

The use of evidence in the audit, inspection and scrutiny of government

Introduction

In recent years audit, inspection and scrutiny (AIS) activity – by which we mean the work of national audit bodies, service inspectorates, parliamentary select committees and scrutiny committees in local authorities - has played an increasingly important role in holding central and local government, the health service and other public service providers in the UK to account.

Researchers have identified the escalation of ‘regulation inside government’ (Hood et al. 2000) with the emergence of the ‘audit society’ (Power 1997) and the UK government’s stated commitment to ‘evidence-based policy’ (Nutley et al. 2007). Policy makers have, it is suggested, become more dependent on what Hoggett (1996) calls the ‘long distant mechanics of control’ offered by AIS as responsibility for frontline service provision has become more dispersed and fragmented (Clarke et al 2000).

The growth of AIS activity has though brought new challenges. There has been widespread disquiet about the costs and other burdens which it places on regulated bodies (c.f. Crerar 2007). There have also been concerns that AIS bodies are being asked to take on new roles which could compromise their traditional independence and objectivity (Martin 2002). And there have been calls for much greater public involvement in AIS work (Broadbent 2003). Most of all though the challenge to AIS bodies is to demonstrate that the assessments they make are grounded in robust and reliable evidence.

To date there has been relatively little independent research into whether and if so how they actually achieve this. What sorts of evidence do they seek? What analytical routines characterise their work? What relationships do they build with the people and organisations they scrutinise? In short, how does the evidence which they gather inform the conclusions that they reach? In 2004/2005 the UK Centre for Evidence Based Policy and Practice explored some of these issues through a series of seminars entitled ‘Evidence for Accountability’. These highlighted five key issues which the convenors of the series believed merited further investigation (Levitt et al. 2005):

- Trust – How far does the authority of AIS work, and public and professional trust in AIS, depend on its evidential basis?
- Purposes – do the nature and uses of evidence differ between the accountability (so-called guard dog) role of AIS and the performance improvement (guide dog) role?
- Interests of users, participants and stakeholders – does evidence appropriate to each of these differ in its content, methods of generation and presentation?
- Methods and skills – are the AIS agencies too insular in their approaches to evidence? Can they learn more from each other?
- Challenges – how are the nature and uses of evidence responding to changes in AIS practice, especially the reduction of burden, the restructuring of some services and partnership working?

A study funded by the Nuffield Foundation over the last two years has sought to provide answers to some of these important questions. The research team, which comprises Professor Sandra Nutley, Dr Ruth Levitt, William Solesbury and Professor Steve Martin, has focused on four issues:

- What forms of evidence collection and evidence use characterise AIS activities in different settings?
- What impact do evidence-related practices in AIS have on the utility and legitimacy of that work?
- What are the key principles and practices for effective evidence use in different situations?
- What are the key similarities and differences between the use of evidence in AIS work and the use of evidence by government?

Our case studies

In order to shed light on these issues we worked with eight agencies which between them undertake a range of different types of AIS activity. We chose matched pairs of broadly similar kinds of activity: two service based inspections; two corporate local authority assessment frameworks; two value for money studies; and two scrutiny committee enquiries. And we included examples from England, Scotland and Wales in order to examine whether there were any national differences in the ways in which evidence was gathered and used by AIS bodies (Table 1).

Table 1: Case study organisations

	England	Scotland	Wales
Inspection of public services	Care Quality Commission	HM Inspector of Education	
Corporate assessments of local government	Audit Commission	Audit Scotland	
Audit and inspection of central government performance	National Audit Office		Wales Audit Office
Scrutiny of local government services	LB Harrow scrutiny committee		National Assembly for Wales

We focused on eight projects that were selected in consultation with these organisations from their ongoing portfolios of work (Table 2).

