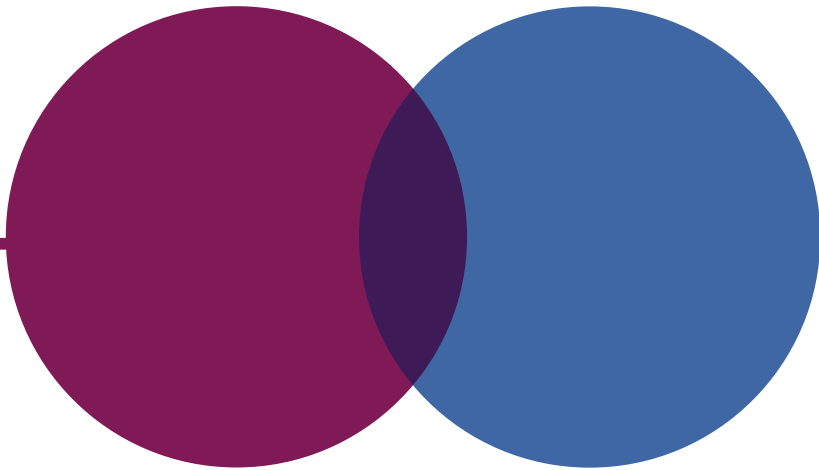




National Audit Office



Evaluating government spending

HM Treasury, Cabinet Office

REPORT

**by the Comptroller
and Auditor General**

**SESSION 2021-22
2 DECEMBER 2021
HC 860**

Key facts

8%

of government spend on major projects (£35 billion of £432 billion total expenditure) had robust evaluation plans in place in 2019

6

departments, out of the 16 that we surveyed, had a single evaluation strategy covering their whole department

4

departments, out of 16, had a readily available estimate of spending on evaluation activity

7 of 16

chief analysts thought a barrier to the use of evaluation evidence in their department was the lack of pressure and demand from senior policy colleagues in support of evaluation evidence

10 of 16

chief analysts told us another barrier was that the opportunity to learn was not adequately built into policy design and delivery

6 of 16

chief analysts told us that only in some or a limited number of cases were they able to publish evaluation findings in a timely manner

9

What Works Centres, which produce evidence about the most effective practices across a number of policy areas

£84 million

our lower estimate of the value of external contracts issued in 2019-20 by the core departments to conduct or support evaluation activity

£885 billion

government spending (Total Managed Expenditure) in 2019-20

Summary

Introduction

1 Evaluation is a systematic assessment of the design, implementation and outcomes of an intervention. Central government guidance makes it clear that departments are expected to undertake comprehensive, robust and proportionate evaluations of their interventions. It is one of many types of evidence that can inform decision-making. Evaluation evidence can help governments understand which approaches work best and support accountability for decisions. Using evaluation evidence requires effective coordination between analysts, decision-makers and officials responsible for developing and implementing policy.

2 Our 2013 report, *Evaluation in government*, concluded that while government spends significant resources on evaluation, coverage of evaluation evidence was incomplete, and the rationale for what government evaluates was unclear. We also found that evaluations were often not robust enough to attribute the impact to the policy being evaluated, and that government did not effectively use the learning from these evaluations to improve impact and cost-effectiveness.

3 Several public bodies, networks and professions are involved in evaluation across government. Within departments, ministers and accounting officers are accountable to Parliament for evaluation of the outputs and outcomes of policies and programmes. Departmental chief analysts and heads of policy also have evaluation responsibilities. HM Treasury publishes guidance on evaluation requirements and expectations for departments. Networks within government, including the Government Analysis Function, the Policy Profession and the Cross Government Evaluation Group (CGEG), also play key roles. In April 2021, the Cabinet Office and HM Treasury established a new Evaluation Task Force to “deliver a step-change in the scale, quality and impact of evaluation practice in government”.

Study scope

4 This report examines government’s progress in developing the provision and use of evaluation evidence across government. Our primary focus for this report is on the role of HM Treasury, Cabinet Office and the Analysis Function in setting out requirements, incentives and oversight arrangements to support accounting officers in fulfilling their evaluation responsibilities. We consider: what actions the government has taken since our 2013 report recommendations, including actions to adopt a strategic approach to evaluation; and progress in addressing systemic barriers to good evaluation and good use of evaluation evidence.

5 Our report considers the provision and use of evaluation evidence in policy-making in the core government departments. It does not cover the wider evaluation eco-system, such as What Works Centres that help departments understand what works in their respective policy areas or external organisations commissioned by departments to undertake evaluations. We did not assess the quality of individual evaluations or how departments have used evidence from evaluations in individual cases. Our report does not consider how policy decision-makers draw on other types of evidence.

