor actors, and manage possible opposition. Accountability imposes costs on actors, such as time and energy to render account, as well as redirecting their behaviour so as to comply with the demands of forums (Kim, 2005^[162]).

In the main, setting up accountability instruments to shape actor behaviour is related to reaping of the so-called ‘promises’ associated with accountability (Dubnick and Frederickson, 2011^[159]; Dubnick, 2002^[163]; Dubnik, 2005^[164]; Dubnick and Frederickson, 2010^[165]) – representing the various motivations in which instruments are expected to serve to enhance certain objectives and meanings to those involved. Yet, promises pursued over-zealously, as well as those that go unfulfilled, can contribute towards dysfunction within accountability exchanges (Bovens, 2009^[118]; Schillemans and Bovens, 2011^[99]).

Some of these promises are intrinsic – that is, they are valued because they are perceived as legitimate and defining characteristics within a political or administrative culture – these are the promises of integrity, democratic legitimacy, and justice. Other promises are considered instrumental – that is, they are valued for what they can accomplish directly – these are the promises of control, appropriate behaviour and performance (Dubnick and Frederickson, 2011^[159]; Bovens, 2010^[3]).

Each of these respective promises are discussed throughout this section – first by elaborating on the intrinsically valued motivations, and then the instrumentally valued motivations.

4.1. Promise of integrity

It is anticipated that account-giving encourages that actors ‘do the right thing’ (Dobel, 1990^[166]) and, for this reason, it is closely related to lofty principles like morality and ethics (Steinbauer et al., 2014^[167]). Indeed, many famous writers have addressed this very issue. For instance, Adam Smith (1759^[168]) supposes that accountability and morality are inseparable (namely that to be moral is to be accountable) and, for Plato, we would all behave unethically in absence of accountability for our actions (Bloom, 1968^[169]).

Beyond the association with morals and ethics, the promise of integrity is more broadly associated with the adherence to prescribed rules [see also, Koppell (2005_[50])]. Efforts to promote integrity and to curb corruption, however, can potentially contribute towards adverse or perverse outcomes if applied over-zealously – potentially resulting in dysfunction (Bovens, 2009_[118]), such as (blind) ‘proceduralism. This can undermine alternative objectives such as the promises of appropriate behaviour and performance – epitomised in the notion that it is ‘not what is expected but what is inspected’.

In schools, accountability instruments for teachers can seek out integrity through, for instance, meeting a set of criteria for professional accreditation. In this case, teachers may render account for their conduct to professional bodies responsible for accreditation or setting standards. In this case, actors may be required to meet specific criteria and follow certain rules against which their rendered account is judged. In addition, school administrators may be called to render account for the use of school finances against the intended purposes to, for instance, provincial governments or departments – perhaps in some cases also to parents.

In practice, forums may be prone to over-zealously applying the principles of integrity in a bias towards negativity and strict compliance²³ (Behn, 2001_[109]), which could effectively reduce discretion of actors, as well as opportunities for learning (Jos and Tompkins, 2004_[148]). This has been associated with the attachment of ‘defensive routines’ by seeking to act in ways that reduce the odds of breaches (typically sticking with the status quo) rather than seeking out and learning what is optimal (Greiling and Halachmi, 2013, p. 400_[170]).

This bias can arise because forums are often encouraged to engage in disproportionate efforts towards seeking out ‘scalps’ – namely by actively searching for breaches, however immaterial, to demonstrate ‘success’. In consequence, actors may augment their behaviour to be disproportionately risk-averse to minimise probability of ‘breaches’ rather than engaging in ‘maximising’ behaviour, particularly when there are high stakes consequences attached.

4.2. Promise of democratic legitimacy

In democratic systems, accountability is a core, if not the defining, characteristic of ‘good governance’ principles (Warren, 2014_[171]; Olsen, 2013_[172]), and, in turn, democratic legitimacy for accountability relationships encompasses a range of components, broadly associated with ‘voice’, such as: representation, participation, transparency, and elections (Dubnick and Frederickson, 2011_[159]). Further, this can articulate to instruments such as petitioning, civil society actions, collective organisation, legal recourses, and demanding access of information to the public.

However, the most iconic accountability in representative democracy pertains to elections [for instance, (Strøm, Müller, Wolfgang and Bergman, 2003_[173])] and the political representation that is attached to this through the ‘democratic chain of delegation’. Through the democratic process, the chain of delegation transfers powers to public officials and ministerial responsibilities, ultimately accountable to the public through elections. Nonetheless, as Romzek (2014_[43]) notes, elections and popular vote can be a relatively

²³ In Behn’s (2001_[109]) influential work, he attaches the term ‘accountability bias’ to this ailment, among others, on the basis that forums may tend towards enforcing compliance standards. This is because it is considered more defensible, less ambiguous, and more readily demonstrable than assessing performance standards.

‘blunt instrument’ for accountability. This is because voting is such a complicated decision making process, undertaken with limited information, limited choices, and only at specified periods, which undermine the ability to effectively hold politicians to account for specific conduct.

Transparency instruments have been a particularly prominent fixture of accountability in education (Rabinowitz, 2018^[174]; Smith, 2016^[175]), in common with governance more generally (Hood and Heald, 2006^[176]). In addition, governance in education has increasingly been categorised in many jurisdictions by decentralisation, which brings with it a multiplicity of stakeholders (participation), and thus, greater demand for negotiation and a diversity of interests to navigate (Burns and Köster, 2016^[49]).

Schools are held to account through institutions such as school boards and parents’ associations (representation). In such instances, a school board requires schools to account for conduct through the avenue of representing community interests. Schools may also render account to interested stakeholders through offering access to information (transparency). In addition, stakeholders may be able to register their complaints and concerns as a means of making schools answerable to expectations (participation).

4.3. Promise of justice

Often invoked by the consequentiality of accountability instruments is the promise ‘to hold them to account’ (that is, to bring about justice for some action) and, in some circles, is synonymous with the so-called ‘off with your head’ notions (particularly in terms of retributive forms of justice). Indeed, this can be related to Koppell’s (2005^[50]) dimension of ‘liability’ (i.e. that actors face consequences for their actions). This marks a point of departure from the promise of democratic legitimacy (having an institutional setting that is consistent with democratic principles) and the promise of justice. This is because the latter relates to whether forums, in practice, *can* (avoiding, for instance, the issue of so-called ‘toothless tigers’) or *do* deliver follow through to consequences.

In addition to simply meting out justice, the promise is also that the accountability instruments apply this justice appropriately (for instance, the penalties which apply to criminal convictions reflect community expectations of justice to suitably hold felons to account) and without exerting undue force or authority (Moore, 2006^[52]). To this end, Dubnick and Frederickson (2011^[159]) note that it is essentially the ‘cathartic value’ that ‘justice will prevail’ that underpins the promise of justice.

The legitimacy of accountability instruments is linked to actors perceiving respective forums as authoritative and committed to justly applying consequences; fairly, proportionally, and objectively assessing accounts; and applying due process. To this end, the perception that an accountability arrangement is fair is associated with, for instance, feeling that a fair hearing has been offered, and that the ‘game is not rigged’ (Levi, 1998^[177]). By contrast, where expectations placed on actors are overbearing (too difficult to feasibly reach or unreasonable consequences), this can result in the forum being perceived as unjust, or holding to an unreasonable standard – ultimately undermining its legitimacy. For instance, in some education systems, the heterogeneity in circumstances that apply – such as in schools with exceptional needs – can undermine legitimacy of expectations, if these circumstances are not reasonably taken into consideration (such as when expectations are unduly standardised, despite heterogeneity).

Relatedly, the perception of ‘being heard’ and having ‘voice’ within the standards of procedural justice – particularly that all parties feel that they have equal opportunities to

present their perspectives – is also of importance (Tyler, 1987_[178]). The notion of a fair hearing is also important for stimulating ‘self-determination’ of actors – namely because it increases actors’ confidence to make judgements regarding their course of conduct, because they can expect to have fair opportunity to justify their actions (Patil, Vieider and Tetlock, 2014_[179]).

In education, negative instruments of this kind include penalties for malpractice or misconduct of teachers, administrators or school leadership, or misappropriation of school finances. In this case, the accountability instrument seeks to ensure that just and equitable outcomes are achieved – say in terms of forums applying sanctions on a school administrator found to have misappropriated public funding – but, moreover, that such an instrument being in place leads actors to behave justly.

4.4. Promise of control

Forums may perceive that some mechanism of control is required to constrain actor behaviour in some way or to limit actor discretion (since if actors were left to their own devices this could deliver undesired outcomes), and to ensure that conduct is undertaken as intended (Gailmard, 2014_[180]; Peters, 2014_[102]). Indeed, the notion of control is often inseparable from accountability (particularly when considering accountability in terms of instruments for regulating behaviour – see Section 3.1). In turn, it is unsurprising that Koppell (2005_[50]) argues that control is the dominant feature of accountability systems.

Nonetheless, it bears reminding that more control does not imply that an actor is more accountable *per se*, and that control in accountability can articulate in a range of ways. For instance, Dubnick and Frederickson (2011_[159]) identify three ways in which control can be exerted in accountability instruments as: substitutes for direct control (such as incentives and choice architectures); complements to direct control (such as reporting and monitoring); or as one among a number of means applied with the intent of controlling behaviour and choices.

Within education, accountability instruments are routinely put in place, for instance, when policy makers impose centralised expectations and general oversight, for instance, of curricular implementation and inspection. In such instances, a forum (the central government) requires that actors (school leaders, provincial governments, line agencies and the like) render account that their conduct is consistent with the forum’s expectations – and possesses the capacity to impose consequences based on the account rendered.

Alternatively, parents exert control when they set expectations of teaching staff and hold them to account through monitoring, as well as with their discretion for exit where feasible²⁴ through demanding account in a controllability sense – that actors (teachers) behave consistently with the forum’s (parents’) expectations (Koppell, 2005_[50]).

When it comes to the promise of control, the key determining factor is the congruence of expectations (between actors and forums) of actors’ substantive conduct. In turn, the distance between expectations is labelled as ‘drift’ – resulting in possible conditions of ‘actor drift’ and ‘forum drift’ (Schillemans and Busuioac, 2015_[181]).

²⁴ For a discussion of implications and conditions for exit (and voice) in accountability arrangements, see, for instance, Paul (1992_[305]).

4.4.1. Actor drift

By extension of so-called ‘agent drift’ in principal-agent theory (Strøm, 2000^[112]; Strøm, 1997^[111]) is the implication that actors tend towards behaviour such as: withholding information, serving their own interests, and generally avoiding accountability to relevant forums (Greiling and Spraul, 2010^[151]). Indeed, accountability relationships, by default, are said to exhibit a tendency towards drift, given that the orientation of control is inherently mistrust-driven (Greiling, 2014^[129]) – namely because if there was complete trust in the actor there would be no need for the forum to introduce an accountability instrument in the first place.

Several factors are likely to attenuate the tendency towards actor drift related to notions of public trust (Kim, 2005^[182]; Greiling, 2014^[129]; Mansbridge, 2014^[183]): credible commitments; perceptions of competence, benevolence, and honesty; the admission of errors and omissions; and engagement in informing evaluative standards. In common across these features is that each element contributes towards more trust within the respective relationships.

Credible commitments

A commitment is ‘credible’ (Levi, 1998^[177]) if intent and action are mutually expected to be consistent (Kim, 2005^[182]) – that is, the trust notion associated with the consistency between what one *says* and what they actually *do* (Lewicki and Bunker, 1996^[131]). In turn, when respective parties have expectations of credible commitments this results in lower levels of uncertainty (Williamson, 1983^[184]).

When information exchanged by actors when rendering accounts is not expected to be credible, this can mean that forums commit additional efforts (say, search costs through requiring additional interrogation of information (Lee et al., 1999^[185]) to be assured of the unbiasedness of accounts rendered (Greiling and Spraul, 2010^[151]; Yang, 2006^[186]).

Similarly, if actors perceive that forums do not make credible commitments this can create apprehension that information provided when rendering account may be misused – which, in turn, can encourage filtering of information when accounts are rendered. This could result, for instance, in biased or incomplete information that is to be assessed (Greiling and Spraul, 2010^[151]).

Perceptions of competence, benevolence, and honesty

Actor’s perceptions of competence, benevolence, and honesty of forums can shape their level of commitment and trust within the accountability relationship (Kim, 2005^[182]). In particular, where actors perceive forums as authoritative or of high status, actors are more likely to refrain from low-effort evasive tactics or defensive behaviour when rendering account (Lerner and Tetlock, 1999^[61]; Mero, Guidice and Brownlee, 2007^[187]). When actors consider their audiences to be authoritative, this reduces tendencies for information to be misrepresented, particularly by enhancing the accuracy of accounts (Mero, Guidice and Brownlee, 2007^[187]), as well as moderating the tendency to misreport conduct – namely, through self-enhancement (inflated sense of one’s own conduct) or to elicit desirability biases (adapting accounts by appealing to perceived desirability) (Deandrea et al., 2012^[188]; Sedikides et al., 2002^[189]). Moreover, accountability to respected audiences can contribute to actors taking account giving more seriously, and, in turn, displaying more responsible conduct (Schillemans, 2016^[14]; Schlenker et al., 1994^[190]).

Perceptions of competence relate to the view as to respective parties' possession of the necessary knowledge and skills to fulfil their functions (Kim, 2005_[182]). To this end, the perception of forum competence is influenced by the forum's technical expertise with respect to the underlying domain of conduct being assessed (put simply, does the forum know what it is talking about?). By extension, trust is lost when forums are perceived to make consistent failures and following discoveries of unexpected incompetence and inadequacies (Kasperson, Golding and Tuler, 1992_[191]), or when forums can be readily dismissed because actors perceive them to be incompetent (Jackson et al., 2015_[192]).

Benevolence and honesty are related to perceptions of 'genuine care and concern' and, similarly, honesty relates to the 'wise and proper use of administrative discretion' (Kim, 2005_[182]). At the heart of the relevance of these notions, according to Fry (1995_[139]) is the perceived congruence of intent (for instance, in terms of 'encapsulated interest' (Hardin, 1993_[193]; Hardin, 1999_[194]) – namely with respect to whether actors perceive forums as being self-interested or if interests are with, and served for, the greater good. To this end, trust is held when actors (forums) perceive that forums (actors) are well-intended (Kasperson, Golding and Tuler, 1992_[191]). For instance, when it is perceived that forums have an ulterior motive (such as using information to a hidden or unknown end), then this alters the information rendered from accounts – such as withholding information or applying excessive risk-aversion in conduct (Greiling and Spraul, 2010_[151]). This, in turn, can contribute towards actors dismissing criticisms made by forums rather than engaging constructively for the purpose of learning (Ron, Lipshitz and Popper, 2006_[195]). On the other hand, perceptions that support may be made available from forums when required can demonstrate a sense of benevolence.

Admission of errors and omissions

The conventional understanding of trust is invariably tied to notions of the 'willingness to be vulnerable' (Mayer, Davis and Schoorman, 1995_[196]), making it particularly important for accountability relations. In environments in which actors feel that they cannot admit errors and omissions (Ebrahim, 2005_[147]), they may engage in withholding or filtering information, which in turn leads to biased or incomplete accounts misrepresenting conduct. Moreover, if there are opportunities to withhold information – say in order to 'cover up' errors – , this can compel forums to engage in additional searching efforts to seek out any 'covered up' errors – which, besides incurring costs, could draw attention away from undertaking due scrutiny based on information that is known (Greiling and Spraul, 2010_[151]).

Engagement in formation of evaluative standards

Actors engaged in the formation of evaluative standards are more likely to be committed to the standards accordingly, perceive them as legitimate, and contribute to greater performance (Linhorst, Eckert and Hamilton, 2005_[197]; Quetzal Tritter and Mccallum, 2006_[198]; Damgaard and Lewis, 2014_[199]). Moreover, standards that are formed in collaboration with actors (Bovaird, 2007_[200]) are less likely to face resistance from actors due to a greater degree of legitimacy that is embedded, and also tend to be more feasible (within the capability of actors to achieve) and relevant (appropriately reflects actor's perception of substantive conduct) than standards that are formed without actor engagement.

4.4.2. Forum drift

While ‘actor drift’ is abundantly discussed in the literature, relatively little attention has been devoted to ‘forum drift’ – namely whereby forums fall short in their account-holding roles (Schillemans and Busuioc, 2015^[181]). Indeed, when accountability is felt sufficiently, forums undertake greater scrutiny (Pennington and Schlenker, 1999^[201]) – for instance, with respect to the auditing of auditors (Hackenbrack and Nelson, 1996^[137]; Koonce, Anderson and Marchant, 1995^[202]). In such occasions, a deficit in accountability can arise in the event that actors render appropriate account but it falls on ‘deaf ears’ – namely forums choose not to hold actors accountable, disregard apparent wrongdoings, or are disinterested in some or all conduct.

Obsolescence of expectations

Forums can be prone to relative invariability in their assessment – that is, the expectations that they hold may only be adjusted infrequently (for instance, when legislative instruments are introduced). This can result in evaluative standards potentially being ‘out of touch’ or no longer contemporary or substantively relevant (Smith, 1995^[149]; Bovens and Schillemans, 2014^[94]). When actors perceive that forums are pursuing obsolete expectations, they are viewed by actors as ‘out of touch’ (incompetent). In turn, they are inclined to engagement in equivalently low effort, perfunctory account rendering (that is, when it is evident that a forum is not genuinely engaged in the actor’s conduct, they respond by allocating little effort to account rendering) (Lerner and Tetlock, 1999^[61]; Green, Visser and Tetlock, 2000^[203]).

Zone of indifference

Whilst the principal-agent logic centres around the tendency of agents to veer from principals, in practice forums can also gravitate towards actors. For instance, the management and regulatory governance literature have long recognised the potential for such conditions as ‘blind-eye’ management and ‘regulatory capture’ (Levine and Forrence, 1990^[204]). At the heart of these matters is that relatively low efforts are dedicated to scrutiny – namely, for the former, managers can be unmotivated to perform disciplinary action on underperforming employees, and for the latter, an exceedingly complacent relationship between the regulator and the regulated may result in favourable relaxation of standards. This arises because of an excessive ‘zone of indifference’ (a term employed largely in the marketing and management literatures, but has its original articulation in Barnard (1938^[205])) may persist, whereby the ‘zone’ in which a forum is willing to pursue a ‘no action’ strategy of scrutiny has extended beyond an appropriate level. This is also a potential perverse result from ‘learning’ in accountability – whereby actors and forums simply learn ‘how to get along’, including how to mutually ‘game’ the system (Schillemans and Smulders, 2016^[72]).

Symbolic accountability

Forums may ostensibly carry out their assessment duties, though they may lack sufficient ‘teeth’ to apply appropriate consequences for malpractice to compel actors towards desired conduct. In turn, the forum may provide accountability ‘in name only’ – say, by nominally carrying out inspections and the like, without the intention, willingness, or ability to regularly and reliably carry out their substantive function (say, to enhance standards or prevent malpractice) (Halachmi, 2002^[206]).

For instance, forums conducting compliance or inspection reports may revert to ‘ticking and flicking’ (or resorting to other low effort strategies, like downgrading breaches to avoid additional efforts such as follow up inspections and the like), rather than engage in efforts to administer consequences. Similarly, forums can deliberately discourage efforts to render full accounts actively, by defensively maintaining ‘plausible deniability’ in the event of possible scrutiny (Greiling and Halachmi, 2013_[170]).