Table 2 Case study projects

AIS organisation	Case study project
Care Quality Commission	Health care associated infection (HCAI) inspection programme
HM Inspectorate of Education	Inspection of the impact of the implementation of the 'teacher's agreement'
Audit Commission	Comprehensive Area Assessment (CAA) Summer Trials
Audit Scotland	Best Value Audit (BVA) of a local council
National Audit Office	Supporting People with Autism through Adulthood
Wales Audit Office	Value for money study of Fleet Management
London Borough of Harrow Scrutiny Committee	Review of delivering a strengthened voluntary and community sector
National Assembly for Wales Sustainability Committee	Inquiry into carbon reduction from energy production

Among a wide range of other activities, the Care Quality Commission assesses whether (acute) NHS trusts' annually are complying with certain duties set out under the Hygiene Code.¹ Inspection is primarily through spot checks undertaken by experts who visit trusts unannounced. Evidence is also sought from policy documents and interviews. The aim is to assist trusts to improve practice, but the Commission also has powers to prosecute in the event of breaches of the code. The programme commenced in April 2008.

Her Majesty's Inspectorate for Education (HMIE) has been asked by the Scottish Government to review implementation of the teacher's agreement (known as TP21). This is being undertaken by a team of its inspectors. It involves the use of secondary data and the gathering of primary through focus groups, interviews and classroom observations. The exercise began in June 2008 and was expected to run for a year.

Comprehensive Area Assessment (CAA) is a joint assessment framework which involves seven inspectorates reaching joint judgements about service outcomes in a local authority area. It focuses on the effectiveness of local strategic partnerships (LSPs) which bring together councils, health, police and other key service providers. Reports flag up areas of weaknesses and examples of good practice which the inspectors believe may be of wider interest. The first reports are due to be published in November 2009. We studied trials of the methodology which were undertaken with a sample of authorities in the summer of 2008.

Best Value Audit (BVAs) are comprehensive assessments of the performance of Scottish councils. They are undertaken by a specialist team and are a selective and tailored audit based on an assessment of risk. Every council was audited in this way once between 2003 and 2009. The BVA we studied involved the use of self-assessment

¹ The Health and Social Care Act 2008. Code of Practice for the NHS on the prevention and control of healthcare associated infections and related guidance, Department of Health, January 2009.

and secondary data and the gathering of primary data primarily through surveys, focus groups, interviews. It ran from March 2008 to January 2009.

The National Audit Office's study of support for adults with autism examined the provision of services - including health and social care, education, benefits and employment support - in England. It was undertaken by value for money specialists within the NAO and used evidence gathered through surveys, focus groups, discussions and interviews with service providers and their clients and cost modelling of alternative provisions. It ran from January 2008 to April 2009.

The Wales Audit Office is undertaking a four year programme of studies of fleet management, buildings management, plant, machinery and equipment and IT assets, areas which are seen as having considerable potential to produce efficiency savings. The work is being undertaken by a central team with input from local audit teams, and is scheduled for 2008 to 2011. Our case study focused on the analysis of fleet management study which was completed in the first half of 2008.

The London Borough of Harrow's scrutiny focused on its relations with the local community and voluntary sector, including issues of collaboration, funding, and the use of community assets and premises. It also addressed the involvement of the police and the Primary Care Trust. It was undertaken by a group of councillors and co-optees from the community and voluntary sector with staff support who gathered evidence from written and oral evidence. It ran from March to December 2008.

The National Assembly for Wales (NAfW) Sustainability Committee is undertaking a programme of work which is scrutinising the Welsh Assembly Government's policies for meeting carbon reduction targets. Evidence is gathered through written evidence, site visits, briefing papers and hearings. The case study was concerned with carbon reduction from electricity generation (including renewable energy). The enquiry started in May 2008 and the report was published in May 2009.

Methods

We have analysed the design and operation of the eight projects through a combination of document analysis, non-participant observation at public and closed meetings and hearings, and interviews with the key players, and we have tested out our emerging findings through bi-lateral meetings with the eight AIS organisations. We have also held a workshop involving all eight organisations to explore the overall conclusions and implications of our work.