6 Our methods are set out in Appendix Two.

Key findings

7 Despite government’s commitment to evidence-based decision-making, much government activity is either not evaluated robustly or not evaluated at all. In December 2019, the Prime Minister’s Implementation Unit (PMIU) concluded that government has little information in most policy areas on what difference is made by the billions of pounds being spent. Out of government’s 108 most complex and strategically significant projects in its Government Major Projects Portfolio, only nine – representing 8% of £432 billion in spending – are evaluated robustly, while 77 (64% of spend) have no evaluation arrangements. Government does not hold data on how far ‘business as usual’ activities are covered by evaluation.¹ Our past reports show many examples of evaluation not being carried out, as well as weaknesses in evaluations or the way evaluation evidence had been used. Approaches to evaluation and evaluation quality vary significantly both between and within departments (paragraphs 1.8 to 1.10, and Figure 2).

Actions to strengthen strategic approaches to evaluation

8 Government has taken steps to strengthen evaluation since our 2013 report. In 2015, it established the cross-government What Works Trials Advice panel and, in 2017, the Analysis Function, whose role is to lead the analytical community, improve analytical capability and share best practice, including in relation to evaluation. Individual departments have undertaken initiatives to improve evaluation. In 2020, government published an update to the Magenta Book – the central government guidance on evaluation. During the 2020 Spending Review, HM Treasury linked funding decisions more clearly to assessments of evaluation evidence quality. In 2021, Cabinet Office and HM Treasury created the Evaluation Task Force. Stakeholders we interviewed welcomed the greater focus on evaluation (paragraphs 2.4 to 2.10, 2.20 and Figure 3).

¹ ‘Business as usual’ activities of government in this context are the normal, expected operations of government in contrast to any projects and interventions associated with change.

9 Despite these greater efforts, roles and responsibilities at the centre of government remain unclear.² The complexity of the evaluation landscape, and the previous lack of a strategic approach to evaluation, has resulted in overlapping roles and responsibilities. For example, HM Treasury has not set out clear arrangements for maintaining and promoting the Magenta Book, including commissioning of future updates. It was updated by the CGEG as a cross-departmental group of evaluation practitioners in 2011 and 2020. There are no systematic arrangements at the centre of government for following up on whether departments are complying with requirements on evaluation. The lack of a coherent central strategy on evaluation in the past is reflected in the variation, in coverage and quality, across departments (paragraphs 2.5 and 2.11 to 2.13).

10 The government said it would review the future of the Evaluation Task Force after two to three years. Government set up the Evaluation Task Force with the purpose to improve how government evaluates programmes and to inform decisions on whether it ought to stop, continue, expand or modify them. Government said that it will decide on the merits of continued funding of the Task Force, based on an assessment of its performance against its objectives. It has not set out the criteria it will use to assess whether the Evaluation Task Force has delivered its purpose (paragraphs 2.14 to 2.16).

11 There is limited oversight from the centre of government to ensure that departments carry out the required evaluations and improve their practice. HM Treasury made greater use of its powers to set requirements for evaluations when it approved funding at the 2020 Spending Review, and intended a similar approach at the 2021 Spending Review. However, departments told us that HM Treasury has not put in place formal arrangements to follow up whether they are complying with conditions it set as part of the 2020 financial settlements. Other than at Spending Reviews, there is little oversight or action to drive improvements in areas where evaluation arrangements are insufficiently robust (paragraphs 2.5 to 2.9, 3.32 and 3.33).

12 Few departments have an evaluation strategy that spans their whole department. Having a strategy can help ensure that the most pressing evaluation gaps are prioritised, planned and resourced. Of the 16 core government departments we surveyed, just over one-third (six) said they had a strategy covering their whole department. A further seven had strategies but only in specific policy areas. Three departments had no evaluation strategy at all. Departments set out their objectives and priority outcomes in Outcome Delivery Plans. However, these Plans do not provide information on how departments prioritise areas for evaluation and how risks or importance of those areas are aligned to priority outcomes (paragraph 2.19).

² In this report, we use the term centre of government to refer to the Cabinet Office and HM Treasury (including the Evaluation Task Force), and the senior leadership of the Analysis Function.

13 Government does not know the full range of evaluation activity, how much it spends or the number of people working on it. Most departments do not collect and hold information on evaluation spending, which tends to be spread across multiple internal budgets and therefore lacks visibility. Of the 16 departments, 12 were unable to provide a readily available estimate of spending on external evaluations, and 11 were unable to provide a readily available estimate of the number of staff working on evaluation. We used publicly available spending data on externally commissioned work to estimate that the 16 core government departments contracted externally for evaluation work worth at least £84 million and £67 million in cash terms in 2019-20 and 2020-21 respectively. Without understanding spending and activity, departments risk not being able to track evaluations, ensure their quality, identify evaluation gaps in their activities or demonstrate that their evaluation activity represents value for money (paragraph 2.21).