Myopic assessment

Because detailed and effortful consideration is required in order to assess and maintain longer-term, complex, or ‘high stakes’ accounts, forums may be motivated to place an emphasis on short-term or easy to monitor accounts and set expectations accordingly²⁵ (Ebrahim, 2005_[147]). For instance, rather than making holistic assessment of conduct, a forum may elect to undertake a more summative assessment, which essentially lets the actor ‘off the hook’ by narrowing the conduct to be assessed (related to ‘measure fixation’ (Smith, 1995_[149])).

Ownership of evaluative standards

In common with actors, when forums assume ownership of evaluative standards, they engage in additional efforts to ‘guard’ standards accordingly more completely (less likely to ‘bend the rules’ or willingness to ‘follow the letter of the law’) and regularly (less likely to ‘let it slide this time’) (Schillemans, 2013_[207]). By contrast, forums tend to be less engaged in policing standards that are formed by others – for instance, those standards inherited from past administrations – (Schillemans and Busuioc, 2015_[181]), which can reduce the seriousness with which actors take the evaluative standards.

4.5. Promise of appropriate behaviour

It is presumed that accountability instruments can ensure that the correct (or most appropriate) decisions are made by actors, which could be associated with, for instance, ethical or professional conduct – essentially by recourse to reflection of what one is ultimately expected to do in a given situation. It is related to, but departs from, promises of control and integrity. The relationship between the promises of integrity and appropriate behaviour are akin to Koppell’s (2005_[50]) observation of tensions between responsiveness (fulfilling substantive expectations) and responsibility (following the rules). One reason why control and responsiveness are not identical, moreover, is that agents are expected to not necessarily strictly follow what is ordered of them, but instead to exercise appropriate judgement under conditions with which they are confronted (Mulgan, 2000_[90]; Friedrich, 1940_[208]).

This is especially true where expectations are subject to changing rules and the like. Within education, school leadership render account to their peers (other school leaders), along with teachers, other professional staff, as well as students and parents. The account rendered is with regard to the professional judgement exercised according to the expectations of the relevant forum(s).

²⁵ This is related to the ‘awakenings syndrome’ in performance monitoring and management literature (Halachmi, 2005_[219]) – associated with the observation of dwelling on the short term while disregarding future goals.

In addition, teachers' professional standards are held to account when forums (departments, students, parents, school leadership) have the opportunity to assess teachers' pedagogical approach applied in classrooms. This can be especially challenging for teachers when rendering account for their conduct with respect to their delivery of delicate course content – given that there may be relatively more 'space' for professional judgement to be applied. They may, at the same time, be required to dedicate attention to, for instance, religious sensitivities, sex education and the like – that is, actors are held to account by forums for the appropriateness or otherwise of their conduct.

Similarly, professional staff in schools working with students with high needs may face especially challenging accountability conditions in terms of meeting diverse needs of students and parents and rendering account accordingly. Moreover, teachers and school leaders may be confronted with managing, for instance, 'zero tolerance' bullying policies with 'letting kids be kids' preferences of stakeholders. In such instances, forums hold the actors to account not for their adherence to rules necessarily (as with integrity accountability instruments) but towards the appropriateness of schools in implementation of relevant policies.

As is apparent, the contention of appropriate behaviour as desired from an accountability exchange, hinges greatly on the interpretation of expectations. For this reason, social psychologists have devoted attention to the behaviour of actors when accountability is 'expected' and 'unexpected' in advance of having carried out the associated conduct to be assessed (Lerner and Tetlock, 1999^[61]; Pollmann, Potters and Trautmann, 2014^[209]).

Put differently, when actors are equipped with 'pre-decisional information' (they are informed *in advance*) of accountability expectations, this facilitates commitment to a course of conduct. Moreover, provided that the expectations are perceived as relevant and feasible, actors can be "motivated to think in relatively flexible, multidimensional ways" and exert "pre-emptive self-criticism" (Tetlock, Skitka and Boettger, 1989^[73]). Moreover, pre-decisional accountability reduces bias judgements resulting from the activation of 'automatic processing routines' (Tetlock, 1985^[210]), as well as biased tendencies of 'first impressions' (a particular type of belief perseverance; see Section 5.2.1) (Tetlock, 1983^[211]).

In contrast, being held retrospectively to account for conduct undertaken can result in engagement in less open-minded self-critical reflection and more defensive views – namely, through pursuing a 'defensive bolstering' strategy (Tetlock, Skitka and Boettger, 1989^[73]). The result is that efforts are concentrated in defending a course of action taken rather than a meaningful, self-critical consideration of conduct.

4.6. Promise of performance

Accountability is commonly considered with respect to the promise of inducing an increase in individual, team, or organisational performance (Camm and Stecher, 2010^[212]). This can diverge in substance from what is ultimately, and substantively, expected (that is, the promise of appropriate behaviour) but represents instead some conduct with respect to a previously set measure of 'performance'.

Namely, this is because a change in conditions might demand that an actor adapt to the revised circumstances (appropriate behaviour's notion of responsiveness) in some manner, but this may not be reflected in the performance expectations themselves. Accordingly, arguments that accountability can instil 'performance for performance sake' are articulated

by critics. For instance, setting performance expectations can result in actors simply seeking to replicate precisely (and perhaps only) what is contained in their target.

Scholars have long drawn attention to a propensity towards unintended consequences arising from a host of performance-motivated accountability efforts (Christensen and Lægreid, 2015_[213]) and indeed to performance monitoring more generally (Hood, 2007_[214]; Hood, 2006_[215]). In particular, there are three key pathologies observed in the wider accountability literature regarding the promise of performance. These relate to:

- The lack of evidence linking accountability pressures to actual performance increases, labelled ‘accountability paradox’ (Thiel and Leeuw, 2002_[150]).
- Accountability pressures prompting actors to fixate on performance as measured rather than improving their substantive conduct, which has been coined ‘accountability trap’²⁶ (Bovens, 2010_[3]).
- Conforming to what is known and focussing on immediate performance, rather than deviating from the known in search of means to potentially sustain performance over time, related to the problem of so-called ‘accountability ping-pong’ (Tetlock and Mellers, 2011_[216]).

4.6.1. The ‘accountability paradox’

Perhaps the most oft-cited observations in the literature and practice regarding the promise of performance has been the seemingly paradoxical lack of a track record in delivering sustained performance improvement (Dubnik, 2005_[164]; Thiel and Leeuw, 2002_[150]; Bouckaert and Peters, 2002_[217]; Halachmi, 2002_[218]). Research brought forward two possible explanations for this.

First, actors may divert resources away from improving their substantive conduct towards avoiding consequences, such as selective reporting, blaming, and justificatory behaviour. Accountability may thus act as a ‘breaker’ to performance increases, rather than ‘driving’ performance (Halachmi, 2002_[218]; Halachmi, 2005_[219]).

Second, improvements may be slowed down or halted by accountability instruments, because actors choose to merely mimic conduct known to increase measured performance, rather than incorporating the logic that underpins the performance improvement in the first place, thus undermining an intended learning function (related to so-called ‘second-loop learning’ in particular – see further Section 5.2.2). This is because it can be difficult to recreate the particular stimulation of innovation associated with a given performance improvement, outside of the specific context in which it originally arose. Namely, because it is rarely retrospectively clear exactly which processes had been undertaken, in which order, by whom, and the like, that generate a particular desired performance outcome – particularly as some performance improvements are the result of unique confluences of factors and cognitive efforts²⁷ (recall from complexity literature – pages 12-14).

²⁶ Authors also have tended to label this predicament as ‘performance paradox’ in some literature (Thiel and Leeuw, 2002_[150]).

²⁷ For a further discussion of innovation-inducing learning and accountability, see, for instance Frederickson (2003_[306]).

4.6.2. *The ‘accountability trap’*

Whilst the accountability paradox pertains to not meeting the desired performance expectation prescribed through an accountability instrument, ailments are also associated with the overzealous adoption of performance expectations (Thiel and Leeuw, 2002_[150]). In particular, this is manifested in the so-called ‘accountability trap’ (Bovens, 2010_[31]), which describes the condition of ‘measure fixation’ in the performance monitoring literature (Smith, 1995_[149]).

Research argues that the more frequently or intensely efforts are employed through accountability expectations, the more that actor conduct converges towards strict accordance to the known expectations of the relevant forum. This results in ‘goal displacement’ where actors’ interpretation of the forum’s goals displaces the actor’s goal of improving substantive conduct (Smith, 1995_[149]) (largely what is implied in the promise of appropriate behaviour). In the extreme, when actors can fully predict consequences or rewards, such as future budgetary allocations based on a set of assessment criteria, actors can manipulate accounts in order to achieve strategic advantages (known as ‘gaming the system’ (Hood, 2002_[143])).

Instruments that incorporate the principle of randomisation (see Section 3.1.1) can go some way to addressing issues relating to goal displacement – namely because the inability to predict which out of the possible indicators that are recorded will be taken as the crucial ones, and when, by those appraising them, or what weight will be put when on which indicator (Hood, 1998_[91]). In turn, actors are encouraged to pursue moderated conduct, as there is no incentive to distort behaviour strategically.

4.6.3. *The ‘accountability ping-pong’*

Another possible reason for perverse results of performance accountability interventions comes from the fixation on the present at the expense of the future (Chang et al., 2017_[220]) – related to what Halachmi (2005_[219]) has termed the ‘awakenings syndrome’ (particularly with respect to equating the achievement of preliminary indications towards desired results with achieving ultimate success). More broadly, there is also an underlying tension between achieving “today’s goals and adaptability for the future” (Simons, 2005, p. 10_[221]). This is because actors are held to account today for achieving short-term results (usually retrospectively), but fixating on these can mean that actors neglect what is required for performance going forward, as forum expectations, nor the context, is likely to remain fixed (Chang et al., 2017_[220]) (recall from complexity’s implications of time interdependence – page 12).

To this end, a dilemma in the literature highlights the trade-off between exploitation and exploration: “balancing the competing goals of developing new knowledge (that is, exploring) and exploiting current competencies” (Levinthal and March, 1993, p. 95_[222]). Despite efforts towards balancing these, however, it has been noted that there is, in practice, a tendency to oscillate between suboptimal degrees of exploitation and exploration, in what Tetlock and Mellers (2011_[216]) call “repeated games of ‘accountability ping-pong’”.

In turn, excess exploitation can result in conformity to processes (in turn, producing ‘mindless bureaucrats’ sticking too close to the rules (Patil, Tetlock and Mellers, 2017_[223]; Patil, Vieider and Tetlock, 2014_[179]) – see also Section 4.1). This is associated with so-called ‘task revision’ (Staw and Boettger, 1990_[152]) – persisting with the known at the expense of potentially better alternatives – thus stymying possible adaptation and learning.

On the other hand, excessive exploration can lead to unnecessary deviation (in turn, producing ‘reckless deviants’ who stray from rules known to be putatively effective) (Patil, Tetlock and Mellers, 2017^[223]; Patil, Vieider and Tetlock, 2014^[179]). Since excessive deviation results in unneeded experimentation with alternatives (March, 1991^[224]), this limits actors’ reliability and efficiency (Canales, 2013^[225]), as well as prompting overly ambitious behaviour rather than what is ‘doable’ (Sitkin et al., 2011^[226]).

In any case, accountability which only produces a temporary shift in performance (exploitation) in order to meet some expectation, at expense of some possible learning (through exploration), can be counterproductive by failing to equip actors with the ability to sustainably perform into the future. At the same time, exploration is not always fruitful, unless learning or innovation is indeed unlocked. This has also stimulated an ongoing debate in organisational studies with respect to the relative virtues of process and outcome accountabilities given that studies have produced contested findings in relation to performance and learning (see further in Section 5.2.2).

4.7. Promises and harsh realities

As with most promises in life, our lived experiences in the real world do not always accord neatly with expectations (Dubnick and Frederickson, 2010^[227]; Dubnick, 2002^[163]). This is made worse because the experiences of well-meaning accountability interventions producing unanticipated results often loom large in the memory of stakeholders and practitioners alike. As Frink and Klimoski (2004, p. 1^[228]) put it: “strewn across the landscape [...] are the wreckages of accountability failures”. In turn, policy makers are cautioned that while accountability is a ‘good thing’, it should be – like everything that is good for us – employed in moderation. This pertains to both substance – that is, the magnitude that is felt – and consequences – that is, in terms of the stakes attached (Bovens, Schillemans and Hart, 2008^[121]).

There are two possible reasons why accountability may not live up to its promises. First is that the weight of accountability is inappropriate (recall from Section 3.4). Second, a specific promise can be compromised when pursued in combination with other promises, such that it may be prone to potential incompatibilities from conflicted motivations.

4.7.1. Unrealised promises

The notion of unrealised promises is hardly unfamiliar in the study of accountability. Table 5 compiles the discussed promises of accountability along with their possible suboptimal outcomes – either as a result of over-zealously applying them (too much accountability felt) or when pursued with too little effect (too little accountability felt) (Bovens, 2009^[118]). In common with the approach developed in Section 3, the weight in which accountability is felt by participants is important in contributing towards possible ailments (Section 3.4). Importantly, this can have differential influences on the experience of actors and forums and across various promises (see Table 5).

Table 5. Suboptimal promises and realities

Common ailments associated with suboptimal accountabilities

Promise		Too little accountability felt	Too much accountability felt
Control	Actor	Actor drift Actors that are mistrustful of forums can withhold information, serve their own interests, and generally avoid or thwart full and complete account rendering.	Goal displacement Actors' goals become completely convergent with forums, irrespective of the relevance of expectations (following blindly) or commitment demanded of them (blind loyalty).
	Forum	Forum drift Accounts are rendered but 'fall on deaf ears' because forums choose not to hold actors accountable, disregard apparent wrongdoings, or are disinterested in conduct.	Oversight fatigue Forums are burdened by being constantly engaged in monitoring and surveillance of actors to ensure that their conduct does not err.
Integrity	Actor	Symbolic accountability Actors know the rules but do not observe them because they are confident they will not be 'found out' for non-compliance. In turn, rules are symbolic and 'in name only'.	Blind proceduralism Actors feel they must not 'draw outside the lines' and blindly follow rules without question of the relevance or appropriateness of the rules.
	Forum	Blind eye Forums are not committed to the enforcement of rules and, so, turn a 'blind eye' to non-compliance or carry out 'tick and flick' exercises.	Accountability bias Forums seek out 'scalps' irrespective of materiality of possible breaches. There is a bias towards negativity and strict compliance.
Appropriate behaviour	Actor	Unexpected accountability Actors find out too late, or not at all, what the expectations for their conduct are. As a result, their conduct is not responsive in the manner that is expected of them.	Predictability of expectations Actors' conduct is prescribed by forums precisely, without room for discretion based on circumstances or actors' judgement. Actors apply 'if-then' conduct and are disinterested in learning or appropriating expectations based on conditions (nonresponsive).
	Forum	Rudder-less Forums set ill-defined expectations for actors, leaving them to largely 'work it out for themselves'. Obsolescence/out of touch Forum expectations are obsolete or 'out of touch' because they do not reflect reality 'on the ground' for actors.	Moving goalposts Expectations are constantly changing which can obscure what forums ultimately want, making it highly demanding for actors to meet expectations. This can be exacerbated when there are conflicted and diverse expectations set by forums.
Democratic legitimacy	Actor	Stretched accountability As more actors are engaged, the accountability felt by a given actor is further diluted amongst the 'many hands'.	Tangled webs Many actors negotiating and interchanging roles confuses expectations and blurs accountability lines (balancing representative and participative democratic principles).
	Forum	Black box A dearth of transparency creates a perception of secrecy and that decisions are reached in a vacuum.	Open book Forums are committed to transparency irrespective of possible sensitivities that this can expose.
Justice	Actor	Toothless tiger Actors are confident that they will not be culpable (identified as responsible) or liable (answerable) for indiscretions based on what they know of their forum.	Risk aversion and information filtering A prevailing 'off with their heads' notion or prospect of excessive penalties ('high stakes') can prompt actors to 'play it safe' or render accounts circumspectly (especially when it comes to disclosing possible misconduct/errors).
	Forum	Unjust trust Forums do not undertake due scrutiny of actor's conduct before conferring trust upon them. Misdeeds unpunished Wrongdoing goes unpunished or there are 'slaps on the wrists' which undermines the legitimacy of the system.	Unjust distrust Forums do not confer trust in actors despite having every reason to do so, leaving them feeling prejudiced. Heavy-handedness Forums mete out disproportionate penalties without due consideration of possible mitigating circumstances.
Performance	Actor	Accountability paradox Rather than focussing efforts on improving conduct, actors divert attention to 'performing', such as 'fudging the numbers', playing blame games, or preparing excuses.	Accountability trap Actors fixate on performance as measured rather than improving their substantive conduct or learning and exploring if measures really reflect better performance.
	Forum	Anything goes Forums are disinterested in substantive performance, instead ensuring that everything appears 'good on paper'.	My way or the highway Forums demand 'results at any cost' without attention to actors' particular needs/resources or what really matters in practice.

Accountability that is motivated towards control from an accountability exchange can suffer from too little felt accountability. In turn, this can result in too little commitment towards the relationship – and in turn to drift of actors and/or forums – namely because actors do not feel sufficiently compelled to the interests or expectations of forums, and forums may err from carrying out their substantive functions. By contrast, excessive control can eliminate agency of actors, as well as require excessive efforts of forums – such as resulting in monitoring fatigue, or ‘burning out’ forums.

In the case of the promise of integrity, when excessively felt, this can result in blind proceduralism driven by the ‘accountability bias’. This can generate a significant burden on forums (searching for scalps) as well as potentially distortionary behaviour (overlooking, say, difficult-to-substantiate indiscretion in favour of more readily substantiated indiscretion) as well as potentially perverse conduct in itself (such as breaking rules – say privacy or the like – in order to catch culprits). For actors, excessively pursued integrity similarly can distort conduct towards following the rules blindly (same way without consideration of appropriateness) and uniformly (same way without consideration of context or participants). Yet, when insufficiently felt, the rules embedded in the accountability instrument can lack conviction – such as inspection regimes that are symbolic only.

In the case of accountability that is motivated by fostering appropriate behaviour, too little accountability felt can result from actors being afforded a sense of discretion to decide upon course of action, however they may be lacking *ex ante* guidance or direction – say, in how and what to do – or in *ex post* evaluation of how discretion is exercised – say, providing corrective feedback²⁸. In turn, actors can feel blind-sided because there is too little formative expectations set with respect to how the conduct should occur in given situations – this is the perception in actors of being ‘rudder-less’ with insufficient sense of direction from forums. On the other hand, where the promise of appropriate behaviour is too heavily felt this can result in ossification of expectations – and indeed in predictability of forums – such that conduct can be distorted to precisely meet known expectations in a routine and predictable pattern.

Democratic legitimacy that is felt too greatly in accountability can result in an overzealous approach to transparency, such that forums may be overloaded with information, or could have the effect of undermining other objectives. In addition, the overzealous invitation of representation and participation can stimulate the ailments associated with ‘tangled webs’. On the other end, democratic legitimacy can be felt too weakly such that, say, excessive participation dilutes accountability – in turn, stimulating the ailments associated with thin, stretched accountability. In addition, a dearth of transparency can create a perception of secrecy and that assessments and consequences are reached in a vacuum.

The promise of justice can be felt too greatly in accountability, resulting in the ‘off with their heads’ motivation and a prevailing sense of universal mistrust of actors (including those that may be justifiably entrusted). In turn, the promise of justice that is lightly pursued can result in unduly entrusted putatively untrustworthy actors, as well as give rise to the ‘toothless tiger’ kind of forum, whereby the instruments for justice exist but have their legitimacy hollowed out.