In order to make systematic comparisons between case study projects we have developed a common descriptive framework. This has provided a structure both for assembling a systematic account of each project's use of evidence and comparing between projects in order to identify similarities and differences. It covers eight main issues:

- The genesis of the project;
- Its purpose and focus;
- Management of the activity;

- The forms of evidence gathering;
- The processes of evidence use;
- The assumptions which underpinned the project;
- How findings were reported; and
- The impacts and in particular benefits of the exercise.

We have been particularly interested in the way in which the use of evidence has been influenced by factors such as the prior knowledge and assumptions of those undertaking the AIS activity; the purpose of the exercise; and the relationships between stakeholders. And we have also examined the ways in which evidence influenced the judgements reached and proportionality i.e. how AIS decided at what point they had amassed sufficient evidence.

Findings

Not surprisingly given the diversity of the case studies, the nature of evidence gathered and used varied considerably. However, there are a number of overall findings which have emerged so far from the AIS activities that we have studied.

What evidence is sought and what counts as good evidence?

Most of the AIS work that we observed involved fairly conventional forms of evidence and evidence gathering. Common methods included the analysis of documents, written and verbal evidence, site visits, enquiry visits, interviews and the use of secondary performance data.

Some case studies (including the work of the NAO and the NafW's sustainability committee) had well developed traditions of evidence gathering. Others (notably the CAA trials) were developing and/or refining their methods. However, most AIS bodies had to balance the desire to minimise the burden that AIS placed on both those being investigated and those providing information with the need to ensure that findings were seen as legitimate. Legitimacy often involved two key considerations. First, that evidence gathering had been inclusive in the sense of taking account of different stakeholders' perspectives. Second, that it was robust and could withstand challenge.

The aim of the AIS work, the status of the organisation undertaking it, and the resources available to it all influenced the approaches that were taken to evidence gathering, as did the organisations' own traditions in terms of their *modus operandi*. Some took a predominantly deductive approach which involved starting with some propositions or hypotheses and gathering evidence to test, refine and possibly refute these starting points. Others used evidence to compare organisations against explicit criteria or standards. A third group took an inductive process which involved more open ended sifting of evidence in search of answers. But even in these cases the kind of evidence that was gathered was shaped by implicit assumptions about and/or prior knowledge of the subject area.

We found significant variations in the approach to data handling across the eight case studies. Most of the AIS bodies experienced significant problems inspecting or judging outcomes. The lack of counterfactual data and the likelihood of time lags made simple

cause and effect relationships unlikely and attribution difficult. There were also big questions about 'whose outcomes' should be taken into account. In some cases for example local service providers had different priorities to those of central government. For all these reasons it was often easier to focus on processes and/or organisational attributes that were assumed to be linked to desired outcomes rather than upon the outcomes themselves.

The sharing of data and evidence with other bodies is becoming an increasingly important issue for several of the AIS bodies that we worked with, and in some cases there was simply too much evidence to handle. In theory developments in electronic information and communications systems should make sharing/pooling of data much easier. In practice though there are often problems with evidence formats, time frames and protocols.

The use of self-assessment also varied across the case studies. In some cases it was used by inspectors as a way of gauging 'self awareness', rather than as part of a process of co-production in which self assessment and external assessment were seen as having parity of esteem.

How does the evidence gathered inform the conclusions and judgements that are reached?

AIS findings were related to the evidence that was gathered, but the relationship between evidence and findings was not simple or straightforward. Findings were always mediated by interpretation. Evidence and judgements often evolved in parallel with the interplay between them shaped by a range of different factors including professional reputation of the AIS body and the organisations that were being scrutinised; the evidence gathering methods; the experience of the AIS bodies; and tacit knowledge and intuition. Political judgements were also often very important, shaping the nature and impacts of AIS work - before, during and after the projects were undertaken.

