

SEMINAR PAPER

Public service reform, scrutiny and inspection: Where to next?

The role of scrutiny and inspection in improving performance of public services¹

Melanie Ehren, University College London, Institute of Education
Christopher Chapman and Nick Watson, What Works Scotland

Introduction: The enigma of operationalising Christie

Scotland, like many other countries in the Global North, faces many challenges as it seeks to meet the needs of a changing population and changing economic conditions. In response to the failure of its public service to meet many of these challenges the Scottish Government commissioned the Christie Commission. It was charged to provide recommendations as to how public services could “change to meet the medium and long-term financial challenges and the expectations of the people of Scotland” (Scottish Parliament 2010). The emerging Scottish model to public service reform² is built around the four key pillars articulated by this Commission (2011, p.72):³

- **public services are built around people and communities**, their needs, aspirations, capacities and skills, and work to build up their autonomy and resilience;
- public service organisations **work together effectively to achieve outcomes** – specifically, by delivering integrated services which help to secure improvements in the quality of life, and the social and economic wellbeing, of the people and communities of Scotland;
- public service organisations **prioritise prevention, reduce inequalities and promote equality**;
- all public services constantly seek to **improve performance and reduce costs**, and are open, transparent and accountable.

One of its key recommendations is to give Audit Scotland “a stronger remit to improve performance and save money across all public service organisations” (2011;ix). Yet despite the centrality of inspection it is rarely mentioned in the final report. The Commission recognised that better scrutiny was required if services are to be improved and also that the call for improved partnership working and the resulting complex lines of accountability and governance can significantly impact on inspection. There is however

¹ This think piece draws on previously published work by Ehren, specifically:

Ehren, M.C.M., Janssens, F.J.G., Brown, M., McNamara, G., O'Hara, J. and Shevlin, P. (submitted). Evaluation and decentralised governance: examples of inspections in polycentric education systems

Janssens, F.J. and Ehren, M.C.M. (2016). Toward a model of school inspections in a polycentric system. *Evaluation and Program Planning*, 56, 88 - 98

Ehren, M.C.M. and Perryman, J. (accepted). Accountability of school networks: who is accountable to whom and for what?. *EMAL*

² Mitchell J (2014) The emerging Scottish model: Avoiding everything becoming nothing What Works Scotland.

<http://whatworksscotland.ac.uk/publications/the-emerging-scottish-model-avoiding-everything-becoming-nothing/>

³ Christie Commission, 2011. Report of the Future Delivery of Public Services, by the Commission chaired by Dr Campbell Christie. At: <http://www.gov.scot/resource/doc/352649/0118638.pdf>. Sourced 10 March 2016.

little or no discussion around how this is to be achieved or how these potential barriers may be broken down. Whilst the Report of the Christie Commission has set down the framework for public service reform there is as yet little evidence of the operationalisation of these principles and this remains 'a key challenge'.⁴

The purpose of this draft paper is to provide background context for the What Works Scotland Seminar on Public Service Scrutiny and Improvement in April 2017. The paper is structured around four key themes that we intend to explore during the seminar. The first theme relates to definitions and models, the second to methodology. The third theme focuses on values and judgements while the fourth discusses user involvement in scrutiny and inspection. In conclusion the paper offers some initial questions for further exploration at the seminar.

Theme one: Defining and describing models of public scrutiny and inspection

White (2015) defines scrutiny as 'any activity that involves examining (and being prepared to challenge) the expenditure, administration and policies of the government of the day', while Sandford (2005) refers to 'a set of clearly defined independent activities - regulation, audit, inspection and complaints handling - directed at public service providers such as local authorities, schools, care homes, housing associations, police services, prisons and hospitals'. Jain and Patnayakuni (2003) talk about public scrutiny as mechanisms of oversight and accountability with regard to an organization's actions. In Sutton and Galunic's view (1995), scrutiny means that a person is examined closely and methodically, the subject of a "minute inquiry". Public scrutiny, in their definition, refers to episodes where leaders and their organizations are examined in a close and obtrusive fashion.