²⁸ In this regard, approaches such as that offered by the ‘responsive regulation’ framework further establish concepts to this effect (Ayres and Braithwaite, 1992_[248]).

Finally, with respect to the promise of performance that is too little felt, then performance accountability does not exert the learning dynamics, nor improvement culture, that is requisite for delivering on the promise of performance. On the other hand, overzealous pursuit of the promise of performance can result in a ‘results at any cost’ approach. There is some inherent risk that such an approach may well stimulate performance in the short run but may be less clear that this fulfils the learning function of the promise of performance.

4.7.2. Compromised promises

Promises of accountability can also come into conflict with one another (Peters, 2014_[102]; Koppell, 2005_[50]), when attempting to pursue varied ends – particularly where accountability means different things to different stakeholders. In this way, promises are compromised because, in attempting to achieve some particular end, there is potential to undermine another. Table 6 consolidates the two-way combinations²⁹ in which promises can be compromised, with each column representing efforts to pursue each respective promise. The adjacent rows indicate the ways that each promise can be potentially compromised when trying to simultaneously pursue the former.

²⁹ It bears noting that there is potentially many possible combinations in which promises could be compromised when working towards multiple ends, but for illustrative purposes, this is constrained to two-way combinations.

Table 6. Potential compromised promises based on two-way combinations

	Integrity	Democratic legitimacy	Justice	Control	Appropriate behaviour	Performance
Integrity		Efforts to invite multiple perspectives for conduct can be undermined by standardising rules.	Maintaining fairness and proportionality can be constrained when rules are strictly adhered to.	Efforts to compel conduct can be undermined by obligations to work 'by the book ('do as I say' vs 'do what it says to do')'.	Efforts to promote responsiveness can be constrained by efforts towards reaching rule-determined decisions.	Performance can be constrained from optimal outcomes through following strict adherence to rules.
Democratic legitimacy	Efforts to achieve clarity in rules can be undermined by efforts to foster inclusion and plurality of perspectives.		Arriving at fair and proportional consequences for conduct may require reaching a judgement that disagrees with a consensus perspective on conduct.	Efforts to control actors' conduct can be diluted by inviting multiple interests into decision making.	Efforts to promote responsiveness can be undermined by adopting conduct based on what is popular.	Performance can be compromised by following consensus courses of action.
Justice	Efforts to apply rules with strict consistency can be undermined by efforts to avoid blanket rule enforcement or to ensure the 'punishment is fit for the crime'.	Efforts to embrace multiple perspectives can be suspended by prosecuting a fair and objective concept of justice.		Efforts to exert control can be constrained by the limits of legitimate and just use of authority.	Efforts to promote responsiveness may be curbed by risk aversion to avoid possible sanction.	Performance can be compromised when trying to reach the fairest course of action rather than necessarily the optimal.
Control	Efforts to prevent 'bending' of rules can be undermined if/when the prescribed rules come into conflict with forum expectations.	Efforts to incorporate multiple voices may be undermined by a forum exercising their control by drowning out of respective voices.	Delivering just and legitimate use of force can be undermined by efforts to exercise control excessively.		Efforts to promote responsiveness of conduct can be compromised by efforts to provide directives on what to do.	Performance can be hampered by the need to report to directors or going out of the way to demonstrate that conduct accords with forums' expectations.
Appropriate behaviour	Efforts towards adherence to rules can be undermined by efforts to adapt to circumstances as appropriate.	Efforts to foster participation can be undermined by efforts to promote adaptiveness of conduct to unique and individual circumstances.	Efforts to objectively assess accounts can be at odds with subjective interpretations of appropriateness of conduct.	Efforts to prevent actors from straying from control may be irreconcilable with also fostering discretion for actors.		Performance can be curbed when actors are compelled to do the right thing in a given situation, which may not accord with performance measures.
Performance	Efforts to impose the strict following of rules can be undermined by efforts to identify more efficient or effective ways of doing things.	Efforts to promote collective decision making can be undermined by efforts to deliver optimal performance outcomes.	Efforts to deliver fairness and justice can be undermined by overbearing expectations of performance.	Efforts to compel strict commitment of actors to expectations undermines opportunities for actors to learn.	Efforts to promote responsive conduct may be undermined by efforts to encourage doing 'anything it takes' to achieve expectations for performance.	

5. Learning and meaning in accountability

An often unspoken promise from accountability is the facilitation of learning (Greiling and Halachmi, 2013_[170]). This tends to be implied in performance monitoring regimes and evaluations, for example, with accountability efforts flowing through to school improvement. Yet, it is not always made clear to those involved, nor is it always explicit through which mechanisms learning is intended to occur.

In addition, because of the occasional emphasis on fault finding – and from time to time, blame attribution – the opportunity for learning can end up being paid little more than lip service. Nonetheless, most would agree that learning from accountability offers a “more promising approach to attain improved results than simply finding mistakes and pointing an accusatory figure” (Greiling and Halachmi, 2013, p. 382_[170]).

This section first looks at different arguments for an emphasis on accountability for processes or outcomes. It explores ways to avoid a trade-off between the two approaches. Second, the section explores insights from social psychology literature for facilitating individual learning under accountability – largely focussed on the conditions of a self-critical attitude by the actor.

5.1. Leveraging benefits from both process and outcome accountability

Independent of the specific means, forums may hold actors accountable for decisions and practices undertaken (process accountability) or for the substantive achievements coming out of actors’ practice (outcome accountability) (Beach and Mitchell, 1978_[229]). In education, process accountability could take the form of setting instructional guidelines for teachers, whereas outcome accountability instead sets a student achievement expectation, but may allow more freedom in how teachers undertake their instruction.

Whether it is more appropriate to hold actors accountable for adhering to processes or for achieving outcomes (or which specific combination thereof) depends on which promises of accountability are prioritised – but also how complex the environment and desired outcomes are and how much emphasis is placed on actors’ learning.

5.1.1. *Process accountability*

Accountability for processes reinforces consistency, because actors render account in terms of how they generated outcomes, including justification of the efforts and strategies that are employed (Chang et al., 2017_[220]), with a focus on “the soundness of the procedures” and with “minimal weight given to the accuracy of the outcomes” (Patil, Tetlock and Mellers, 2017, p. 283_[223]). Process accountability is most suited to conditions in which “decision rules are effective, and tasks are unpredictable to the point where exploring new rules is futile” (Patil, Tetlock and Mellers, 2017, p. 283_[223]) – hence making exploitation of what is already known more appropriate than exploration (recall from Section 4.6.3).

Process accountability has earned considerable support within the psychology literature (De Dreu, Koole and Steinel, 2000_[230]; Brtek and Motowidlo, 2002_[231]). This is because of the appropriateness of holding actors accountable only for what they have sufficient control over (Lerner and Tetlock, 1999_[61]; Tetlock and Mellers, 2011_[86]), and not unduly rewarding actors for outcomes not associated with their own actions (such as luck) (Bertrand and Mullainathan, 2001_[232]). Moreover, other desirable effects include: improved judgement quality (Langhe, Van Osselaer and Wierenga, 2011_[233]); more accurate self-perception of performance³⁰ (Siegel-Jacobs and Yates, 1996_[234]); and greater opportunities to withdraw from a failing course of action (Tetlock, 1985_[210]) because defective practices are more readily identified and replaced.

Yet, under conditions of complex decision-making, or situations demanding improvisation, process accountability is found to be relatively less effective than outcome accountability (Langhe, Van Osselaer and Wierenga, 2011_[233]; Skitka, Mosier and Burdick, 1999_[235]). In addition, despite the potential cognitive benefits of process accountability, these can be diminished by perceptions of unwarranted ‘surveillance’, as accountability procedures can employ an overly ‘watchful eye’ (Enzle and Anderson, 1993_[236]) – which can signal a lack of confidence in actors’ capabilities, and thus stimulate perceptions of distrust (Mansbridge, 2009_[130]).

5.1.2. Outcome accountability

While the psychology literature has tended to favour process accountability, the management and organisational studies literature have lent support to outcome accountability (or the preferred term in this literature – ‘accountability for results’). Under outcome accountability, actors render account according to “accuracy or the achievement of end-state goals, with minimal weight on the means used to reach those goals” (Patil, Tetlock and Mellers, 2017, p. 283_[223]). This is most appropriate when decision rules are contestable, and “tasks are sufficiently predictable that there is considerable room for improving on the rules” (Patil, Tetlock and Mellers, 2017, p. 283_[223]).

Because of the scope afforded to actors in terms of discretion and exploration, outcome accountability is considered to be relatively more supportive in fostering innovation and adaptive behaviour (Simons, 2005_[221]). This is because it can serve to incentivise actors to “challenge standard practices to effectively tackle novel situations” (Patil, Tetlock and Mellers, 2017, p. 283_[223]), including questioning the appropriateness of designated processes (Skitka, Mosier and Burdick, 2000_[237]) – associated with so-called second-loop learning (see Section 4.6.1). Thanks to the potential of unlocking ‘ingenious analytic strategies’ beyond adherence to processes (Wilson, 1989_[238]), this has led some researchers to conclude that outcome accountability can make actors’ conduct more likely to meet the substantive expectations of forums (Eisenhardt, 1989_[239]).

Yet, the conditions of outcome accountability are also regularly contested. For instance, critics point out that following a ‘no questions asked’ approach can incentivise ‘corner-cutting’ behaviour. In addition, the potential benefits of outcome accountability can be limited to very specific conditions. Namely, to successfully nurture actors’ innovation, outcomes must be genuinely controllable and achievable. This can be dubious in education settings, given the range of potential confounding factors in educational outcomes.

³⁰ This is referred to as ‘calibration’ in psychology research (Keren, 1991_[321]) and related applications (Budescu and Du, 2007_[322]).

Moreover, a focus on outcome accountability can obscure attention from the very processes that are responsible for generating innovations in the first place. In turn, this can undermine efforts to reproduce purported successful performance (Douglas and Judge, 2001^[240]) – related to the accountability trap (Section 4.6.2). The level of ambiguity within evaluative standards (Spreitzer, 1996^[241]), the correlation between actors’ effort and the outcome, as well as the opportunity for justifications of decision-making, can contribute to various considerations when assessing the utility of outcome accountability instruments.

5.1.3. Overcoming trade-offs between process and outcome accountability

It is evident that a complete emphasis on either process or outcome accountability tends to be suboptimal and distort conduct (Tetlock and Mellers, 2011^[86]). Instead, it is argued that the particular instruments and their implementation are more relevant, along with the perceptions and attitudes of participants (Patil, Vieider and Tetlock, 2014^[179]).

There are recognised trade-offs between process and outcome accountability relating to the dual objective of encouraging both compliance to essential processes, while also fostering the rationalisation and redesign of redundant or obsolete practices (March, 1991^[224]; Levinthal and March, 1993^[222]). More concretely, Patil, Vieider and Tetlock (2014, p. 78^[179]) specify this ‘design dilemma’ as follows, to:

- “implement process accountability to minimize variance in decision-making and increase reliability (ensuring some control over how decisions are made), but run the risk of prolonged reliance on deficient practices with little regard for outcomes”, or
- “implement outcome accountability to encourage attention to actual outcomes (ensuring some innovation and flexibility), but run the risk of encouraging gaming of poorly understood metrics”.

One suggested way to address the ‘design dilemma’ is informed by research into psychological empowerment within the management literature (Conger and Kanungo, 1988^[242]; Spreitzer, 1995^[243]; Spreitzer, 1996^[241]). This is because ‘empowering experiences’ – the process through which empowerment is conferred from a forum to an actor – is found to stimulate, in actors, a sense of self-efficacy (Conger and Kanungo, 1988^[242]). This follows from the observation that “forms of process and outcome accountability that ‘empower’ employees are more likely to stimulate innovation, whereas those that ‘disempower’ employees are more likely to yield perfunctory compliance (if not passive or active resistance)” (Patil, Vieider and Tetlock, 2014, p. 76^[179]).

Two extensions in the literature provide insights for how to attempt to overcome trading off accountability for process and outcomes. One option is to segment actors, contingent on a selection of those for which outcome accountability is most appropriate and those which processes is most appropriate (Mansbridge, 2009^[130]). In this case, selected actors are entrusted with autonomy while others are not (Mansbridge, 2014^[183]). Another option is to essentially hybridise process and outcome accountabilities (Tetlock and Mellers, 2011^[86]), particularly through focussing efforts on elements which contribute towards empowerment of actors as a means to “compensate for the deficiencies of both systems” (Patil, Vieider and Tetlock, 2014, p. 81^[179]).

Contingent selection for autonomy

Within the education field, the appropriate role of autonomy – particularly with respect to teachers in classrooms – has been a source of active debate. For some, entrusting teachers

with greater autonomy is closely associated with the success of education systems (Sahlberg, 2007^[55]; Sahlberg, 2010^[56]). More broadly, too, the role of autonomy is important in the accountability literature, given its association with psychological wellbeing (similar to the empowerment imbued by the hybridised formulation, page 57), and with this, greater task involvement and greater task persistence under adversity (Ryan and Deci, 2006^[244]). However, unchecked autonomy can be detrimental – for instance, by facilitating ‘excessive deviation’ (see Section 4.6.3).

For this reason, decision makers in education systems adjudicate on the extent to which autonomy is conferred in their systems – and to whom. The question of selection amongst actors can be a contentious one, particularly where autonomy is afforded to some, but not to others. Nonetheless, segmentation can be applied across a population (Besley, 2006^[245]) when it is contingent upon relevant factors such as demonstrated reputations, commitment, past performance, intrinsic motivations, and the like. In doing so, rather than assuming homogeneity of actors (Brennan, 1996^[246]; Besley, 2006^[245]), forums undertake selection (Mansbridge, 2009^[130]) that is ‘contingent’ on both actors and forums in the relevant context (Mansbridge, 2014^[183]). This is related to ‘responsive’ approaches that are adopted in the regulation literature, whereby escalation of ‘supports’ and ‘sanctions’ are employed, based upon the attitudes towards compliance that actors demonstrate (Braithwaite, 2010^[247]; Ayres and Braithwaite, 1992^[248]).

However, segmentation should be applied cautiously when it comes to autonomy. Some key considerations include: the alignment in pre-existing expectations between actors and forums; the relative burden that monitoring and sanctioning might cause forums; the potential that monitoring and sanctioning could inhibit performance; and the willingness to accept actors’ initiative and flexibility under uncertainty (Mansbridge, 2014^[183]).

From the accountability system-design level, the intention is to maintain as light a footprint as possible, affording as much autonomy to actors as appropriate, while maintaining the recourse for monitoring and sanction among those actors where it is deemed necessary (such as those that have yet to demonstrate a credible track record with the forum). In turn, relative efforts are shifted to the selection process – to select those considered as suitable for being afforded greater autonomy and otherwise – and with a lighter burden of efforts placed on monitoring and sanctioning.

Such an approach appreciates that applying instruments universally amongst actors can leave an impression amongst those actors that already meet expectations set of them – including those that have demonstrated their capacity and willingness to ‘get the job done’ in absence of external pressure – feeling distrusted. With this, arises an ‘institutionalised suspicion’ (Behn, 2001^[109]) towards accountability exchanges that can itself manifest in distrust – potentially inhibiting efforts and motivation towards exchange of account (recall from Section 4.4.1).

Nonetheless, a number of challenges may arise in implementing this approach. First is that the assumption of heterogeneity of actors can be at odds with other promises. For instance, notions of justice (namely, associated with equality before the law) and integrity (namely, associated with consistency of rules). Second is that of spurious selection. This applies both in the case of ‘unjust trust’ – assigning autonomy to an actor whom putatively ought to have been monitored – and ‘unjust distrust’ – applying monitoring and sanction to an actor that is unwarranted. In the case of the latter, it is argued that a result of this is that unjustly distrusting an actor can perpetuate a mutual cycle of distrust, and with this, detrimental results on accountability exchange.

Empowering actors to stimulate innovation and learning

Actors are commonly exposed to both process and outcome accountabilities simultaneously (Tetlock and Mellers, 2011_[86]), with a view to unlocking the respective hypothesised benefits of each (Spreitzer, 1995_[243]). Though some research (Chang et al., 2017_[220]) suggests that hybrid process-outcome instruments can potentially mitigate against downsides of pure outcome and process systems, Patil, Vieider and Tetlock (2014_[179]) warn that maintaining such a mix can be neither easy to deliver, nor immune to bringing out the worst in both.

Instead, focussing attention to elements associated with ‘empowering messages’ into instrument design when explicitly employing either process or outcome based instruments (but implicitly inducing a sense of the other accordingly) is argued to attenuate undesirable traits of the other (Patil, Vieider and Tetlock, 2014_[179]). So, in attempting to overcome the process-outcome trade-off, it is suggested that both control and innovation can be optimally maintained through the incorporation of such ‘messages’ in the design of accountability instruments. This includes helping actors to see that forums genuinely: find their joint efforts to be towards a meaningful endeavour; value their contributions towards collective efforts; take actors’ perspectives seriously; and proceed towards shared goals in a mutually respectful manner (Patil, Vieider and Tetlock, 2014_[179]).

To this end, Thomas and Velthouse³¹ (1990_[249]), discuss four additive factors that are found to be positively associated with innovation and initiative from actors’ conduct:

- Meaning: how an actor’s functions fit with their beliefs and intrinsic caring (Thomas and Velthouse, 1990_[249]).
- Competence: actor’s confidence in their capacity to perform functions (White, 1959_[250]).
- Self-determination: sense of personal control and autonomy over actions (Ryan and Deci, 2000_[251]; Deci, Connell and Ryan, 1989_[252]).
- Impact: perceived capacity to influence outcomes (Abramson, Seligman and Teasdale, 1978_[253]).

For actors, meaning and impact are anchored by the perception that their actions ‘make a difference’ and to desirable effect, such as through improving somebody’s welfare (say, improving student’s wellbeing) (Grant, 2007_[254]). Forums can motivate actors accordingly by drawing association to identifiable beneficiaries, particularly stimulating actors’ notions of their ‘task significance’ by frequent communication of the difference made to the lives of others (Grant, 2008_[255]).

Together, in order to be meaningful for actors, they should be able to see how their processes fit within the bigger picture (as opposed to be feeling that their day-to-day work ‘goes nowhere’) and be positively inclined about what is achieved (at a minimum, feeling that the conduct is taken seriously by forums). In turn, the establishment and maintenance of actor’s association between their processes (say, routine daily tasks) and outcomes (say, an organisational mission) is an important determinant of ensuring meaning and impact.

Forums perceived to undertake procedurally fair practice engender self-determination and autonomy in actors (Tyler, 1987_[178]). This is because the opportunity for ‘voice’ can

³¹ See also Patil, Vieider and Tetlock (2014_[179]); Spreitzer (1995_[243]); and Conger and Kanungo (1988_[242]).

engender a sense of autonomy in their decision making (for processes or outcomes) – including with respect to the abandonment of defective practices or conflicted goals. Given this ‘voice’, actors perceive an opportunity to defend their choices (if not the substantive ends *per se*) (Patil, Vieider and Tetlock, 2014_[179]), and in turn, this provides an avenue to exercise their autonomy.

Competence in this context is proxied by the perception of respect – because of the potential positive affect associated with external validation of actors’ competence, beyond self-validation. When authoritative forums confer respect (such as inviting actors to participate in discussion of the salience of processes in achieving desired outcomes), actors feel more valued and attach more meaning to the tasks they undertake, and also are more motivated to engage in accountability exchange. Moreover, by increasing perceived competence (while attenuating over-confidence (Tetlock and Kim, 1987_[256])) this can motivate critical assessment of how outcomes can best be achieved and also lowers reliance on intuition (Patil, Vieider and Tetlock, 2014_[179]) – stimulating the open-minded critical thinking required to facilitate learning through accountability (Lerner and Tetlock, 1999_[61]). Moreover, with competence comes more efficient processing of decision considerations (Huneke, Cole and Levin, 2004_[257]).