Evidence was used in different ways in different case studies. In some instances it was used to arrive at a set of findings. In others it was used to verify propositions that had been arrived at quite early on in the process. In some it acted as a trigger of intervention in failing organisations. In others it acted as a means of highlighting good practice. In one it could be used to prosecute organisations that were failing to meet minimum standards.

Moderation and clearance processes were in different ways built into all of the case studies. They were often very important and could change not just the tone of reports but also the judgements that were made, sometimes without consideration of any significant new evidence.

What impact do evidence-related practices in AIS have on the impact and legitimacy of that work?

We found that the perceived legitimacy of AIS reports depended in part but not exclusively on the evidence base. It was also influenced by professional expertise of

those undertaking AIS work, the composition of AIS project team or committee, adherence to expert frameworks, and co-production of findings.

The mechanisms for checking on the implementation of recommendations and for evaluating impacts vary. In several cases follow up was built into the process, but on the whole follow-up processes after publication of the report are not yet well developed.

Most of the AIS bodies saw their work as being conducted on behalf of the public. However, interestingly, there was often minimal direct public involvement. Evidence from citizens and service users did not loom large in many of the case studies, although some pro-actively sought views from the voluntary and community sectors. Nor were members of the public usually involved in judgement processes. The mechanisms for informing the public about findings varied considerably. Some AIS bodies went to considerable lengths to broadcast or publicise evidence gathering sessions and/or reports. In general there was however relatively little public interest in proceedings, although in some cases user groups were active in responding to or commenting on reports. There was a widespread view that AIS bodies needed to guard against a divide opening up between professionals who do the investigations and those who are recipients of them. Evidence gathering and analysis can become too sophisticated for non-specialists to be able to deal with, and it is important that reports are accessible. This requires attention to issues such as presentation and formatting, though it is important not to 'dumb down' reports.

The impact of the case study projects was difficult to assess, partly because many have only just been published and some have still to be published. Some exercises appeared likely to have a direct impact (feeding into policy and/or practice in an instrumental way), but more often it seemed that any influence would be indirect, for example helping to shape subsequent policy debates and ways of thinking. However, there was a general feeling that the economic downturn is likely to lead to increasing attention on the impacts and value added by AIS activities. Several of the bodies which we worked with believed that the development of robust approaches by which to evaluate the effectiveness of AIS was a priority for future research.

Next steps

We are currently writing up a report for the Nuffield Foundation, which funded the research, and disseminating key findings through a number of workshops and events.

For further information please contact the project director Professor Sandra Nutley (telephone: 0131-650-4307/8342; email: Sandra.Nutley@ed.ac.uk).

References

- Broadbent J (2003), 'The crock of gold at the end of the performance rainbow?', *Public Money & Management*, 23:1, pp 5-7.
- Clarke, J., Gewirtz, S., Hughes, G. and Humphrey, J. (2000) 'Guarding the public interest? Auditing public services', in Clarke, J., Gerwitz, S. and McLaughlin, E. (eds), *New Managerialism, New Welfare?* London: Sage

For publication in the Evaluator (UK Evaluation Society) - Summer 2009 edition

Hoggett, P. (1996) 'New modes of control in the public service' *Public Administration* 74, 9-32.

Hood, C., James. O. and Scott. C. (2000) 'Regulation inside Government: has it increased, it is increasing, should it be diminished?' *Public Administration* 78:2, 283-304.

Levitt, R., Solesbury, W., and Steele J. (2005) *Evidence for accountability: Final report of seminar series*. Centre for Evidence Based Policy and Practice, Queen Mary, University of London.

Martin, S. J. (2002) 'The modernisation of UK local government: markets, managers, monitors and mixed fortunes', *Public Management Review*, 4:3, 291-307

Nutley, S.M., Walter, I. and Davies, H.T.O. (2007) *Using Evidence: How research can inform public services*. Bristol: The Policy Press.

Power, M. (2003) 'Evaluating the Audit Explosion' *Law and Policy* 25:3, 185-202.