Scrutiny can take place at any point – it may be forward looking (horizon-scanning to inform future policy), retrospective (e.g. evaluation implemented policy) or assess ongoing activity and ask questions about processes and outcomes. White (2015) explains how:

- *Scrutiny of process* asks the question: 'Did those in authority do what they were required to in reaching this decision?'
- and
- *Scrutiny of outcome* asks: 'Was the outcome what the Government intended?' 'Could that outcome have been achieved more effectively?' and 'Was that outcome the best possible one?'

Here we are particularly interested in the scrutiny of public services as described by Sandford (2005), where scrutiny is carried out by professional auditors such as inspection bodies and is focused on the (improvement) of public services by those who are audited (the auditee) to improve outcomes for those who use these services (users/clients). In Scotland, such *scrutiny activities can involve the collaboration of a number of inspection bodies who jointly evaluate the performance of a multiple of organisations to understand outcomes for the public that require the collaboration of public bodies*. An example is the place-based scrutiny of a local area undertaken by a consortium of scrutiny bodies, led by Education Scotland which looked at the performance of a range of public bodies, such as schools, care centres and local authorities from the perspective of the community they served.

Scrutiny over such outcomes is relatively straightforward when the network is established and has a formalized structure, governance and set of guiding principles to organize the collaboration between service providers in the network. In Scotland, public service

⁴ Mitchell J., 2014. The emerging Scottish model: Avoiding everything becoming nothing. What Works Scotland <http://whatworksscotland.ac.uk/publications/the-emerging-scottish-model-avoiding-everything-becoming-nothing/>

scrutiny is however tied to Government's public service reform agenda (eg, health and social care integration and the Community Empowerment Act) and Shared Risk Assessment by local councils. As a result, the scrutiny of public services will focus on more informal types of collaboration and multi-agency approaches where joint outcomes of organizations providing these services are the result of compromise and negotiation where partners try to work towards a magnitude of outcomes, and try to find the most effective way to share decision-making authority, risks, resources and benefits of the collaboration.

The potentially diffuse and ambiguous nature of such informal collaborative arrangements (at times, and increasingly across local government boundaries), the changing (contribution of members to) network-level outcomes, and the sometimes conflicting expectations of stakeholder and client groups makes it difficult to scrutinize the joint outcomes of such collaboration, and develop standardized frameworks and methodologies to guide the scrutiny work. The involvement of multiple inspection bodies, who each have their own line of accountability with the organisations they inspect, further complicates the development of standardized methodology, judgement and ways to involve users in public scrutiny. Such complications are clearly summarized by Mayne (2003), Schwartz (2003) and Ehren and Perryman (submitted) who talk about the problematic nature of accountability of networks in answering the question of 'who is accountable to whom and for what'.

Ehren and Perryman (submitted) and Ehren et al (submitted) provide a framework for answering this question, building on the work of Christie and Alkin (2013), Popp et al (2013), Aviram (2003), Gray et al (2003), Mandell and Keast (2007) and Provan and Kenis (2008). They juxtapose (monocentric) methods which evaluate single interventions, programmes, or organisational entities to (polycentric) models which include multiple levels of analysis (individual, interpersonal and collective) at which influence occurs. The common value of these polycentric approaches is their ability to understand and validate local and context-specific approaches to shaping educational quality by different partners in a network, looking at the bigger picture of how the many different parts in a network operate and the ways they interact and evolve over time in mutually reinforcing ways. Such approaches are particularly relevant for the public scrutiny in Scotland, given the involvement of multiple stakeholders (both auditors, inspectors and auditees and inspected) in the scrutiny exercise and the fact that the functioning and performance of mostly informal networks and sometimes diffuse and ambiguous collaborative arrangements is evaluated.

Table 1 summarises the two ends of the continuum while the remainder of this chapter further elaborates the differences in methodology, valuing/judging and use/user involvement in monocentric systems (hierarchical coordination of single organisations with single lines of accountability) versus polycentric systems (consisting of multiple centres of decision-making and network coordination)⁵.