5.2. Meaningful accountability: effective exchanges and learning

Scholars have often argued that learning and accountability are “an unhappy couple” (Schillemans, Van Twist and Vanhommerig, 2013, p. 409_[71]). Theoretically, accountability can provide a platform to induce actors to learn, for example by prompting actors to revisit their conduct and, from this, to learn for future actions (Bovens, Schillemans and Hart, 2008_[121]). However, accountability often does not realise this potential in practice but instead introduces dynamics that curb learning (Behn, 2001_[109]; Neale and Anderson, 2000_[258]).

With respect to tensions between accountability and learning, authors point to the potential of accountability to encourage risk-aversion (Schillemans, Van Twist and Vanhommerig, 2013_[259]) (Behn, 2001_[109]; Neale and Anderson, 2000_[258]), goal displacement (Section 4.6.2), and diversion of resources (Halachmi, 2002_[260]). For instance, preoccupation to following rules and procedures enforced by accountability mechanisms is counterproductive to exploration, and thus, to learning (Behn, 2001_[109]). Others argue the very nature of accountability discourages innovation due to the emphasis placed on control and regularity (Ebrahim, 2005_[147]).

In any case, effectively functioning exchanges are a necessary precondition to stimulate learning through accountability (see further Section 5.2.2). Social psychology research identifies promising inroads to foster effective accountability exchanges in complex environments, particularly because of its independence of specific instruments and contexts (Tetlock, 1992_[12]; Schillemans, 2018_[261]; Harari and Rudolph, 2017_[262]; Hall, Frink and Buckley, 2017_[66]).

In the case of actors, effective accountability exchanges build on informative and effortful account rendering, as well as careful consideration and pursuit of conduct towards substantive expectations (Frink and Ferris, 1999_[263]). In relation to forums, accountability exchanges can be considered successful if they build on accurate, unbiased, carefully considered, and moderated assessment of actors’ conduct and, in turn, enable just, reliable, and appropriate judgements (Levy and Williams, 2004_[264]; Ashton, 1992_[265]).

5.2.1. *Promoting effective accountability exchanges*

Accountability shapes decision-making processes and how actors behave or “cope with” the situation of being held to account (Schillemans, 2018_[261]) (Tetlock, 1992_[12]; Lerner and Tetlock, 1999_[61]).

In general, forums holding actors accountable improves actors’ cognitive processing of information and reduces decision-making biases (Lerner and Tetlock, 1999_[61]; Hall, Frink and Buckley, 2017_[66]; Schillemans, 2018_[261]). Accountability’s effect on behaviour is more ambiguous. Actors and forums tend to be more cautious and risk-averse when held to account. This can be positive by deterring rash judgements and ‘jumping to conclusions’. However, risk-averse behaviour may also prompt decision-making deadlocks, decision avoidance and shifting of blame (Tetlock and Boettger, 1994_[266]; Schillemans, 2018_[261]; Tetlock, 1992_[12]).

Table 7 summarises potential stresses on accountability exchanges. The listed stresses are the main obstacles to effective decision-making processes and drive undesirable behaviours when coping with accountability. These pertain to actors and forums conforming to one another, inadequate information about accountability expectations and the human tendency to minimise cognitive efforts. Conformity pertains to participants in an accountability exchange placing more focus on pleasing their respective other than on reflecting critically about conduct. Inadequate information about accountability expectations prompts defensive behaviour, and minimising cognitive efforts leads to biased decisions.

Inroads to mitigate these stresses pertain to reducing predictability of expectations, raising awareness of substantive expectations before actors embark on a course of action (the timing of expectations), and mitigating the tendencies to minimise cognitive efforts. These inroads are discussed in detail below.

Table 7. Implications of accountability stresses on conduct, exchanges and possible learning

Accountability stress		Implications for conduct	Implications for exchange	Implications for learning
Conformity	Actor	Course of conduct is directed according to what is perceived as desirable for the forum (acceptability heuristic). Behaviours are based on what is predicted that forums are likely to look into or consider most salient.	Account rendering adopts an ingratiation strategy: 'Tell them what they want to hear' rather than 'what they need to hear'. Actors may attempt to adopt persuasion of forums rather than presenting the 'cold hard facts'.	Learning is inhibited because there is limited consideration of what is most appropriate substantively (such as evaluating possible trade-offs of courses of action). Actors instead focus on predicting forum's preferences.
	Forum	Forums are guided by avoiding disapproval of actors (e.g. preserving favourable reputations, averting unpleasant situations) based on acceptability ('go along to get along') heuristics. Forums that are held accountable for their ratings face similar acceptability heuristics as actors do ('keeping the boss happy').	Forums arrive at inflated (overly lenient) assessments of actors' conduct. Rating can also be inflated in accordance with superiors' expectations (such as middle managers wanting to keep executives happy – e.g. 'sweeping things under the rug'). Forums can be susceptible to groupthink (avoiding reputation as 'the hard marker') when forums are aware of how other forums assess actors.	Rather than being critical, lenient assessments do not constructively point to where actors can make improvements to their conduct, nor do they provide forums' superiors with accurate information.
Inadequate information about accountability expectations	Actor	Without pre-decisional information of expectations, actors interpret expectations based on potentially limited cues, and thus commit to a course of action relatively blind of forum expectations.	Actors find it harder to admit error (sunk cost effect) and take it personally (reacting and bolstering defensively) when their conduct is assessed. In turn, actors commit to self-justification (rather than self-criticism), following retrospective rationality. There is little to no opportunity for 'cognitive preparation' to allow for constructive account rendering or justification (such as recalling relevant mitigating factors affecting conduct).	Actors have limited opportunity to engage in pre-emptive self-criticism (weighing up the merits of how their conduct compares with forums' expectations relatively impartially). Adopting a defensive posture contributes to dismissal of forum assessment rather than genuine engagement in the content of assessments.
	Forum	Forums that are unexpectedly held accountable themselves (such as being identified or giving reasons for ratings) lean their assessments on what is most readily recalled, potentially overlooking relevant information.	Because forums are unprepared to provide assessment, they are more likely to make rash judgements and unfounded conclusions. They may also be less discerning of information rendered. Assessments may lack differentiation of contexts and integration into the 'bigger picture' ('integrative complexity').	Forums do not reach accurate assessments and judgements, which may misdirect feedback from exchanges. In turn, while learning may be induced it is possible to reinforce inaccurate conclusions reached by forums.
Low-effort tendencies and biases	Actor	Rather than analytically (weighing up alternatives/needs, etc.) considering how to behave in a given context and expectations, actors adopt intuition and heuristics.	Actors suffer from self-enhancement biases which can frustrate account rendering and justification posed. Actors can be overconfident in their own conduct and falsely assume that others think the way they do.	Without open-minded critical reflection on their own potential biases and shortcomings, actors do not constructively perceive of the opportunity to learn ('there is nothing for me to learn as I know it all').
	Forum	Forum assessments and judgements favour heuristics and introduce biased decisions through limited comprehension and evaluation of relevant arguments.	Forums succumb to biases leading to evaluative consistency, particularly: belief perseverance (dominance of first impressions or most recent observations) rather than being more discriminating; and attribution errors (under-appreciation of relevant mitigating factors).	Following heuristics leads to poorer assessments (less relevant and informed feedback to actors) and lowers actors' engagement and motivation in the underlying activity ('they never check this anyway, so why bother').

Reducing predictability of forums to overcome conformity

Predictability of a forum's preferences may prompt actors to focus on seeking approval of the forum over engaging in the effort to reflect critically on their substantive conduct. This is underpinned by an acceptability heuristic, which pertains in essence to the desire to fit in socially and liking to be liked (Tetlock, Skitka and Boettger, 1989^[73]; Tetlock, 1983^[211]; Wayne and Ferris, 1990^[267]; Tetlock, 1992^[12]).

Actors are inclined to undertake lesser efforts weighing up their options in terms of alternative courses of action and instead render accounts reflecting 'what they want to hear' (rather than 'what they need to hear') (Lerner and Tetlock, 1999^[61]). For forums, conforming behaviour implies that assessments are less regularly based on a balanced evaluation of conduct (say, a weighting of positive and negative conduct, or over a relevant period of time) and well-considered judgement of consequences.

Research finds that conformity is an underlying factor in actors seeking out favour of forums and forums arriving at lenient assessments (Bretz, Milkovich and Read, 1992^[268]; Shore and Tashchian, 2007^[269]; Wayne and Ferris, 1990^[267]; Harari and Rudolph, 2017^[262]). Examples include forums seeking to preserve interpersonal reputations (Spence and Keeping, 2011^[270]); discomfort of management in handing down critical performance appraisals (Levy and Williams, 2004^[264]); and participants seeking to project inflated assertions of their own competence (Harris, Smith and Champagne, 1995^[271]).

Forums not being predictable has been found to contribute to more balanced considerations of perspectives and trade-offs (Tetlock, 1983^[272]); produce more thorough justifications (Koonce, Anderson and Marchant, 1995^[202]); and qualified assessments (Lord, 1992^[273]). It helps mitigating undesirable behaviour such as actors 'gaming' their conduct and forums engaging in conflicts of interest.

Actors are clear about what is expected before they are held to account

When participants are made aware of what is expected of them matters for their decision-making processes and how they cope with being held accountable. This includes expectations related to information required for account rendering, the criteria used for account assessment, and the evaluative judgements (the 'stakes') to be applied.

Being confronted with accountability expectations *after* having engaged in conduct can prompt diverting cognitive efforts to making excuses and justifications in a backward-looking, defensive search to rationalise past conduct ('defensive bolstering') (Lerner and Tetlock, 1999^[61]). When held to account unexpectedly for conduct, actors may feel compelled to recommit to previously held positions when rendering account. This includes courses of action and positions to which actors may have had little commitment initially (Schillemans, 2016^[14]; Arkes and Blumer, 1985^[274]).

Empirical research finds that conveying accountability expectations to actors *before* they engage in a course of action ('pre-decisional information') contributes to better decisions, forward-looking rationality, and self-critical reflection (Schillemans, 2018^[261]; Schillemans, 2016^[14]; Tetlock, 1992^[12]). Information rendered is more carefully recounted and more likely to be free of superfluous information. Actors are less likely to obfuscate or misrepresent conduct (Arkes, 1991^[275]).

Overcoming low effort tendencies and biases

Motivated and effortful engagement in accountability exchanges are fundamental to effective exchanges and can improve decision quality (Scholten et al., 2007_[276]). This relates to efforts and time invested in searching information and weighing up options when making a decision, and being personally invested in one's work (Frink and Ferris, 1999_[263]; Huneke, Cole and Levin, 2004_[257]; De Dreu, Koole and Steinel, 2000_[230]; Lee et al., 1999_[185]).

People generally seek to avoid cognitive efforts that are consciously or unconsciously perceived as unnecessary (Tetlock, 1992_[12]). This includes reducing efforts in accountability exchanges such as engaging in the exchange superficially, using heuristics rather than reflecting critically on conduct and avoiding accountability exchanges altogether (Chaiken, 1980_[277]). Low engagement and motivation in the underlying work is linked to greater tendency to avoid cognitive efforts in the accountability exchange.

Low-effort strategies linked to low motivation tend to produce biases such as: preconceptions not being revised despite new information (belief perseverance), overemphasising personal efforts and characteristics while discounting situation and context (fundamental attribution error), and over-confidence in one's own capabilities and actions (Fischhoff, 1982_[278]; Sedikides et al., 2002_[189]; Heidemeier and Moser, 2009_[279]; Tetlock, 1983_[211]).

Such biases and strategies to save cognitive efforts tend to produce greater consistency in forums' assessment and actors' self-assessment ('lazy decision making') (Lerner and Tetlock, 1999_[61]; Tetlock, Skitka and Boettger, 1989_[73]).

5.2.2. Promoting learning through accountability exchanges

Central to learning under accountability are 'feedback-based inducements' (Bovens, Schillemans and Hart, 2008_[121]; Bovens, 2007_[11]) prompting actors' reflection on conduct (that is, recognising and systematically correcting deficiencies vis-à-vis the expected standard) or acting on feedback from forums (Moynihan, 2005_[280]). However, not all learning is equal – with adaptive learning the most desired given the conditions of complexity. Moreover, it bears noting that there are some 'pathologies' of learning, particularly those pertaining out of low-effort diffusion which can produce dysfunctional results, such as groupthink, bad habits, and the like (Dunlop, 2017_[281]).

A distinction has already been made in terms of whether 'things are done right' (processes) or if 'the right things are done' (outcomes). The corollary in literature on learning is that of single-loop (doing things right) and double-loop (doing the right thing) learning. The learning literature also points to 'triple-loop learning', which is about assessing what is appropriate in given context. Under complexity's dynamic and adaptive conditions, learning of this sort is evidently desirable. Accordingly, so-called adaptive performance, which considers the ability to adjust to uncertain, complex and dynamic tasks has tended to be supported within settings of outcome accountability, including those beset with uncertainty (Chang et al., 2017_[220]; Huang et al., 2014_[282]).

In the remainder of this subsection, the state of learning within accountability literature within public administration and organisational studies is briefly considered (largely informed by organisational learning). Scholars in the accountability literature have tended to focus on barriers and opportunities for learning (Greiling and Halachmi, 2013_[170]; Behn, 2001_[109]), including the conditions which are likely to facilitate this (Table 8).

The potential link between accountability mechanisms and learning is thus argued to hinge on two primary factors if it is to bring about self-critical examination of past performance, decisions, and practices (Schillemans, Van Twist and Vanhommerig, 2013_[71]). First is to provide relevant feedback information that would be closed off otherwise. Namely, the need for sufficient (and sufficiently accurate) information so as to guide prospective behaviour (Ebrahim, 2005_[147]), as well as providing an environment that does not trigger counterproductive responses (avoiding circumstances in which actors fear retribution or sanction, for instance) (Schillemans, Van Twist and Vanhommerig, 2013_[259]).

Table 8. Conditions for learning under accountability

	Findings with respect to learning
Substantive focus	<ul style="list-style-type: none"> • Learning can be enabled when accountability is focussed on long-term goals, rather than dwelling on narrow, short-term goals or targets – see also Section 4.6.3. • Exchanges should find new ways to improve future performance rather than being concerned with dwelling on past mistakes, if learning is to be mobilised – see also Section 4.6.3. • To be sustainable, learning is most present when there is an ‘internal accountability to mission’ (recall Section 1). That is, actors (recall from page 57) and forums (recall from pages 44) are committed to their underlying conduct. • Forums that apply an excessive focus on control (Section 4.6.2) and regularity (recall from pages 34-35 and 43) may create an environment too tight to facilitate learning. This is because actors are more likely to frustrate the supply of information (recall page 42), rather than engage in accountability with learning at top of mind.
Interpersonal trust	<ul style="list-style-type: none"> • Relatively trusting approaches, which treat accountability exchanges as professional relationships, are found to encourage learning within particular settings. • Anticipation that forums are legitimate and of ‘high status’ (such as competence, expertise, and the like (recall pages 41-42)) can motivate pre-emptive self-criticism. • Suspicious signals (recall pages 54, 41), such as controlling (rather than empowering (recall page 55) messages, can contribute towards actors limiting the information they render, thereby impairing possibilities for learning. • Relatively informal relationships (recall Section 3.3.4) are considered more accommodating for learning, because of lesser tensions and defensive responses – conditional on a willingness to improve in the first instance. • Open and ongoing dialogue, and the encouragement of sharing knowledge and insights can contribute towards reduced defensive and evasive tendencies.
Availability of sanctions	<ul style="list-style-type: none"> • The threat of punitive sanctions can trigger counterproductive responses that inhibit actors’ admissions of error or omissions (recall page 42). In turn, inhibiting honest, unbiased provision of all the relevant facts. Instead, when errors are embraced as opportunities, and absent of low-trust, high-blame situations, this can stave off potential blame games (recall page 33) – which orient efforts towards blame avoidance and shifting, rather than learning. • At least a moderate recourse to sanctions is argued as appropriate for learning, particularly in instances where actors are unwilling to adapt on the basis of reasoned arguments only. Moderate sanctions are considered appropriate in order to push actors to reform in response to critical performance feedback, whereas the threat of severe sanctions are more likely to trigger distortions, such as strategic behaviour shifts.

Source: Adapted from Schillemans and Smulders (2016_[72]); Greiling and Halachmi (2013); Greiling and Spraul (2010_[151]); Ebrahim (2005_[147]); Patil, Vieider, & Tetlock (2014_[179]); Hood (2010_[283]); De Bruijn (2002_[284]).

To this end, the quality of information that is rendered necessarily influences the prospects for learning, since it not only drives the extent of self-critical reflection, but also determines the quality of feedback information (Schillemans, Van Twist and Vanhommerig, 2013_[259]). In the ‘learning perspective’ offered by Bovens, Schillemans and Hart (2008_[121]), learning depends upon the provision of accurate, timely, and unambiguous accounts. When information rendered is deemed to be novel (Schillemans, 2011_[285]), is collected flexibly (Ebrahim, 2005_[147]), and is derived from a diverse range of sources (Strøm, 2000_[112]), it is suggested that this can facilitate learning (Schillemans, Van Twist and Vanhommerig, 2013_[259]).

Given the importance of the structure of relationships between actor and forum in stimulating effective exchanges, the nature of feedback generated is also necessarily relevant for learning. To begin with, the perceptions of an authoritative (Mero, Guidice and Brownlee, 2007^[187]; Curtis, Harvey and Ravden, 2005^[286]) and legitimate (Cvetkovich, 1978^[287]; Gordon and Stuecher, 1992^[288]) audience enhances efforts, motivation and engagement in accountability exchanges. In addition, forums that are more authoritative are less lenient in their assessment (Pennington and Schlenker, 1999^[201]).

Moreover, research identifies that forums that are required to justify their assessments (Mero and Motowidlo, 1995^[289]), particularly when in person (Mero, Guidice and Brownlee, 2007^[187]), offer more accurate and more detailed assessment (Gordon and Stuecher, 1992^[288]). Forums that provide judgements in person have also been found to result in greater likelihood that feedback is incorporated into performance improvement efforts of actors (Walker and Smither, 1999^[290]). However, in person feedback can potentially contribute towards forums reporting inflationary ratings for poor performing actors (Curtis, Harvey and Ravden, 2005^[286]) – since some face-to-face exchanges may reinforce a general desire to eschew unpleasant social interactions.

6. Conclusion and outlook

This paper set out to review interdisciplinary scholarly literature incorporating contributions beyond education – including public administration, social psychology and organisational studies – and apply lessons to the governance of complex education systems. This is particularly relevant given that accountability in education literature and practice has to date been highly contentious and almost universally defined according to the respective instruments employed by practitioners.

First, there are important implications of the complexity paradigm for governance of accountability, which reorient perspectives and priorities. In particular, efforts to improve accountability systems need to focus on the *accountability relationships* themselves beyond individual accountability instruments. Accountability relationships are grounded in each participant’s subjectivity, history and local context. The review focussed on the micro-level of accountability, revolving around the relationship between forums and actors in the process of rendering, assessing and judging accounts of conduct. The relational approach allowed identifying enablers and barriers of successful accountability relationships.

Second, a core finding from reviewing the interdisciplinary literature is that within complex systems, approaches to accountability in governance should seek to *accommodate* rather than reduce complexity. Managing accountability is less about ‘solving’ challenges associated with accountability exchanges than contending with accountability stresses in a continuous, dynamic process. Moving towards continuous improvement of accountability exchanges in education systems necessitates persistent efforts to attenuate possible stresses.

Accountability stresses have been identified and discussed throughout this paper. These stresses generally refer to any elements diminishing effort, motivation, and engagement in rendering, assessing, or making judgements in an accountability exchange. At the heart of many such stresses are several common ailments, often with root causes located in social psychology research’s understanding of conformity to known audiences, a lack of pre-emptive information, and suffering from other low-effort decision making tendencies and biases. In addition, potential stressors have been discussed with respect to appropriateness of combinations of dimensions (Section 3.3) and promises (Section 4) within a given accountability exchange.

The review concluded with a sketch of approaches to attenuate accountability stresses by fostering effortful, motivated and engaged behaviour in exchanges of both actors and forums. For actors, an effective exchange stimulates detailed, accurate and complete accounts of conduct. For forums, an effective exchange assesses these accounts accurately, fairly, carefully and cautiously. With this, the verdict reached – and consequences doled out – are just and applied consistently. With the approaches contributing directly to individual behaviour, they offer inroads to improving accountability exchanges independent of specific contexts and instruments. To this end, a common element of such exchanges is that participants find the exchange meaningful. Namely, how meaningful an exchange is for those involved can be appraised in light of how well it avoids or limits the accountability stresses revealed in research. For the governance of accountability in complex education systems, this suggests that practitioners could benefit from closer examination of the experience of participants in accountability exchanges.

Nonetheless, the study of accountability remains as lively as ever, with promising prospects for further development into a field of inquiry in its own right (Flinders, 2014_[291]). The

discipline's single greatest deficiency to date has been its lack of conceptual clarity and indeed the lack of a universal conceptual framework (Willems and Van Dooren, 2012^[157]). On both counts, this paper has made first steps towards mitigating these deficiencies, with scholars and practitioners encouraged to build upon this.

For the discipline to fully grasp its potential to arrive at meaningful insights for practice, there are two primary obstacles research may address. The first is that the predominantly segmented disciplinary approaches could benefit from drawing on broader fields of inquiry. Indeed, recent progress in the literature has resulted from interdisciplinary applications – of particular note, the import of the insights of social psychology into public administration literature (Schillemans, 2016^[14]; Schillemans, 2018^[261]).

In common with this recent trend, this paper has been guided analytically by complexity sciences and social psychology, in particular, along with a comprehensive review of the general accountability literature from public administration and organisational sciences. Scholars are encouraged to continue efforts, including through importation across, as yet, uncharted relevant disciplines. In particular, further integration from the literature on learning is critical in this, including in terms of identification of what content is actually retained from exchanges (Schillemans and Smulders, 2016^[72]).

The second fundamental obstacle for accountability studies is to establish a firmer grounding in empirical study. To date, both qualitative and quantitative empirical branches remain relatively nascent (Brandsma, 2014^[292]; Yang, 2014^[293]). It is pertinent that scholars take opportunities to validate and challenge the experimental findings of social psychology beyond the laboratory and into observational, real world practice (Schillemans, 2016^[14]).

Finally, this paper has been underlined by complexity science and the associated systems approach. As one of the most promising opportunities within interdisciplinary studies, further application of systems approaches represent an, as yet, relatively untapped opportunity for accountability scholars [barring, for instance, Vibert (2014^[62])]. Such approaches offer promises both to establish additional conceptual grounding, but most importantly, opportunities for empirical application.

Conceptually, this paper has drawn on several of the principal implications of complexity and systems theory, but especially with regard to the focus of a relational paradigm at the micro-level. Empirical analysis (particularly in the form of case studies) could offer valuable insight of these interactions and how they manifest in behaviour and decision making in practice.

In common with other applications of systems approaches in the social sciences, the mechanisms through which to appropriately link the micro and macro levels of accountability systems remains relatively under-explored. To connect accountability exchanges at the micro-level to the macro level, scholars are encouraged to investigate the role of environmental macro conditions, such as institutional settings, domains of governance, and other contextual features in influencing micro level accountability exchanges.

Moreover, the notion of meaning – particularly in terms of ‘meaning making’ – at the macro level could offer a fruitful conceptual grounding for underpinning of these various elements, given the attention paid to coherence and purpose as meaning making mechanisms, in particular. Together, uncovering mechanisms to embed meaning in accountability exchanges from the macro perspective could offer a fruitful means to further support governance in complex education systems and complement the micro foundations explored in this paper.

References

- Abramson, L., M. Seligman and J. Teasdale (1978), “Learned Helplessness in Humans: Critique and Reformulation”, *Journal of Abnormal Psychology*, Vol. 87/1, pp. 49-74. [253]
- Acar, M., C. Guo and K. Yang (2008), “Accountability when hierarchical authority is absent: Views from public-private partnership practitioners”, *American Review of Public Administration*, Vol. 38/1, pp. 3-23, <http://dx.doi.org/10.1177/0275074007299481>. [120]
- Antonioni, D. (1994), “The effects of feedback accountability on upward appraisal ratings”, *Personnel Psychology*, Vol. 47, pp. 349-356. [114]
- Arkes, H. (1991), “Costs and benefits of judgment errors: Implications for debiasing”, *Psychological Bulletin*, Vol. 110/3, pp. 486-498, <http://dx.doi.org/10.1037/0033-2909.110.3.486>. [275]
- Arkes, H. and C. Blumer (1985), “The Psychology of Sunk Cost”, *Behavior and Human Decision Processes*, Vol. 35, pp. 124-140. [274]
- Ashton, R. (1992), “Effects of justification and a mechanical aid on judgment performance”, *Behavior and Human Decision Processes*, Vol. 52/2, pp. 292-306. [265]
- Ayres, I. and J. Braithwaite (1992), *Responsive Regulation*, Oxford University Press, New York. [248]
- Barnard, C. (1938), *The Functions of the Executive*, Harvard University Press, Cambridge, MA. [205]
- Beach, L. and T. Mitchell (1978), “A Contingency Model for the Selection of Decision Strategies”, *Academy of Management Review*, Vol. 3/3, pp. 439-449, <http://dx.doi.org/10.5465/amr.1978.4305717>. [229]
- Behn, R. (2001), *Rethinking democratic accountability*, Brookings Institute Press, Washington D.C. [109]
- Bertrand, M. and S. Mullainathan (2001), “Do People Mean What They Say? Implications for Subjective Survey Data”, *American Economic Review*, Vol. 91/2, pp. 67-72. [232]
- Besley, T. (2006), *Principled Agents: The Political Economy of Good Government*, Oxford University Press, Oxford. [245]
- Bloom, A. (1968), *The Republic of Plato*, Basic Books, <http://dx.doi.org/10.4324/9781912281992>. [169]
- Bouckaert, G. and B. Peters (2002), “Performance Measurement and Management: The Achilles’ Heel in Administrative Modernization Dilemmas and Issues of Performance-Based Management”, *Public Performance & Management Review*, Vol. 25/4, pp. 359-362. [217]

- Boulton, J. (2015), *Embracing complexity : strategic perspectives for an age of turbulence*, Oxford University Press, Oxford. [23]
- Bovaird, T. (2007), “Beyond Engagement and Participation”, *Public Administration Review*, Vol. 67/Sept/Oct, pp. 846–60. [200]
- Bovens, M. (2010), “Two Concepts of Accountability: Accountability as a Virtue and as a Mechanism”, *West European Politics*, Vol. 33/5, pp. 946-967. [3]
- Bovens, M. (2009), “Public Accountability”, in *The Oxford Handbook of Public Management*, <http://dx.doi.org/10.1093/oxfordhb/9780199226443.003.0009>. [118]
- Bovens, M. (2007), “Analysing and Assessing Accountability: A Conceptual Framework”, *European Law Journal*, <http://dx.doi.org/10.1111/j.1468-0386.2007.00378.x>. [1]
- Bovens, M. et al. (2014), “Public Accountability”, in *The Oxford Handbook of Public Accountability*, <http://dx.doi.org/10.1093/oxfordhb/9780199641253.013.0012>. [141]
- Bovens, M. and T. Schillemans (2014), “Meaningful Accountability”, in Bovens, M., R. Goodin and T. Schillemans (eds.), *The Oxford Handbook of Public Accountability*, <http://dx.doi.org/10.1093/oxfordhb/9780199641253.013.0038>. [94]
- Bovens, M., T. Schillemans and P. Hart (2008), “Does public accountability work? An assessment tool”, *Public Administration*, Vol. 86/1, pp. 225-242, <http://dx.doi.org/10.1111/j.1467-9299.2008.00716.x>. [121]
- Braithwaite, J. (2010), *The essence of responsive regulation*. [247]
- Brandsma, G. (2014), “Quantitative Analysis”, in Bovens, M., R. Goodin and T. Schillemans (eds.), *The Oxford Handbook of Public Accountability*, <http://dx.doi.org/10.1093/oxfordhb/9780199641253.013.0041>. [292]
- Brennan, G. (1996), “Selection and the Currency of Reward”, in Goodin, R. (ed.), *The Theory of Institutional Design*, Cambridge University Press, Cambridge. [246]
- Bretz, R., G. Milkovich and W. Read (1992), “The current state of performance appraisal research and practice: Concerns, directions, and implications”, *Journal of Management*, Vol. 18/2, pp. 321-352. [268]
- Brtek, M. and S. Motowidlo (2002), “Effects of procedure and outcome accountability on interview validity.”, *The Journal of applied psychology*, Vol. 87/1, pp. 185-191, <http://dx.doi.org/10.1037/0021-9010.87.1.185>. [231]
- Budescu, D. and N. Du (2007), “Coherence and Consistency of Investors’ Probability Judgments”, *Management Science*, Vol. 53/11, pp. 1731-1744, <http://dx.doi.org/10.1287/mnsc.1070.0727>. [322]
- Burnheim, J. (1985), *Is democracy possible? The alternative to electoral politics*, University of California Press, Berkeley. [104]

- Burns, T. and F. Köster (2016), *Governing Education in a Complex World*, OECD Educational Research and Innovation, OECD Publishing, Paris. [16]
- Burns, T. and F. Köster (2016), “Modern governance challenges in education”, in *Governing Education in a Complex World*, OECD Publishing, Paris, <https://dx.doi.org/10.1787/9789264255364-3-en>. [49]
- Burns, T., F. Köster and M. Fuster (2016), *Education Governance in Action: Lessons from Case Studies*, Educational Research and Innovation, OECD Publishing, Paris, <https://dx.doi.org/10.1787/9789264262829-en>. [18]
- Busuioac, M. and M. Lodge (2017), “Reputation and Accountability Relationships: Managing Accountability Expectations through Reputation”, *Public Administration Review*, Vol. 77/1, pp. 91–100, <http://dx.doi.org/10.1111/puar.12612>. [136]
- Byrne, D. (1998), *Complexity Theory and the Social Sciences: An Introduction*, Routledge, New York. [26]
- Camm, F. and B. Stecher (2010), “How a Performance-Based Accounting System Changes Service Delivery: An Analytic Overview”, in RAND (ed.), *Analyzing the Operation of Performance-Based Accountability Systems for Public Services*, RAND Corporation, Santa Monica, <http://dx.doi.org/10.7249/tr853.9>. [212]
- Canales, R. (2013), “Weaving Straw into Gold: Managing Organizational Tensions Between Standardization and Flexibility in Microfinance”, *Organization Science*, Vol. 25/1, pp. 1-28, <http://dx.doi.org/10.1287/orsc.2013.0831>. [225]
- Carroll, T. and R. Burton (2000), “Organizations and Complexity: Searching for the Edge of Chaos”, *Computational and Mathematical Organization Theory*, Vol. 6/4, pp. 319-337. [89]
- Chaiken, S. (1980), “Heuristic Versus Systematic Information Processing and the Use of Source Versus Message Cues in Persuasion”, *Journal of Personality and Social Psychology*, Vol. 39/5, pp. 752-766. [277]
- Chakravarthy, B. and Y. Doz (1992), “Strategy Process Research: Focusing on Corporate Self-Renewal”, *Strategic Management Journal*, Vol. 13, pp. 5-14, <https://about.jstor.org/terms>. [313]
- Chang, W. et al. (2017), “Accountability and adaptive performance under uncertainty: A long-term view”, *Judgment and Decision Making*, Vol. 12/6, pp. 610-626, <https://www.researchgate.net/publication/321339998>. [220]
- Chen, Z., W. Lam and J. Zhong (2007), “Leader-member exchange and member performance: A new look at individual-level negative feedback-seeking behavior and team-level empowerment climate”, *Journal of Applied Psychology*, Vol. 92/1, pp. 202-212, <http://dx.doi.org/10.1037/0021-9010.92.1.202>. [316]
- Christensen, T. and P. Lægreid (2015), “Performance and Accountability—A Theoretical Discussion and an Empirical Assessment”, *Public Organization Review*, Vol. 15, pp. 207-225, <http://dx.doi.org/10.1007/s11115-013-0267-2>. [213]

- Cialdini, R. and N. Goldstein (2004), “Social Influence: Compliance and Conformity”, *Annual Review of Psychology*, Vol. 55, pp. 591–621, <http://dx.doi.org/10.1146/annurev.psych.55.090902.142015>. [156]
- Cochran-Smith, M., P. Piazza and C. Power (2013), “The Politics of Accountability: Assessing Teacher Education in the United States”, *Educational Forum*, Vol. 77/1, pp. 6-27, <http://dx.doi.org/10.1080/00131725.2013.739015>. [59]
- Conger, J. and R. Kanungo (1988), “The Empowerment Process: Integrating Theory and Practice.”, *Academy of Management Review*, Vol. 13/3, pp. 471-482, <http://dx.doi.org/10.5465/AMR.1988.4306983>. [242]
- Considine, M. (2002), “The end of the line? Accountable governance in the age of networks, partnerships, and joined-up services”, *Governance*, Vol. 15/1, pp. 21–40, <http://dx.doi.org/10.1111/1468-0491.00178>. [144]
- Curtis, A., R. Harvey and D. Ravden (2005), “Sources of political distortions in performance appraisals: Appraisal purpose and rater accountability”, *Group and Organization Management*, Vol. 30/1, pp. 42-60, <http://dx.doi.org/10.1177/1059601104267666>. [286]
- Cvetkovich, G. (1978), “Cognitive Accommodation, Language, and Social Responsibility*”, *Social Psychology*, Vol. 41/2, pp. 149-155, <https://about.jstor.org/terms>. [287]
- Damgaard, B. and J. Lewis (2014), “Accountability and Citizen Participation”, in Bovens, M., R. Goodin and T. Schillemans (eds.), *The Oxford Handbook of Public Accountability*, Oxford University Press, Oxford, <http://dx.doi.org/10.1093/oxfordhb/9780199641253.013.0020>. [199]
- de Bruijn, H. (2002), “Performance measurement in the public sector: Strategies to cope with the risks of performance measurement”, *International Journal of Public Sector Management*, Vol. 15/7, pp. 578-594, <http://dx.doi.org/10.1108/09513550210448607>. [284]
- De Dreu, G., S. Koole and W. Steinel (2000), “Unfixing the fixed pie: A motivated information-processing approach to integrative negotiation”, *Journal of Personality and Social Psychology*, Vol. 79/6, pp. 975-987, <http://dx.doi.org/10.1037/0022-3514.79.6.975>. [230]
- Deandrea, D. et al. (2012), “When Do People Misrepresent Themselves to Others? The Effects of Social Desirability, Ground Truth, and Accountability on Deceptive Self-Presentations”, *Journal of Communication*, Vol. 62, pp. 400-417, <http://dx.doi.org/10.1111/j.1460-2466.2012.01646.x>. [188]
- Deci, E., J. Connell and R. Ryan (1989), “Self-Determination in a Work Organization”, *Journal of Applied Psychology*, Vol. 74/4, pp. 580-590. [252]
- Dezoort, F. et al. (2006), “The Effects of Accountability Pressure Strength and Decision Aid Availability on Auditors’ Materiality Judgments”, *Accounting, Organizations and Society*, Vol. 31/4-5, pp. 373-390. [304]
- Dobel, J. (1990), “Integrity in the Public Service”, *Public Administration Review*, Vol. 50/3, pp. 354-366, <http://dx.doi.org/10.2307/976617>. [166]

- Dooley, K. and A. Van de Ven (1999), “Explaining Complex Organizational Dynamics”, [79]
Organization Science, Vol. 10/3, pp. 358-372, <http://dx.doi.org/10.1287/orsc.10.3.358>.
- Douglas, T. and W. Judge (2001), “Total Quality Management Implementation and Competitive [240]
 Advantage: The Role of Structural Control and Exploration”, *The Academy of Management Journal*, Vol. 44/1, pp. 158-169.
- Dubnick, M. (2003), “Accountability and ethics: reconsidering the relationships”, [161]
International Journal of Organization Theory & Behavior, Vol. 6/3, pp. 405-441,
<http://dx.doi.org/10.1108/ijotb-06-03-2003-b002>.
- Dubnick, M. (2002), *The Rhetoric and Promises of Accountability*, Remarks prepared as [163]
 Fulbright Senior Specialist presentation, Leiden University.
- Dubnick, M. and H. Frederickson (2011), “Introduction: The Promises of Accountability [159]
 Research”, in Dubnick, M. and H. Frederickson (eds.), *Accountable Government : Problems and Promises*, Routledge.
- Dubnick, M. and H. Frederickson (2010), “Accountable agents: Federal performance [165]
 measurement and third-party government”, *Journal of Public Administration Research and Theory*, Vol. 20/1, pp. 143–159, <http://dx.doi.org/10.1093/jopart/mup039>.
- Dubnick, M. and H. Frederickson (2010), *Accountable Governance: Problems and Promises*, [227]
 M.E. Sharpe.
- Dubnick, M. and J. Justice (2004), *Accounting for Accountability*. [160]
- Dubnik, M. (2005), “Accountability and the promise of performance: In Search of the [164]
 Mechanisms”, *Public Performance & Management Review*, Vol. 28/3, pp. 376-417,
<http://dx.doi.org/10.1080/15309576.2005.11051839>.
- Duit, A. and V. Galaz (2008), “Governance and Complexity-Emerging Issues for Governance [82]
 Theory”, *Governance*, Vol. 21/3, pp. 311–335.
- Dunlop, C. (2017), “Pathologies of policy learning: what are they and how do they contribute to [281]
 policy failure?”, *Policy and Politics*, Vol. 45/1, pp. 19-37.
- Ebrahim, A. (2005), “Accountability myopia: Losing sight of organizational learning”, [147]
Nonprofit and Voluntary Sector Quarterly, Vol. 34/1, pp. 56-87,
<http://dx.doi.org/10.1177/0899764004269430>.
- Ebrahim, A. (2003), “Making sense of accountability: Conceptual perspectives for northern and [107]
 southern nonprofits”, *Nonprofit Management and Leadership*, Vol. 14/2, pp. 191-212,
<http://dx.doi.org/10.1002/nml.29>.
- Ehren, M. and A. Visscher (2006), “Towards a theory on the impact of school inspections”, [297]
British Journal of Educational Studies, Vol. 54/1, pp. 51-72, <http://dx.doi.org/10.1111/j.1467-8527.2006.00333.x>.