⁵ The following section was adapted from: Ehren, M.C.M., Janssens, F.J.G., Brown, M., McNamara, G., O'Hara, J. and Shevlin, P. (submitted). Evaluation and decentralised governance: examples of inspections in polycentric education systems

Table 1. Public service scrutiny in monocentric versus polycentric systems (adapted from Ehren et al (submitted))

	Monocentric	Polycentric
Methodology Changes: from monocentric to polycentric: - Who defines standards and criteria and methods for inspection/public scrutiny? - What is the object of inspection/public scrutiny, who is evaluated?	Top down, 'objectivist' and standardized approach to evaluation. Focus on establishing causality, predicting and explaining quality, and an evaluation of single service providers.	Bottom-up, 'subjectivist' approach to evaluate performance of networks and collaborative efforts, aimed at validating, interpreting, understanding quality of context-specific approaches and solutions.
Valuing: - Who decides inspection criteria and targets for the scrutiny exercise?	Prescriptive assessment criteria to judge quality of individual service providers, pass-fail judgment decided by Inspectorate/scrutiny body	Inspectorate(s)/scrutiny bodies facilitate(s) evaluation, goal-free, flexible and specific to context and information needs of (network of) service providers and their users/clients.
User involvement: - What is the role of users/clients in the inspection/public scrutiny exercise and the use of findings? - Which phase of the inspection/scrutiny exercise are they involved in? - Who decides on consequences of an inspection assessment/outcome of the scrutiny exercise?	'Distanced evaluation approaches' Service providers and their clients/users are end users of inspection assessments/public scrutiny and treated only as the object of evaluation. Only primary 'decision-makers' are target of consequences (e.g. head teachers)	'Collaborative/participatory evaluation' Service providers and their clients/users are end users are involved in all inspection/scrutiny phases Intelligent intervention strategies targeted at all service providers and their clients/users to improve performance of collaborative work

We now move on to offer an overview of the second theme of methodology.

Theme two: Methodologies in scrutiny and inspection

Methodology involves the collection and analysis of empirical data for the study and judgment of particular aspects of social life. The methods of research provide the genesis of evaluation and the framework for evaluation practice, according to Christie and Alkin (2013). Using scientific research methods and techniques improves the rigor in evaluations and moves evaluation beyond what Suchman (1967) describes as 'evaluation as a common sense usage'. Methodology concerns the techniques used to conduct evaluation studies and these can range from the traditional research methods of (quasi) experimental research methods to evaluate the effects of an intervention or programme, to broader and more comprehensive conceptualisations of evaluation of human activity, policies or organizations.

Alkin (2013) provides a summary of the main (Northern American) theorists that have developed evaluation methodologies, such as Campbell, Suchman, Boruch, Cook,

Cronbach, Rossi, Weiss, Chen, Henry, Mark and Tyler. His overview of the main ideas of these evaluation theorists suggests a continuum of approaches from on the one hand those which are concerned with

- investigating causal inference and making generalisations to other subjects and settings,
 - evaluating single interventions, programmes, or organisational entities, and
 - using objectivist and standardized techniques in the evaluation, to those which aim to capture
 - the mechanisms and conditions that explain the functioning and performance of an intervention, programme or organisation. Such approaches would
 - include multiple levels of analysis (individual, interpersonal and collective) at which influence occurs
- and
- using constructivist approaches to develop and test theories of 'how something works'.

The first set of theories is particularly relevant for centrally organised inspections of single organizations, using standardized frameworks, aiming to produce evaluation findings that are reproducible and that would lead to similar assessments of school quality across different inspectors (Janssens and Dijkstra, 2013). Such objectivist and standardised approaches are informed by policy frameworks that describe performance indicators and quality of individual bodies (e.g. in a school context targets for school improvement) and define the remit of inspection bodies. Methods often rely to a large extent on available quantitative data, such as students' test scores or student drop out rates to evaluate a school's performance, sometimes using risk-based approaches which assume a cause and effect relation between risks of failing school quality (e.g. staff turnover, low teaching quality) and student performance.

The second set of theories is more relevant for Scotland's public service scrutiny as these theories suggest ways to understand the outcomes of interactions between public bodies, to look at interdependencies between different organisations and how these organisations have different roles and expertise in defining and improving school quality. The models summarized by Alkin (2013) provide examples of how public scrutiny bodies can adapt to local context, the outcomes of collaborative efforts they are inspecting and to create the conditions in which public bodies can effectively steer themselves and their joint work. Examples are 'theory-driven evaluation', 'comprehensive evaluation', 'programme evaluation', 'realist evaluation', 'modus operandi methods', 'principled discovery', 'objectives-oriented evaluation', 'decision-oriented evaluation', 'meta-evaluation', and 'system dynamics'. The common value of these approaches is their ability to understand and validate local and context-specific approaches to collaboration and service provision, looking at the bigger picture of how the many different parts in a network operate and the ways they interact and evolve over time in mutually reinforcing ways.