- Eisenhardt, K. (1989), "Agency Theory: An Assessment and Review.", *Academy of Management Review*, Vol. 14/1, pp. 57-74, <http://dx.doi.org/10.5465/AMR.1989.4279003>. [239]
- Enzle, M. and S. Anderson (1993), "Surveillant Intentions and Intrinsic Motivation", *Journal of Personality and Social Psychology*, Vol. 64/2, pp. 257-266, <http://dx.doi.org/10.1037/0022-3514.64.2.257>. [236]
- Eppel, E. and M. Rhodes (2018), "Complexity theory and public management: a 'becoming' field", *Public Management Review*, Vol. 20/7, pp. 949-959, <http://dx.doi.org/10.1080/14719037.2017.1364414>. [314]
- Epstein, J. and R. Axtell (1996), *Growing artificial societies: Social science from the bottom up*, MIT Press, Cambridge. [60]
- Eraut, M. (1994), *Developing Professional Knowledge and Competence Acknowledgments*, Falmer Press, London. [301]
- Espeland, W. and M. Sauder (2016), *Engines of anxiety: Academic rankings, reputation, and accountability*, Russell Sage Foundation, New York. [105]
- Ethiraj, S. and D. Levinthal (2009), "Hoping for A to Z While Rewarding Only A. Complex Organizations and Multiple Goals", *Organization Science*, Vol. 20, pp. 4-21, <http://dx.doi.org/10.1287/orsc.1080.0358>. [155]
- Fandt, P. (1991), "The Relationship of Accountability and Interdependent Behavior to Enhancing Team Consequences", *Group & Organization Studies*, Vol. 16/3, pp. 300-312. [317]
- Fischhoff, B. (1982), "Debiasing", in Kahneman, D., P. Slovic and A. Tversky (eds.), *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge University Press, New York, <http://dx.doi.org/10.1046/j.1365-2753.2001.00284.x>. [278]
- Flinders, M. (2014), "The Future and Relevance of Accountability Studies", in Bovens, M., R. Goodin and T. Schillemans (eds.), *The Oxford Handbook of Public Accountability*, Oxford University Press, New York, <http://dx.doi.org/10.1093/oxfordhb/9780199641253.013.0004>. [291]
- Frederickson, H. (2003), *Best Practice, Benchmarking, Ranking, and the Management of Organizational Creativity*, Kettering Foundation, Washington DC. [306]
- Friedrich, C. (1940), "Public Policy and the Nature of Administrative Responsibility", in Friedrich, C. and E. Mason (eds.), *Public Policy: A Yearbook of the Graduate School of Public Administration*, Harvard University Press, Cambridge, MA. [208]
- Frink, D. and G. Ferris (1999), "The moderating effects of accountability on the conscientiousness-performance relationship", *Journal of Business and Psychology*, Vol. 13/4, pp. 515-524, <http://dx.doi.org/10.1023/A:1022918904256>. [263]
- Frink, D. and R. Klimoski (2004), "Advancing accountability theory and practice: Introduction to the human resource management review special edition", *Human Resource Management Review*, Vol. 14, pp. 1-17, <http://dx.doi.org/10.1016/j.hrmmr.2004.02.001>. [228]

- Frink, D. and R. Klimoski (1998), “Toward a theory of accountability in organizations and human resources management”, in Ferris, G. (ed.), *Research in Personnel and Human Resources Management*, Elsevier Science/JAI Press. [11]
- Fry, R. (1995), “Accountability in organizational life: Problem or opportunity for nonprofits?”, *Nonprofit Management and Leadership*, Vol. 16/2, pp. 181-195, <http://dx.doi.org/10.1002/nml.4130060207>. [139]
- Gailmard, S. (2014), “Accountability and Principal–Agent Theory”, in Bovens, M., R. Goodin and T. Schillemans (eds.), *The Oxford Handbook of Public Accountability*, Oxford University Press, Oxford, <http://dx.doi.org/10.1093/oxfordhb/9780199641253.013.0016>. [180]
- Gailmard, S. and J. Patty (2012), “Formal Models of Bureaucracy”, *Annual Review of Political Science*, Vol. 15, pp. :353–77, <http://dx.doi.org/10.1146/annurev-polisci-031710-103314>. [135]
- Gains, F. and G. Stoker (2009), “Delivering ‘public value’: Implications for accountability and legitimacy”, *Parliamentary Affairs*, Vol. 62/3, pp. 438-455, <http://dx.doi.org/10.1093/pa/gsp007>. [54]
- Gelfand, M., B. Lim and J. Raver (2004), “Culture and accountability in organizations: Variations in forms of social control across cultures”, *Human Resource Management Review*, Vol. 14, pp. 135-160, <http://dx.doi.org/10.1016/j.hrmr.2004.02.007>. [76]
- Gordon, R. and U. Stuecher (1992), “The effect of anonymity and increased accountability on the linguistic complexity of teaching evaluations”, *Journal of Psychology: Interdisciplinary and Applied*, Vol. 126/6, pp. 639-649, <http://dx.doi.org/10.1080/00223980.1992.10543394>. [288]
- Granovetter, M. (1978), “Threshold Models of Collective Behavior”, *American Journal of Sociology*, Vol. 83/6, pp. 1420-1443, <http://dx.doi.org/10.1086/226707>. [83]
- Grant, A. (2008), “The Significance of Task Significance: Job Performance Effects, Relational Mechanisms, and Boundary Conditions”, *Journal of Applied Psychology*, Vol. 93/1, pp. 108 – 124, <http://dx.doi.org/10.1037/0021-9010.93.1.108>. [255]
- Grant, A. (2007), “Relational job design and the motivation to make a prosocial difference”, *Academy of Management Review*, Vol. 32/2, pp. 393–417, <http://dx.doi.org/10.1200/JCO.2005.02.2368>. [254]
- Green, J. (2011), *Education, Professionalism, and the Quest for Accountability*, Routledge, New York. [300]
- Green, M., P. Visser and P. Tetlock (2000), “Coping with accountability cross-pressures: Low-effort evasive tactics and high-effort quests for complex compromises”, *Personality and Social Psychology Bulletin*, Vol. 26/11, pp. 1380-1391, <http://dx.doi.org/10.1177/0146167200263006>. [203]
- Greiling, D. (2014), “Accountability and Trust”, in Bovens, M. and T. Schillemans (eds.), *The Oxford Handbook of Public Accountability*, Oxford University Press, <http://dx.doi.org/10.1093/oxfordhb/9780199641253.013.0023>. [129]

- Greiling, D. and A. Halachmi (2013), “Accountability and Organizational Learning in the Public Sector”, *Public Performance & Management Review*, Vol. 36/3, pp. 380-406, <http://dx.doi.org/10.2753/pmr1530-9576360301>. [170]
- Greiling, D. and K. Spraul (2010), “Accountability and the challenges of information disclosure”, *Public Administration Quarterly*, Vol. 34/3, pp. 338-377. [151]
- Grimmelikhuijsen, S. et al. (2017), “Behavioral Public Administration: Combining Insights from Public Administration and Psychology”, *Public Administration Review*, Vol. 77/1, pp. 45-56, <http://dx.doi.org/10.1111/puar.12609>. [13]
- Gunderson, L. (2002), “Adaptive dancing: interactions between social resilience and ecological crises”, in Berkes, F., J. Colding and C. Folke (eds.), *Navigating Social-Ecological Systems : Building Resilience for Complexity and Change*, Cambridge University Press, Cambridge, <http://dx.doi.org/10.1017/cbo9780511541957.005>. [67]
- Hackenbrack, K. and M. Nelson (1996), “Auditors’ Incentives and Their Application of Financial Accounting Standards”, *The Accounting Review*, Vol. 71/1, pp. 43-59. [137]
- Halachmi, A. (2014), “Accountability Overloads”, in Bovens, M., R. Goodin and T. Schillemans (eds.), *The Oxford Handbook of Public Accountability*, <http://dx.doi.org/10.1093/oxfordhb/9780199641253.013.0011>. [39]
- Halachmi, A. (2005), “Performance measurement: Test the water before you dive in”, *International Review of Administrative Sciences*, Vol. 71/2, pp. 255-266, <http://dx.doi.org/10.1177/0020852305053884>. [219]
- Halachmi, A. (2002), “Performance Measurement, Accountability, and Improved Performance”, *Public Performance & Management Review*, Vol. 25/4, pp. 370-374, <http://dx.doi.org/10.1080/15309576.2002.11643674>. [218]
- Halachmi, A. (2002), “Work Study Performance measurement: a look at some possible dysfunctions”, *Work Study International Journal of Productivity and Performance Management Work Study*, Vol. 51/2, pp. 230-239, <https://doi.org/10.1108/00438020210437240/>. [206]
- Halachmi, A. (2002), “Work Study Performance measurement: a look at some possible dysfunctions”, *Work Study International Journal of Productivity and Performance Management Work Study*, Vol. 51/2, pp. 230-239, <https://doi.org/10.1108/00438020210437240/>. [260]
- Hall, A., D. Frink and M. Buckley (2017), “An accountability account: A review and synthesis of the theoretical and empirical research on felt accountability”, *Journal of Organizational Behavior*, Vol. 38, pp. 204– 224, <http://dx.doi.org/10.1002/job.2052>. [66]
- Hall, A. et al. (2006), “Relationships between felt accountability as a stressor and strain reactions: The neutralizing role of autonomy across two studies”, *Journal of Occupational Health Psychology*, Vol. 11/1, pp. 87–99, <http://dx.doi.org/10.1037/1076-8998.11.1.87>. [65]

- Hanushek, E. and M. Raymond (2002), “Lessons about the Design of State Accountability Systems”, in Peterson, P. and M. West (eds.), *No Child Left Behind? The Politics and Practice of Accountability*, Brookings Institution Press, Washington D.C. [138]
- Harari, M. and C. Rudolph (2017), “The effect of rater accountability on performance ratings: A meta-analytic review”, *Human Resource Management Review*, Vol. 27/1, pp. 121-133, <http://dx.doi.org/10.1016/j.hrmr.2016.09.007>. [262]
- Hardin, R. (1999), “Do we want trust in government?”, in Warren, M. (ed.), *Democracy and trust*, Cambridge University Press. [194]
- Hardin, R. (1993), “The street-level epistemology of trust”, *Politics and Society*, Vol. 21/4, pp. 505-529. [193]
- Hargreaves, A. et al. (2009), *The Learning Mosaic: A multiple perspective review of the Alberta Initiative for School Improvement (AISI)*, Alberta Education, Edmonton. [126]
- Harris, M., D. Smith and D. Champagne (1995), “A field study of performance appraisal based ratings purpose: research versus administrative”, *Personnel Psychology*, Vol. 48, pp. 151-160. [271]
- Head, B. (2008), “Wicked Problems in Public Policy”, *Public Policy*, Vol. 3/2, pp. 101-118. [30]
- Head, B. and J. Alford (2015), “Wicked Problems: Implications for Public Policy and Management”, *Administration and Society*, Vol. 47/6, pp. 711-739, <http://dx.doi.org/10.1177/0095399713481601>. [28]
- Heidemeier, H. and K. Moser (2009), “Self-Other Agreement in Job Performance Ratings: A Meta-Analytic Test of a Process Model”, *Journal of Applied Psychology*, Vol. 94/2, pp. 353-370, <http://dx.doi.org/10.1037/0021-9010.94.2.353>. [279]
- Hextall, I. et al. (2007), *Changing Teacher Roles, Identities and Professionalism: An Annotated Bibliography*, Changing Teacher Roles, Identities and Professionalism’. [127]
- Hochwarter, W. et al. (2007), “Political skill as neutralizer of felt accountability—job tension effects on job performance ratings: A longitudinal investigation”, *Organizational Behavior and Human Decision Processes*, Vol. 102, pp. 226-239, <http://dx.doi.org/10.1016/j.obhdp.2006.09.003>. [140]
- Hochwarter, W. et al. (2005), “Negative affectivity as a moderator of the form and magnitude of the relationship between felt accountability and job tension”, *Journal of Organizational Behavior*, Vol. 26, pp. 517-534, <http://dx.doi.org/10.1002/job.324>. [84]
- Holland, J. (1995), *Hidden order: How adaptation builds complexity*, Addison Wesley, Reading. [63]
- Hondeghem, A. (ed.) (1998), *Remedies for management: changing the mix but not the ingredients?*, European Group of Public Administration., International Institute of Administrative Sciences., & EGPA Annual Conference, IOS Press. [91]

- Hood, C. (2010), *The blame game: Spin, bureaucracy and self-preservation in government.*, [283]
Princeton University Press, Princeton.
- Hood, C. (2007), “Public service management by numbers: Why does it vary? Where has it come from? What are the gaps and the puzzles?”, *Public Money and Management*, Vol. 27/2, pp. 95-102, <http://dx.doi.org/10.1111/j.1467-9302.2007.00564.x>. [214]
- Hood, C. (2006), “Gaming in Targetworld: The Targets Approach to Managing British Public Services”, *Public Administration Review*, Vol. Jul/Aug, pp. 515-521. [215]
- Hood, C. (2004), *Controlling Modern Government. Variety, Commonality and Change*, Edward Elgar Press. [93]
- Hood, C. (2002), “The Risk Game and the Blame Game*”, *Government and Opposition*, Vol. 11/4, pp. 15-37, <http://www.bseinquiry.gov.uk>. [143]
- Hood, C. (1998), “Control and Regulation in Public Management”, in *The Art of the State : Culture, Rhetoric, and Public Management*, Oxford University Press. [103]
- Hood, C. (1998), “Doing Public Management the Fatalist Way?”, in *The Art of the State : Culture, Rhetoric, and Public Management*, Oxford University Press, <http://ebookcentral.proquest.com/lib/uts/detail.action?docID=3052801>. [96]
- Hood, C. (1998), “Taking Stock: The State of the Art of the State”, in *The Art of the State : Culture, Rhetoric, and Public Management*, Oxford University Press. [92]
- Hood, C. and D. Heald (2006), *Transparency: The Key to Better Governance?*, Oxford University Press, Oxford, <http://dx.doi.org/10.5871/bacad/9780197263839.001.0001/upso-9780197263839-chapter-1>. [176]
- Hooge, E. (2016), “Making multiple school accountability work”, in Burns, T. and F. Köster (eds.), *Governing Education in a Complex World*, OECD Publishing, Paris, <https://dx.doi.org/10.1787/9789264255364-7-en>. [47]
- Hooge, E., T. Burns and H. Wilkoszewski (2012), “Looking Beyond the Numbers: Stakeholders and Multiple School Accountability”, *OECD Education Working Papers*, No. 85, OECD Publishing, Paris, <https://dx.doi.org/10.1787/5k91dl7ct6q6-en>. [33]
- Hooghe, L. and G. Marks (2003), “Unraveling the Central State, but How? Types of Multi-level Governance”, *American Political Science Review*, Vol. 97/2, pp. 233-243. [34]
- Hoos, F., J. Pruijssers and M. Lander (2017), “Who’s Watching? Accountability in Different Audit Regimes and the Effects on Auditors’ Professional Skepticism”, *Journal of Business Ethics*, pp. 1-13, <http://dx.doi.org/10.1007/s10551-017-3603-6>. [97]
- Horgan, J. (1995), “From Complexity to Perplexity”, *Scientific American* June, pp. 104-109, <http://dx.doi.org/10.1038/scientificamerican0695-104>. [21]

- Hoxby, C. (2003), "School Choice and School Productivity. Could School Choice Be a Tide that Lifts All Boats?", in Hoxby, C. (ed.), *The Economics of School Choice*, University of Chicago Press, NBER. [303]
- Huang, J. et al. (2014), "Personality and adaptive performance at work: A meta-analytic investigation", *Journal of Applied Psychology*, Vol. 99/1, pp. 162–179, <http://dx.doi.org/10.1037/a0034285>. [282]
- Hui, P. and J. Bonan (1991), *Decentralization and Accountability in Public Education*, RAND Corporation, Santa Monica. [46]
- Huneke, M., C. Cole and I. Levin (2004), "How varying levels of knowledge and motivation affect search and confidence during consideration and choice", *Marketing Letters*, Vol. 15/2-3, pp. 67-79, <http://dx.doi.org/10.1023/B:MARK.0000047385.01483.19>. [257]
- Isett, K. et al. (2011), "Networks in public administration scholarship: Understanding where we are and where we need to go", *Journal of Public Administration Research and Theory*, Vol. 21/1, pp. 151-173, <http://dx.doi.org/10.1093/jopart/muq061>. [117]
- Jaafar, S. and S. Anderson (2007), "Policy trends and tensions in accountability for educational management and services in Canada", *Alberta Journal of Educational Research*, Vol. 53/2, pp. 207-227. [58]
- Jackson, A. et al. (2015), "Too arrogant for their own good? Why and when narcissists dismiss advice", *Organizational Behavior and Human Decision Processes*, Vol. 131, pp. 33-50, <http://dx.doi.org/10.1016/j.obhdp.2015.07.006>. [192]
- Jos, P. and M. Tompkins (2004), "The Accountability Paradox in an Age of Reinvention", *Administration & Society*, Vol. 36/3, pp. 255-281, <http://dx.doi.org/10.1177/0095399704263479>. [148]
- Kapucu, N. and Q. Hu (2014), "Understanding Multiplexity of Collaborative Emergency Management Networks", *American Review of Public Administration*, Vol. 46/4, pp. 399–417, <http://dx.doi.org/10.1177/0275074014555645>. [310]
- Karsten, N. (2015), "Scrutinize me, please! the drivers, manifestations and implications of accountability-seeking behaviour", *Public Administration*, Vol. 93/3, pp. 684–699, <http://dx.doi.org/10.1111/padm.12164>. [133]
- Kasperson, R., D. Golding and S. Tuler (1992), "Social Distrust as a Factor in Siting Hazardous Facilities and Communicating Risks", *Journal of Social Issues*, Vol. 48/4, pp. 161-187, <http://dx.doi.org/10.1111/j.1540-4560.1992.tb01950.x>. [191]
- Kauffman, S. (1993), *The origins of order: self-organisation and selection in evolution*, Oxford University Press, Oxford. [32]
- Kearns, K. (1994), "The Strategic Management of Accountability in Nonprofit Organizations: An Analytical", *Public Administration Review*, Vol. 54/2, pp. 185-192. [70]

- Keren, G. (1991), “Calibration and probability judgements: Conceptual and methodological issues”, *Acta Psychologica*, Vol. 77/3, pp. 217-273, [http://dx.doi.org/10.1016/0001-6918\(91\)90036-Y](http://dx.doi.org/10.1016/0001-6918(91)90036-Y). [321]
- Kim, S. (2005), “Balancing Competing Accountability Requirements: Challenges in Performance Improvement of the Nonprofit Human Services Agency”, *Public Performance & Management Review*, Vol. 29/2, pp. 145-163. [162]
- Kim, S. (2005), “The role of trust in the modern administrative state an integrative model”, *Administration and Society*, Vol. 37/5, pp. 611-635, <http://dx.doi.org/10.1177/0095399705278596>. [182]
- Klijn, E. (2008), “Complexity theory and public administration: What’s new?”, *Public Management Review*, Vol. 10/3, pp. 299-317, <http://dx.doi.org/10.1080/14719030802002675>. [31]
- Klijn, E. and J. Koppenjan (2014), “Complexity in Governance Network Theory”, *Complexity Governance & Networks*, pp. 61-70, <http://dx.doi.org/10.7564/14-CGN8>. [68]
- Klingner, D., J. Nalbandian and B. Romzek (2002), “Politics, administration, and markets - Conflicting expectations and accountability”, *American Review of Public Administration*, Vol. 32/2, pp. 117-144. [51]
- Knapp, M. and S. Feldman (2012), “Managing the intersection of internal and external accountability: Challenge for urban school leadership in the United States”, *Journal of Educational Administration*, Vol. 50/2, pp. 666-694, <http://dx.doi.org/10.1108/09578231211249862>. [128]
- Koning, P. and K. Van Der Wiel (2010), “Ranking the Schools: How Quality Information Affects School Choice in the Netherlands”, *IZA Discussion Paper*, No. 4984, Institute for the Study of Labor, Bonn. [295]
- Koonce, L., U. Anderson and G. Marchant (1995), “Justification of Decisions in Auditing”, *Journal of Accounting Research*, Vol. 33/2, pp. 369-384. [202]
- Koop, C. (2014), “Theorizing and explaining voluntary accountability”, *Public Administration*, Vol. 92/3, pp. 565-581, <http://dx.doi.org/10.1111/padm.12058>. [134]
- Koppell, J. (2005), “Pathologies of accountability: ICANN and the challenge of “multiple accountabilities disorder””, *Public Administration Review*, Vol. 65/1, pp. 94-108, <http://dx.doi.org/10.1111/j.1540-6210.2005.00434.x>. [50]
- Landau, M. (1969), “Redundancy, Rationality, and the Problem of Duplication and Overlap”, *Public Administration Review*, Vol. 29/4, pp. 346-358, <http://dx.doi.org/10.2307/973247>. [98]
- Langhe, B., S. Van Osselaer and B. Wierenga (2011), “The effects of process and outcome accountability on judgment process and performance”, *Organizational Behavior and Human Decision Processes*, Vol. 115, pp. 238–252, <http://dx.doi.org/10.1016/j.obhdp.2011.02.003>. [233]

- Lanivich, S. et al. (2010), "P-E Fit as moderator of the accountability - employee reactions relationships: Convergent results across two samples", *Journal of Vocational Behavior*, Vol. 77, pp. 425-436, <http://dx.doi.org/10.1016/j.jvb.2010.05.004>. [75]
- Latane, B., K. Williams and S. Harkins (1979), "Many Hands Make Light the Work: The Causes and Consequences of Social Loafing", *Journal of Personality and Social Psychology*, Vol. 37/6, pp. 822-832. [319]
- Lawrence, P. and J. Lorsch (1967), *Organization and Environment*, Harvard University Press, Cambridge. [74]
- Le Grand, J. (ed.) (2007), "School Education", in *The Other Invisible Hand : Delivering Public Services Through Choice and Competition*, Princeton University Press, Princeton. [320]
- Lee, H. et al. (1999), "Motivated search: Effects of choice accountability, issue involvement, and prior knowledge on information acquisition and use", *Journal of Business Research*, Vol. 45, pp. 75-88, [http://dx.doi.org/10.1016/S0148-2963\(98\)00067-8](http://dx.doi.org/10.1016/S0148-2963(98)00067-8). [185]
- Lee, S. and C. Lee (2015), "Creative Interaction and Multiplexity in Intraorganizational Networks", *Management Communication Quarterly*, Vol. 29/1, pp. 56-83, <http://dx.doi.org/10.1177/0893318914553120>. [311]
- Lerner, A. (1987), "There is more than one way to be redundant", *Administration and Society*, Vol. 18/3, pp. 334-359. [100]
- Lerner, J. and P. Tetlock (1999), "Accounting for the Effects of Accountability", *Psychological Bulletin*, Vol. 125/2, pp. 255-275. [61]
- Levi, M. (1998), "A State of Trust", in Braithwaite, V. and M. Levi (eds.), *Trust and Governance*, Russell Sage Foundation. [177]
- Levine, M. and J. Forrence (1990), "Regulatory Capture, Public Interest, and the Public Agenda: Toward a Synthesis", *Journal of Law, Economics, and Organization*, Vol. 6/SI, pp. 167-198, http://dx.doi.org/10.1093/jleo/6.special_issue.167. [204]
- Levinthal, D. and J. March (1993), "The myopia of learning", *Strategic Management Journal*, Vol. 14, pp. 95-112, <http://dx.doi.org/10.1002/smj.4250141009>. [222]
- Levy, P. and J. Williams (2004), "The social context of performance appraisal: A review and framework for the future", *Journal of Management*, Vol. 30/6, pp. 881-905, <http://dx.doi.org/10.1016/j.jm.2004.06.005>. [264]
- Lewicki, R. and B. Bunker (1996), "Developing and Maintaining Trust in Work Relationships", in Kramer, R. and T. Tyler (eds.), *Trust in Organizations: Frontiers of Theory and Research*, <http://dx.doi.org/10.4135/9781452243610.n7>. [131]
- Lewis, J., J. O'Flynn and H. Sullivan (2014), "Accountability: To Whom, in Relation to What, and Why?", *Australian Journal of Public Administration*, Vol. 73/4, pp. 401-407, <http://dx.doi.org/10.1111/1467-8500.12104>. [9]

- Lewis, J. and P. Triantafillou (2012), “From performance measurement to learning: A new source of government overload?”, *International Review of Administrative Sciences*, Vol. 78/4, pp. 597–614, <http://dx.doi.org/10.1177/0020852312455993>. [145]
- Lindberg, S. (2013), “Mapping accountability: core concept and subtypes”, *International Review of Administrative Sciences*, Vol. 79/2, pp. 202-226, <http://dx.doi.org/10.1177/0020852313477761>. [2]
- Linhorst, D., A. Eckert and G. Hamilton (2005), “Participation in Organizational Decision Making by Clients with SMI”, *Social Work*, Vol. 50/1, pp. 21-30. [197]
- Linn, R. (2003), “Accountability: Responsibility and Reasonable Expectations”, *Educational Researcher*, Vol. 32/7, pp. 3-13, <http://dx.doi.org/10.3102/0013189X032007003>. [48]
- Liu, X. et al. (2019), “Multidimensionality: A Cross-Disciplinary Review and Integration”, *Journal of Management*, Vol. 45/1, pp. 197–230, <http://dx.doi.org/10.1177/0149206318807285>. [312]
- Lord, A. (1992), “Pressure: A Methodological Consideration for Behavioral Research in Auditing”, *Auditing*, Vol. 11/2, pp. 90-109. [273]
- Lorenz, E. (1972), *Predictability: Does the flap of a butterfly’s wings in Brazil set off a tornado in Texas?*. [80]
- Macintosh, R. and D. Maclean (1999), “Conditioned emergence: A Dissipative structures approach to transformation”, *Strategic Management Journal*, Vol. 20, pp. 297-316, <http://dx.doi.org/10.4324/9781315887784>. [85]
- Maguire, S. and B. McKelvey (1999), “Complexity and Management: Moving From Fad To Firm Foundations”, *Emergence*, Vol. 1/2, pp. 19-61, http://dx.doi.org/10.1207/s15327000em0102_3. [24]
- Majone, G. (2011), “Strategic issues in risk regulation”, in Levi-Faur, D. (ed.), *Handbook on the Politics of Regulation*. [95]
- Mansbridge, J. (2014), “A Contingency Theory of Accountability”, in Bovens, M., R. Goodin and T. Schillemans (eds.), *The Oxford Handbook of Public Accountability*, <http://dx.doi.org/10.1093/oxfordhb/9780199641253.013.0019>. [183]
- Mansbridge, J. (2009), “A selection model of political representation”, *Journal of Political Philosophy*, Vol. 17/4, pp. 369–398, <http://dx.doi.org/10.1111/j.1467-9760.2009.00337.x>. [130]
- Mansouri, M. and J. Rowney (2014), “The Dilemma of Accountability for Professionals: A Challenge for Mainstream Management Theories”, *Journal of Business Ethics*, Vol. 123, pp. 45–56, <http://dx.doi.org/10.1007/s10551-013-1788-x>. [124]
- Mante, B. and G. O’Brien (2002), “Efficiency measurement of Australian public sector organisations”, *Journal of Educational Administration*, Vol. 40/3, pp. 274-296, <http://dx.doi.org/10.1108/09578230210427181>. [302]

- March, J. (1991), “Exploration and Exploitation in Organizational Learning”, *Organization Science*, Vol. 2/1, pp. 71-87, <http://dx.doi.org/10.1287/orsc.2.1.71>. [224]
- Marks, H. and J. Nance (2007), “Contexts of accountability under systemic reform: Implications for principal influence on instruction and supervision”, *Educational Administration Quarterly*, Vol. 43/1, pp. 3-37, <http://dx.doi.org/10.1177/0013161X06291414>. [299]
- Mashaw, J. (2014), “Accountability and Time”, in Bovens, M., R. Goodin and T. Schillemans (eds.), *The Oxford Handbook of Public Accountability*, Oxford University Press, <http://dx.doi.org/10.1093/oxfordhb/9780199641253.013.0031>. [142]
- Mason, M. (2008), “What is complexity theory and what are its implications for educational change?”, *Educational Philosophy and Theory*, Vol. 40/1, pp. 35-49, <http://dx.doi.org/10.1111/j.1469-5812.2007.00413.x>. [17]
- Matthews, P. and M. Ehren (2017), “Accountability and Improvement in Self-improving School Systems”, in Earley, P. and T. Greany (eds.), *School Leadership and Education System Reform*, Bloomsbury Publishing, London. [125]
- Mayer, R., J. Davis and F. Schoorman (1995), “An Integrative Model of Organizational Trust”, *The Academy of Management Review*, Vol. 20/3, pp. 709-734. [196]
- Mero, N., R. Guidice and A. Brownlee (2007), “Accountability in a performance appraisal context: The effect of audience and form of accounting on rater response and behavior”, *Journal of Management*, Vol. 33/2, pp. 223-252, <http://dx.doi.org/10.1177/0149206306297633>. [187]
- Mero, N. and S. Motowidlo (1995), “Effects of Rater Accountability on the Accuracy and the Favorability of Performance Ratings”, *Journal of Applied Psychology*, Vol. 80/4, pp. 517-524, <http://dx.doi.org/10.1037/0021-9010.80.4.517>. [289]
- Mitchell, M. (2009), “What is complexity?”, in *Complexity: A Guided tour*, Oxford University Press, New York, <http://dx.doi.org/10.1002/bies.10192>. [20]
- Mitleton-Kelly, E. (2003), “Ten Principles of Complexity and Enabling Infrastructures”, in Mitleton-Kelly, E. (ed.), *Complex Systems and Evolutionary Perspectives of Organisations: The Application of Complexity Theory to Organisations*, Elsevier, Amsterdam. [22]
- Moore, M. (2006), “Recognising Public Value: The Challenge of Measuring Performance In Government”, in Wanna, J. (ed.), *A Passion for Policy: Essays in Public Sector Reform*, ANU Press, Canberra, <http://about.jstor.org/terms>. [52]
- Moynihan, D. (2005), “Goal-Based Learning and the Future of Performance Management”, *Public Administration Review*, Vol. 65/2, pp. 203-216, <http://dx.doi.org/10.1111/j.1540-6210.2005.00445.x>. [280]
- Mulgan, R. (2014), “Accountability Deficits”, in Bovens, M., R. Goodin and T. Schillemans (eds.), *The Oxford Handbook of Public Accountability*, <http://dx.doi.org/10.1093/oxfordhb/9780199641253.013.0010>. [38]

- Mulgan, R. (2000), “‘Accountability’: An Ever-Expanding Concept?”, *Public Administration*, Vol. 78/3, pp. 555-573. [90]
- Neale, A. and B. Anderson (2000), “Performance reporting for accountability purposes: lessons, issues, future”, *International Public Management Journal*, Vol. 30, pp. 93-106. [258]
- O’Connell, L. (2005), “Program accountability as an emergent property: The role of stakeholders in a program’s field”, *Public Administration Review*, Vol. 65/1, pp. 85-93, <http://dx.doi.org/10.1111/j.1540-6210.2005.00433.x>. [115]
- O’Day, J. (2002), “Complexity, Accountability, and School Improvement”, *Harvard Educational Review*, Vol. 72/3, pp. 293-329, <http://dx.doi.org/10.17763/haer.72.3.021q742t8182h238>. [27]
- O’Donnell, G. (1998), “Horizontal Accountability in New Democracies”, *Journal of Democracy*, Vol. 9/3, pp. 112-126, <http://dx.doi.org/10.1353/jod.1998.0051>. [113]
- OECD (2017), *Systems Approaches to Public Sector Challenges: Working with Change*, OECD Publishing, Paris, <https://dx.doi.org/10.1787/9789264279865-en>. [19]
- Olsen, J. (2013), “The Institutional Basis of Democratic Accountability”, *West European Politics*, Vol. 36/3, pp. 447-473, <http://dx.doi.org/10.1080/01402382.2012.753704>. [172]
- Osborne, S. (2006), “The New Public Governance?”, *Public Management Review*, Vol. 8/3, pp. 377-387, <http://dx.doi.org/10.1080/14719030600853022>. [35]
- Page, S. (2004), “Measuring accountability for results in interagency collaborates”, *Public Administration Review*, Vol. 64/5, pp. 591-606, <http://dx.doi.org/10.1111/j.1540-6210.2004.00406.x>. [116]
- Pascale, R. (1999), “Surfing the Edge of Chaos”, *Sloan Management Review*, Vol. 40/3, pp. 83-94. [88]
- Patil, S. and P. Tetlock (2014), “Punctuated incongruity: A new approach to managing trade-offs between conformity and deviation”, *Research in Organizational Behavior*, Vol. 34, pp. 155-171, <http://dx.doi.org/10.1016/j.riob.2014.08.002>. [87]
- Patil, S., P. Tetlock and B. Mellers (2017), “Accountability Systems and Group Norms: Balancing the Risks of Mindless Conformity and Reckless Deviation”, *Journal of Behavioral Decision Making*, Vol. 30, pp. 282–303, <http://dx.doi.org/10.1002/bdm.1933>. [223]
- Patil, S., F. Vieider and P. Tetlock (2014), “Process Versus Outcome Accountability”, in Bovens, M., R. Goodin and T. Schillemans (eds.), *The Oxford Handbook of Public Accountability*, <http://dx.doi.org/10.1093/oxfordhb/9780199641253.013.0002>. [179]
- Paul, S. (1992), “Accountability in public services: Exit, voice and control”, *World Development*, Vol. 20/7, pp. 1047-1060, [http://dx.doi.org/10.1016/0305-750X\(92\)90130-N](http://dx.doi.org/10.1016/0305-750X(92)90130-N). [305]

- Pennington, J. and B. Schlenker (1999), “Accountability for consequential decisions: Justifying ethical judgments to audiences”, *Personality and Social Psychology Bulletin*, Vol. 25/9, pp. 1067-1081, <http://dx.doi.org/10.1177/01461672992512001>. [201]
- Perryman, J. (2009), “Inspection and the fabrication of professional and performative processes”, *Journal of Education Policy*, Vol. 24/5, pp. 611-631, <http://dx.doi.org/10.1080/02680930903125129>. [296]
- Peters, B. (2014), “Accountability in Public Administration”, in *The Oxford Handbook of Public Accountability*, Oxford University Press, <http://dx.doi.org/10.1093/oxfordhb/9780199641253.013.0032>. [102]
- Pierre, J. and B. Peters (2005), *Governing complex societies : trajectories and scenarios*, Palgrave Macmillan, New York. [7]
- Pollitt, C. and P. Hupe (2011), “The role of magic concepts”, *Public Management Review*, Vol. 13/5, pp. 641-658, <http://dx.doi.org/10.1080/14719037.2010.532963>. [10]
- Pollmann, M., J. Potters and S. Trautmann (2014), “Risk taking by agents: The role of ex-ante and ex-post accountability”, *Economics Letters*, Vol. 123/3, pp. 387-390, <http://dx.doi.org/10.1016/j.econlet.2014.04.004>. [209]
- Power, M. (1997), *The Audit Society: Rituals of Verification*, Oxford University Press. [146]
- Prigogine, I. and I. Stengers (1984), *Order Out of Chaos: Man’s New Dialogue with Nature*, Bantam, New York. [78]
- Quetzal Tritter, J. and A. McCallum (2006), “The snakes and ladders of user involvement: Moving beyond Arnstein”, *Health Policy*, Vol. 76, pp. 156-168, <http://dx.doi.org/10.1016/j.healthpol.2005.05.008>. [198]
- Rabinowitz, S. (2018), *Australia: using open school data to improve transparency and accountability: Ethics and corruption in education*, UNESCO International Institute for Educational Planning, Paris, <http://www.unesco.org/open-access/terms-use-ccbysa-en>. [174]
- Radin, B. and B. Romzek (1996), “Accountability Expectations in an Intergovernmental Arena: The National Rural Development Partnership”, *The Journal of Federalism*, Vol. 26/2, pp. 59-81. [108]
- Reichersdorfer, J., T. Christensen and K. Vrangbæk (2013), “Accountability of immigration administration: comparing crises in Norway, Denmark and Germany”, *International Review of Administrative Sciences*, Vol. 79/2, pp. 271-291, <http://dx.doi.org/10.1177/0020852313478251>. [57]
- Rhodes, R. (2008), “Policy Network Analysis”, in Goodin, R., M. Moran and M. Rein (eds.), *The Oxford Handbook of Public Policy*, Oxford University Press, <http://dx.doi.org/10.1093/oxfordhb/9780199548453.003.0020>. [119]
- Rhodes, R. (1997), *Understanding governance: policy networks, reflexivity, and accountability*, Open University Press, Buckingham. [45]

- Rhodes, R. and J. Wanna (2007), “The limits to public value, or rescuing responsible government from the platonic guardians”, *Australian Journal of Public Administration*, Vol. 66/4, pp. 406-421, <http://dx.doi.org/10.1111/j.1467-8500.2007.00553.x>. [154]
- Roch, S. and L. McNall (2007), “An investigation of factors influencing accountability and performance ratings”, *Journal of Psychology*, Vol. 141/5, pp. 499-524, <http://dx.doi.org/10.3200/JRLP.141.5.499-524>. [110]
- Romzek, B. (2014), “Living accountability: Hot rhetoric, cool theory, and uneven practice”, *Political Science and Politics*, Vol. 48/1, pp. 27-34, <http://dx.doi.org/10.1017/S1049096514001553>. [43]
- Romzek, B. (2011), “The Tangled Web of Accountability in Contracting Networks: The Case of Welfare Reform”, in Dubnick, M. and H. Frederickson (eds.), *Accountable Government : Problems and Promises*, Routledge. [42]
- Romzek, B. (2000), “Dynamics of public sector accountability in an era of reform”, *International Review of Administrative Sciences*, Vol. 66/1, pp. 21-44. [4]
- Romzek, B. and M. Dubnick (1998), “Accountability”, in Shafritz, J. (ed.), *International Encyclopedia of Public Policy and Administration*, Westview Press. [37]
- Romzek, B. and M. Dubnick (1987), “Accountability in the Public Sector: Lessons from the Challenger Tragedy”, *Public Administration Review*, Vol. 47/3, p. 227, <http://dx.doi.org/10.2307/975901>. [122]
- Romzek, B. and P. Ingraham (2000), “Cross Pressures of Accountability: Initiative, Command, and Failure in the Ron Brown”, Vol. 60/3, pp. 240-253, <http://www.jstor.org/stable/>. [153]
- Romzek, B., K. LeRoux and J. Blackmar (2012), “A Preliminary Theory of Informal Accountability among Network Organizational Actors”, *Public Administration Review*, Vol. 72/3, pp. 442-453. [41]
- Romzek, B. et al. (2014), “Informal accountability in multisector service delivery collaborations”, *Journal of Public Administration Research and Theory*, Vol. 24/4, pp. 813-842, <http://dx.doi.org/10.1093/jopart/mut027>. [44]
- Ron, N., R. Lipshitz and M. Popper (2006), “How organizations learn: Post-flight reviews in an F-16 fighter squadron”, *Organization Studies*, Vol. 27/8, pp. 1069–1089, <http://dx.doi.org/10.1177/0170840606064567>. [195]
- Rosenkvist, M. (2010), “Using Student Test Results for Accountability and Improvement: A Literature Review”, *OECD Education Working Papers*, No. 54, OECD Publishing, Paris, <https://dx.doi.org/10.1787/5km4htwzvbv30-en>. [323]
- Ryan, R. and E. Deci (2006), “Self-regulation and the problem of human autonomy: Does psychology need choice, self-determination, and will?”, *Journal of Personality*, Vol. 74/6, pp. 1557-1586, <http://dx.doi.org/10.1111/j.1467-6494.2006.00420.x>. [244]