Theme three: Values and judgements in scrutiny and inspection

Valuing and judging distinguish evaluation from other forms of research as evaluators must place value on their findings and, in some cases, determine which outcomes to examine. According to Eisner (1998), evaluation concerns the making of value judgments about the quality of some object, situation or process (p.80). Valuing and judging are an important part of the inspection bodies whose work is often structured by a set of clear protocols and guidelines to judge the quality and performance of public bodies, categorizing the performance of these bodies on a predefined scale (ranging from failing

to good), and using a hierarchical model of aggregating judgements on lower level indicators to s a summary score on the overall quality and performance of the organisation, making 'pass/fail' decisions or using standardized evaluation criteria to compare similar entities or benchmarking organisations against a set of inspection indicators. These approaches are strongly in line with the objectivist methodologies described in the previous section, putting the onus of the judgement on the inspector(s) who is/are evaluating a public body

Such a standardized 'objectivist' approach however does not fit well in a polycentric system where a variety of different networks emerge which include a range of (sometimes different and changing) actors working on a variety of different network-level outcomes in response to context-specific problems.

Christie and Alkin (2013) describe more 'subjectivist' approaches to valuing which are more responsive to the object of evaluation and are guided by the meanings people construct in particular places, times and situations, recognizing the dynamic nature of human activity. Stakeholders and users of an evaluation (such as the network that is inspected) are actively involved in making judgements as the evaluator ensures that multiple realities are taken into account when making a value judgement. Specific approaches that have been described by a range of theorists (e.g. Scriven, Levin, Stake, Eisner, House, Greene, Guba, Lincoln, Mertens) include 'responsive evaluation', 'consumer reports', 'goal-free evaluations', 'deliberative democratic evaluation', 'value-engaged approaches to evaluation', and 'inclusive and transformative models of evaluation'.

In 'goal-free' evaluation (Scriven, 1997) the evaluator for example assumes the responsibility for determining which outcomes to examine, rejecting the objectives of organisations or specific programmes as a starting point (as cited in Christie and Alkin, 2013, p.33). This would allow the evaluator to identify the real accomplishments of organisations or programmes, using a more qualitative approach when trying to describe events, reactions and interactions. Goal-free evaluation would allow public scrutiny bodies to portray a picture of a situation when attempting to understand and make sense of what she/he has seen and decide on the value of what she/he has seen. Such goal-free evaluation also allows scrutiny bodies to be adaptable to stakeholder needs and concerns. In this approach the locus of judgement moves from the scrutiny body or individual inspector as the primary person to value an organisation, programme or service to facilitating the valuing by others, such as stakeholders and users and organizing the process by which those stakeholders and users are involved in making value judgments.

Theme four: User involvement in scrutiny and inspection

The third branch in Christie and Alkin's (2013) framework includes evaluation theories that focus on the use of evaluation findings, recognizing the importance of involving stakeholders when determining the evaluation questions to gather useful information. Alkin (2013) explains that a primary problem with evaluations is that the evaluator often does not consult the decision makers or any other interested parties when forming the research questions that will be addressed by the evaluation. Involvement of the general public in scrutiny is an issue often discussed in the public administration literature, particularly as something that is undervalued and often lacking. Involving users however allows scrutiny bodies to draw on multiple stakeholder perspectives and it promotes the possibility that relevant value perspectives are represented, fostering a comprehensive evaluation of programme and service value where evaluation outcomes are relevant and

timely. Involvement of a range of stakeholders also promotes the use of the findings from the scrutiny exercise.

Christie and Alkin (2013) describe how user involvement is strongly informed by decision-oriented theories that are concerned with designing evaluations that are intended to inform decision making and organizational change. This approach suggests that decision-makers (e.g. managers and policy-makers) are involved in evaluations. Other theorists such as House (1991), emphasize the involvement of minority or underrepresented groups whose voices are usually not heard, bringing empowerment and emancipatory goals into the evaluation. Stakeholders may vary from 'primary intended users' who have a stake in the evaluation and who personally care about the findings it generates (see Patton, 1997) to a broader spectrum of potential stakeholders who are not directly involved in the (improvement of the) programme or service being evaluated.

Stakeholders can be involved in different stages of an evaluation, such as the definition stage in which the goals, processes and resources of an evaluation are specified, the installation stage which aims to identify discrepancies in the implementation of the program, the process stage in which the extent of attainment of short-term outcomes or objectives are determined, and the product stage that aims to determine the attainment of terminal or ultimate goals. A potential fifth stage includes a cost-benefit analysis (see Provus, 1971).

The role of stakeholders in evaluations has consequences for the choice of methodology and for how judgements are made and informed. Guba and Lincoln (1989) for example describe how the claims, concerns and issues of stakeholders can be the organizing elements in the choice of methodology. In their view, the role of the evaluator becomes one that focuses on a joint construction of claims, concerns and issues to facilitate valuing by stakeholders. Stufflebeam (2003) also describes how evaluators can engage 'a representative stakeholder panel to help define the evaluation questions, shape evaluation plans, review draft reports and disseminate the findings' (as cited in Alkin, 2013, p.42). The evaluator has frequent interactions with this stakeholder panel to discuss the formative information produced from the evaluation, to make decisions about both the programme or service that is evaluated and the design of the evaluation to inform such decisions. Evaluations are shaped as a cyclical process of interactions between stakeholders and evaluators, and the role of the evaluator moves towards helping to develop the intervention, programme or service.

Involvement of stakeholders in inspection and public scrutiny however has major implications for the roles of Inspectorates and scrutiny bodies who generally position users of their findings and other stakeholder groups at the end of the evaluation cycle where they are expected to use findings as, and when presented to them, and to align their work to the inspection standards. Inspection and scrutiny bodies generally do not view the use of their findings as a two-way process, stimulated by the involvement of users in the entire cycle of the scrutiny exercise: taking an active part in the design of standards for the evaluation, the data collection and in deciding on improvements. The definition of stakeholders is also generally restricted to include primary users: staff working in inspected public bodies, instead of the broader community in which these organisations operate (e.g. including other service providers such as youth services, childcare or feeder schools).

Involvement users and a broader stakeholder group early on however allows for the public scrutiny exercise to become a process of joint learning among all participating agencies and organisations involved in the exercise through a process of collaborative evaluation and knowledge development; reciprocal relationships and joint activities are essential

strategies in such an approach which need to purposefully developed on the back of the public scrutiny exercise. The common concepts underlying these approaches are:

- a focus on the process of evaluation and a continuous cycle of evaluation with the purpose of transformation and learning (instead of seeing evaluation as an end product to be used for improvement by stakeholders)
- involvement of stakeholders throughout the evaluation process (instead of treating them as end users), and
- a shift in the role of evaluators from objective outsiders to one which fosters continuous interaction with the major stakeholders in an evaluation; evaluations should be actively involved in developing intended users' commitment to utilization and engaging intended users actively and directly in all stages of the evaluation.

Commentary

The four themes outlined above offer a starting point for thinking about the future of some aspects of public service scrutiny and inspection in Scotland. It is our intention that the presentations, responses and discussion at the forthcoming seminar will generate many more themes, questions and insights that will feed into this debate. We offer the following questions as a stimulus to initiate the conversation:

Theme one: Is polycentric scrutiny and inspection a helpful construct within the Scottish context? What forms of polycentric scrutiny would support the development of improved outcomes across a range of services? What might this look like in practice?

Theme two: What methodologies will provide rigorous quality assurance and generate quality improvement across professional and geographical boundaries?

Theme three: How do we combine objectivist and subjectivist approaches to optimise the impact of public service scrutiny and inspection?

Theme four: How can public service users best contribute to public service scrutiny and inspection? How does the system ensure this is meaningful and leads to changes in provision and practice?

While the above questions may, or may not provide a helpful starting point, we are clear that this debate is both necessary and timely. It will be for the system to assess the available evidence and to decide on the direction of travel about how we can best use scrutiny and inspection to improve outcomes across public services.