- Ryan, R. and E. Deci (2000), “Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions”, *Contemporary Educational Psychology*, Vol. 25, pp. 54-67, <http://dx.doi.org/10.1006/ceps.1999.1020>. [251]
- Sahlberg, P. (2010), “Rethinking accountability in a knowledge society”, *Journal of Educational Change*, Vol. 11/1, pp. 45–61, <http://dx.doi.org/10.1007/s10833-008-9098-2>. [56]
- Sahlberg, P. (2007), “Education policies for raising student learning: The Finnish approach”, *Journal of Education Policy*, Vol. 22/2, pp. 147–171, <http://dx.doi.org/10.1080/02680930601158919>. [55]
- Salamon, L. (2001), “Introduction”, in Salamon, L. (ed.), *Handbook of Policy Instruments*, Oxford University Press, New York. [307]
- Schillemans, T. (2018), *Lessons from five decades of experimental and behavioral research on accountability: A systematic literature review*. [261]
- Schillemans, T. (2016), “Calibrating Public Sector Accountability: Translating experimental findings to public sector accountability”, *Public Management Review*, Vol. 18/9, pp. 1400-1420, <http://dx.doi.org/10.1080/14719037.2015.1112423>. [14]
- Schillemans, T. (2013), “Moving Beyond The Clash of Interests: On stewardship theory and the relationships between central government departments and public agencies”, *Public Management Review*, Vol. 15/4, pp. 541-562, <http://dx.doi.org/10.1080/14719037.2012.691008>. [207]
- Schillemans, T. (2011), “Does horizontal accountability work? Evaluating potential remedies for the accountability deficit of agencies”, *Administration and Society*, <http://dx.doi.org/10.1177/0095399711412931>. [285]
- Schillemans, T. (2010), “Redundant Accountability: The Joint Impact of Horizontal and Vertical Accountability on Autonomous Agencies”, *Public Administration Quarterly*, Vol. 34/3, pp. 300-337. [101]
- Schillemans, T. and M. Bovens (2011), “The Challenge of Multiple Accountability: Does Redundancy Lead to Overload?”, in Dubnick, M. and H. Frederickson (eds.), *Accountable Government : Problems and Promises*, Routledge. [99]
- Schillemans, T. and M. Busuioc (2015), “Predicting public sector accountability: From agency drift to forum drift”, *Journal of Public Administration Research and Theory*, Vol. 25/1, pp. 191-215, <http://dx.doi.org/10.1093/jopart/muu024>. [181]
- Schillemans, T. and R. Smulders (2016), “Learning from accountability?! Whether, what, and when”, *Public Performance and Management Review*, Vol. 39/1, pp. 248-271, <http://dx.doi.org/10.1080/15309576.2016.1071175>. [72]
- Schillemans, T., M. Van Twist and I. Vanhomerig (2013), “Innovations in Accountability”, *Public Performance & Management Review*, Vol. 36/3, pp. 407-435, <http://dx.doi.org/10.2753/pmr1530-9576360302>. [71]

- Schillemans, T., M. Van Twist and I. Vanhommerig (2013), “Innovations in Accountability”, [259]
Public Performance & Management Review, Vol. 36/3, pp. 407-435,
<http://dx.doi.org/10.2753/PMR1530-9576360302>.
- Schlenker, B. et al. (1994), “The triangle model of responsibility.”, [190]
Psychological Review, Vol. 101/4, pp. 632-652, <http://dx.doi.org/10.1037/0033-295X.101.4.632>.
- Schneider, V. (2012), “Governance and complexity”, in Levi-Faur, D. (ed.), *Oxford Handbook of Governance*, Oxford University Press, Oxford. [8]
- Scholten, L. et al. (2007), “Motivated information processing and group decision-making: [276]
 Effects of process accountability on information processing and decision quality”, *Journal of Experimental Social Psychology*, Vol. 43, pp. 539-552,
<http://dx.doi.org/10.1016/j.jesp.2006.05.010>.
- Sedikides, C. et al. (2002), “Accountability as a deterrent to self-enhancement: The search for [189]
 mechanisms”, *Journal of Personality and Social Psychology*, Vol. 83/3, pp. 592-605,
<http://dx.doi.org/10.1037/0022-3514.83.3.592>.
- Shepard, L. (2000), “The Role of Assessment in a Learning Culture”, *Educational Researcher*, [298]
 Vol. 29/7, pp. 4-14.
- Shore, T. and A. Tashchian (2007), “Effects of feedback accountability and self-rating [269]
 information on employee appraisals: A replication and extension”, *Psychological Reports*,
 Vol. 100, pp. 1091-1100.
- Siegel-Jacobs, K. and J. Yates (1996), “Effects of procedural and outcome accountability on [234]
 judgment quality”, *Organizational Behavior and Human Decision Processes*, Vol. 65/1,
 pp. 1-17, <http://dx.doi.org/10.1006/obhd.1996.0001>.
- Simons, R. (2005), *Levers of Organization Design How Managers Use Accountability Systems [221]
 for Greater Performance and Commitment*, Harvard Business School Press, Boston, MA.
- Simons, T. and R. Peterson (2000), “Task conflict and relationship conflict in top management [318]
 teams: The pivotal role of intragroup trust”, *Journal of Applied Psychology*, Vol. 85/1,
 pp. 102-111, <http://dx.doi.org/10.1037/0021-9010.85.1.102>.
- Sinclair, A. (1995), “The chameleon of accountability: Forms and discourses”, *Accounting, [123]
 Organizations and Society*, Vol. 20/2-3, pp. 219-237, [http://dx.doi.org/10.1016/0361-3682\(93\)E0003-Y](http://dx.doi.org/10.1016/0361-3682(93)E0003-Y).
- Sirotnik, K. (2004), *Holding Accountability Accountable: What Ought to Matter in Public [294]
 Education*, Teachers College Press, New York.
- Sitkin, S. et al. (2011), “The paradox of stretch goals: Organizations in pursuit of the seemingly [226]
 impossible”, *Academy of Management Review*, Vol. 36/3, pp. 544–566,
<http://dx.doi.org/10.5465/AMR.2011.61031811>.

- Skitka, L., K. Mosier and M. Burdick (2000), "Accountability and automation bias", [237]
International Journal of Human Computer Studies, Vol. 52, pp. 701-717,
<http://dx.doi.org/10.1006/ijhc.1999.0349>.
- Skitka, L., K. Mosier and M. Burdick (1999), "Does automation bias decision-making?", [235]
International Journal of Human Computer Studies, Vol. 51, pp. 991-1006,
<http://dx.doi.org/10.1006/ijhc.1999.0252>.
- Smith, A. (1759), *The theory of moral sentiments*, A. Millar. [168]
- Smith, P. (1995), "On the unintended consequences of publishing performance data in the public [149]
sector", *International Journal of Public Administration*, Vol. 18/2-3, pp. 277-310,
<http://dx.doi.org/10.1080/01900699508525011>
[org/10.1080/01900699508525011](http://dx.doi.org/10.1080/01900699508525011).
- Smith, W. (2016), "Exploring accountability: National testing policies and student achievement", [175]
in Burns, T. and F. Köster (eds.), *Governing Education in a Complex World*, OECD
Publishing, Paris, <http://dx.doi.org/10.1787/9789264255364-6-en>.
- Snyder, S. (2013), "The Simple, the Complicated, and the Complex: Educational Reform [15]
Through the Lens of Complexity Theory", *OECD Education Working Papers*, No. 96, OECD
Publishing, Paris, <https://dx.doi.org/10.1787/5k3txnpt1lnr-en>.
- Spence, J. and L. Keeping (2011), "Conscious rating distortion in performance appraisal: A [270]
review, commentary, and proposed framework for research", *Human Resource Management
Review*, Vol. 21, pp. 85-95, <http://dx.doi.org/10.1016/j.hrmr.2010.09.013>.
- Spreitzer, G. (1996), "Social Structural Characteristics of Psychological Empowerment", [241]
The Academy of Management Journal, Vol. 39/2, pp. 483-504, <https://about.jstor.org/terms>.
- Spreitzer, G. (1995), "Psychological empowerment in the workplace: dimensions, measurement, [243]
and validation", *Academy of Management Journal*, Vol. 38/5, pp. 1442-1465.
- Stacey, R. (1995), "The Science of Complexity: An Alternative Perspective for Strategic Change [64]
Processes", *Management Journal*, Vol. 16/6, pp. 477-495.
- Staw, B. and R. Boettger (1990), "Task Revision: A Neglected Form of Work Performance", [152]
Academy of Management Journal, Vol. 33/3, pp. 534-559, <http://dx.doi.org/10.5465/256580>.
- Steinbauer, R. et al. (2014), "Ethical Leadership and Followers' Moral Judgment: The Role of [167]
Followers' Perceived Accountability and Self-leadership", *Journal of Business Ethics*,
Vol. 120, pp. 381-392, <http://dx.doi.org/10.1007/s10551-013-1662-x>.
- Stoker, G. (2006), "Public value management: A new narrative for networked governance?", [53]
American Review of Public Administration, Vol. 36/1, pp. 41-57,
<http://dx.doi.org/10.1177/0275074005282583>.
- Streeter, C. (1992), "Redundancy in Organizational Systems", *Social Service Review*, Vol. Mar, [309]
pp. 97-111.

- Strøm, K. (2000), “Delegation and accountability in parliamentary democracies”, *European Journal of Political Research*, Vol. 37, pp. 261-289. [112]
- Strøm, K. (1997), “Democracy, accountability, and coalition bargaining The 1996 Stein Rokkan Lecture*”, *European Journal of Political Research*, Vol. 31, pp. 47-62. [111]
- Strøm, K., C. Müller, Wolfgang and T. Bergman (2003), *Delegation and Accountability in Parliamentary Democracies*, Oxford University Press, New York. [173]
- Styhre, A. (2002), “Nonlinear change in organizations: organization change management informed by complexity theory”, *Leadership & Organization Development Journal*, Vol. 23/6, pp. 343-351, <http://dx.doi.org/10.1108/01437730210441300>. [69]
- Suspitsyna, T. (2010), “Accountability in American education as a rhetoric and a technology of governmentality”, *Journal of Education Policy*, Vol. 25/5, pp. 567-586, <http://dx.doi.org/10.1080/02680930903548411>. [5]
- Teisman, G. and A. Van Buuren (2007), “Implementing NPM: A complexity perspective on public management reform trajectories”, in Pollitt, C., S. van Thiel and V. Homburg (eds.), *New Public Management in Europe: Adaptation and Alternatives*, Palgrave Macmillan, <http://dx.doi.org/10.1057/9780230625365>. [25]
- Tetlock, P. (1992), “The impact of accountability on judgment and choice: Toward a social contingency model”, *Advances in Experimental Social Psychology*, Vol. 25, pp. 331-376, [http://dx.doi.org/10.1016/S0065-2601\(08\)60287-7](http://dx.doi.org/10.1016/S0065-2601(08)60287-7). [12]
- Tetlock, P. (1985), “Accountability: A Social Check on the Fundamental Attribution Error”, *Social Psychology Quarterly*, Vol. 48/3, pp. 227-236, <http://dx.doi.org/10.2307/3033683>. [210]
- Tetlock, P. (1983), “Accountability and complexity of thought”, *Journal of Personality and Social Psychology*, Vol. 45/1, pp. 74-83, <http://dx.doi.org/10.1037/0022-3514.45.1.74>. [272]
- Tetlock, P. (1983), “Accountability and the perseverance of first impressions”, *Social Psychology Quarterly*, Vol. 46/4, pp. 285-292. [211]
- Tetlock, P. and R. Boettger (1994), “Accountability amplifies the status quo effect when change creates victims”, *Journal of Behavioral Decision Making*, Vol. 7, pp. 1-23, <http://dx.doi.org/10.1002/bdm.3960070102>. [266]
- Tetlock, P. and J. Kim (1987), “Accountability and Judgment Processes in a Personality Prediction Task”, *Journal of Personality and Social Psychology*, Vol. 52/4, pp. 700-709, <http://dx.doi.org/10.1037/0022-3514.52.4.700>. [256]
- Tetlock, P. and B. Mellers (2011), “Intelligent Management of Intelligence Agencies: Beyond Accountability Ping-Pong”, *American Psychologist*, Vol. 66/6, pp. 542-554, <http://dx.doi.org/10.1037/a0024285>. [216]
- Tetlock, P. and B. Mellers (2011), “Structuring Accountability Systems in Organizations: Key Trade-Offs and Critical Unknowns”, in Fischhoff, B. and C. Chauvin (eds.), *Intelligence Analysis : Behavioral and Social Scientific Foundations*, National Academies Press. [86]

- Tetlock, P., L. Skitka and R. Boettger (1989), “Social and Cognitive Strategies for Coping With Accountability: Conformity, Complexity, and Bolstering”, *Journal of Personality and Social Psychology*, Vol. 57/4, pp. 632-640. [73]
- Tetlock, P. et al. (2013), “Accountability and ideology: When left looks right and right looks left”, *Organizational Behavior and Human Decision Processes*, Vol. 122, pp. 22-35, <http://dx.doi.org/10.1016/j.obhdp.2013.03.007>. [77]
- Thiel, S. and F. Leeuw (2002), “The Performance Paradox in the Public Sector”, *Public Performance & Management Review*, Vol. 25/3, pp. 267-281, <http://dx.doi.org/10.2307/3381236>. [150]
- Thomas, K. and B. Velthouse (1990), “Cognitive Elements of Empowerment: An “Interpretive” Model of Intrinsic Task Motivation”, *Academy of Management Review*, Vol. 15/4, pp. 666-681, <http://dx.doi.org/10.5465/amr.1990.4310926>. [249]
- Thompson, D. (1980), “Moral Responsibility of Public Officials: The Problem of Many Hands”, *American Political Science Review* 74, pp. 905-916, <http://dx.doi.org/10.2307/1954312>. [36]
- Torfinn, J. et al. (2012), “Horizontal, vertical, and diagonal governance”, in Torfinn, J. et al. (eds.), *Interactive Governance: advancing the paradigm*, Oxford University Press, New York. [40]
- Tversky, A. and D. Kahneman (1991), “Loss Aversion in Riskless Choice: A Reference-Dependent Model”, *The Quarterly Journal of Economics*, Vol. 106/4, pp. 1039-1061, <http://dx.doi.org/10.2307/2937956>. [132]
- Tyler, T. (1987), “Conditions Leading to Value-Expressive Effects in Judgments of Procedural Justice: A Test of Four Models”, *Journal of Personality and Social Psychology*, Vol. 52/2, pp. 333-344, <http://dx.doi.org/10.1037/0022-3514.52.2.333>. [178]
- Van de Walle, S. and G. Bouckaert (2003), “Public Service Performance and Trust in Government: The Problem of Causality”, *International Journal of Public Administration*, Vol. 26/8-9, pp. 891-913, <http://dx.doi.org/10.1081/PAD-120019352>. [308]
- van Twist, M. et al. (2013), “Coping with Very Weak Primary Schools: Towards Smart Interventions in Dutch Education Policy”, *OECD Education Working Papers*, No. 98, OECD Publishing, Paris, <https://dx.doi.org/10.1787/5k3txnpnhld7-en>. [81]
- Vibert, F. (2014), “The Need for a Systemic Approach”, in Bovens, M., R. Goodin and T. Schillemans (eds.), *The Oxford Handbook of Public Accountability*, Oxford University Press, New York, <http://dx.doi.org/10.1093/oxfordhb/9780199641253.013.0001>. [62]
- Walker, A. and J. Smither (1999), “A five-year study of upward feedback: What managers do with their results matters”, *Personnel Psychology*, Vol. 52, pp. 393-423, <http://dx.doi.org/10.1111/j.1744-6570.1999.tb00166.x>. [290]
- Wallace, J. et al. (2011), “Structural and Psychological Empowerment Climates, Performance, and the Moderating Role of Shared Felt Accountability: A Managerial Perspective”, *Journal of Applied Psychology*, Vol. 96/4, pp. 840-850, <http://dx.doi.org/10.1037/a0022227>. [315]

- Warren, M. (2014), “Accountability and Democracy”, in Bovens, M., R. Goodin and T. Schillemans (eds.), *The Oxford Handbook of Public Accountability*, <http://dx.doi.org/10.1093/oxfordhb/9780199641253.013.0018>. [171]
- Waslander, S., C. Pater and M. van der Weide (2010), “Markets in Education: An Analytical Review of Empirical Research on Market Mechanisms in Education”, *OECD Education Working Papers*, No. 52, OECD Publishing, Paris, <https://dx.doi.org/10.1787/5km4pskmkr27-en>. [106]
- Wayne, S. and G. Ferris (1990), “Influence Tactics, Affect, and Exchange Quality in Supervisor-Subordinate Interactions: A Laboratory Experiment and Field Study”, *Journal of Applied Psychology*, Vol. 75/5, pp. 487-499. [267]
- Weber, E. and A. Khademian (2008), “Wicked Problems, Knowledge Challenges, and Collaborative Capacity Builders in Network Settings”, *Public Administration Review*, Vol. 68/2, pp. 334-349. [29]
- White, R. (1959), “Motivation reconsidered: The concept of competence”, *Psychological Review*, Vol. 66/5, pp. 297-333, <http://dx.doi.org/10.1037/h0040934>. [250]
- Willems, T. and W. Van Dooren (2012), “Coming to Terms with Accountability: Combining multiple forums and functions”, *Public Management Review*, Vol. 14/7, pp. 1011-1036, <http://dx.doi.org/10.1080/14719037.2012.662446>. [157]
- Willems, T. and W. van Dooren (2011), “Lost in diffusion? How collaborative arrangements lead to an accountability paradox”, *International Review of Administrative Sciences*, Vol. 77/3, pp. 505-530, <http://dx.doi.org/10.1177/0020852311408648>. [6]
- Williamson, O. (1983), “Credible Commitments: Using Hostages to Support Exchange”, *The American Economic Review*, Vol. 73/4, pp. 519-540. [184]
- Wilson, J. (1989), *Bureaucracy: What Government Agencies Do and Why They Do It*, Basic Books, New York. [238]
- Yang, K. (2014), “Qualitative Analysis”, in Bovens, M., R. Goodin and T. Schillemans (eds.), *The Oxford Handbook of Public Accountability*, Oxford University Press, New York, <http://dx.doi.org/10.1093/oxfordhb/9780199641253.013.0040>. [293]
- Yang, K. (2012), “Further Understanding Accountability in Public Organizations: Actionable Knowledge and the Structure-Agency Duality”, *Administration and Society*, Vol. 44/3, pp. 255–284, <http://dx.doi.org/10.1177/0095399711417699>. [158]
- Yang, K. (2006), “Trust and citizen involvement decisions: Trust in citizens, trust in institutions, and propensity to trust”, *Administration and Society*, Vol. 38/5, pp. 573-595, <http://dx.doi.org/10.1177/0095399706292095>. [186]