Understanding the barriers

14 Barriers to good evaluation and use of evaluation evidence have persisted since our previous report. In 2019, the PMIU produced analysis identifying similar factors to those we had noted in 2013, including lack of political engagement with evaluations, capacity concerns and a lack of incentives for departments to produce and use evaluation evidence, together with few adverse consequences for not doing so. Our surveys of departments for this 2021 study found general agreement that these barriers still apply (paragraphs 3.3 to 3.6 and Figure 6).

15 Chief analysts and heads of policy profession differ on which are the most common barriers to using evaluation evidence. Chief analysts most commonly identified insufficient understanding of evaluation evidence by policy-makers, and lack of demand from senior policy colleagues in their departments as the major barriers.³ Heads of policy profession most commonly mentioned evaluation evidence not being available when needed and the insufficient capacity of analysts to help them understand the evidence.⁴ Effective collaboration and coordination between the various communities will be important for ensuring there are the right strategies in place to address these barriers (paragraphs 3.7 to 3.9 and Figure 6).

³ Chief analysts are responsible for overseeing research and analysis, including evaluation, across departments' policy areas.

⁴ Heads of policy profession are responsible for monitoring and improving policy capability in their departments.

Government actions to address barriers

16 Building evaluation into policy design and delivery remains challenging.

Ten chief analysts and eight heads of policy profession in the 16 departments we surveyed, identified that the opportunity to learn was not being built into policy design and delivery. We found examples where departments have tried to combine evaluation and policy design more effectively. These included: bringing analysts and policy officials together within programme or project teams; initiatives to raise awareness and skills among policy officials; and formal processes, such as the scrutiny of evaluation plans at the investment approval stage, which require policy officials to consult with analysts at specified points in policy development and implementation (paragraphs 3.6, 3.21 and 3.22 and Figure 9).

17 Government is taking steps to improve the skills of analysts and policy officials, although it recognises that more needs to be done.

Departmental chief analysts told us they face challenges in recruiting and retaining skilled evaluators and senior analytical capability. HM Treasury's updated version of the Magenta Book published in March 2020 was accompanied by a framework setting out the knowledge and skills needed by analysts to deliver quality evaluations. The Analysis Function Career Framework sets out the skills and experience needed for analytical roles across government – including those involved in evaluation. A challenge is ensuring that policy officials have sufficient 'evaluation literacy' to understand the evidence before making decisions. The Analysis Function told us it has started an audit of analytical skills among policy officials and expects to conclude its work, with planned actions, in March 2022 (paragraphs 3.23 to 3.25).

18 Chief analysts have mixed views on the quality of support that the centre of government gives them on evaluation.

Departments that received support from the centre were most satisfied with access to advice on evaluation design, integrating evaluation plans into policy, and recruiting evaluation specialists. They were least satisfied with support for working across departments on evaluating shared outcomes and sharing data. As part of a wider initiative, government is developing an Integrated Data Service to make sharing data easier across government. In some cases, chief analysts were not aware of support that the centre of government provides (paragraphs 3.17 to 3.19, 3.28 and Figure 8).

19 Poor understanding of the value of evaluation at senior levels is still a challenge.

The 2019 PMIU review found evidence that it was hard to embed a culture of open enquiry and overcome the temptation to use evaluation to justify chosen policies. The Cabinet Office is looking at cultural norms across the civil service and plans to draw lessons from this to inform its actions on improving evaluation. Seven out of 16 chief analysts thought that the lack of pressure and demand from senior policy colleagues in support of evaluation evidence was a barrier to its use (paragraphs 3.10 and 3.11, and Figure 6).

20 Departments are falling short of government requirements on transparency and publication of evaluation findings. Government’s guidance is that “the presumption should be for maximum openness and transparency to allow others to critique the methods used, as well as learn from and replicate them. Publishing the communications plan, so external observers are aware of what will be published when, is also good practice.” We heard that departments could find it difficult to get approval from senior civil servants and Cabinet Office to publish evaluations and protocols. More than one-third of chief analysts (six out of 16) told us that they could publish evaluation findings in a timely manner only in some or a limited number of cases. The Evaluation Task Force told us it is planning to improve the transparency of the evaluations that are commissioned and published, by collating the information in a public register (paragraphs 3.12 to 3.14 and Figure 4).

Achieving change through evaluation

21 Government is not managing the knowledge gained from evaluation effectively or using it widely. Of the 16 core departments, the large majority of chief analysts (15) and heads of policy profession (13) agreed that the primary purpose of evaluation is to learn what works. There are, however, difficulties in accessing and understanding that knowledge. In a limited number of policy areas, What Works Centres collate existing evidence on the effectiveness of programmes, and produce synthesis reports and systematic reviews. But, overall, government is not taking full advantage of opportunities to bring together evaluation findings, extract the learning and apply lessons across different departments (paragraphs 3.29 and 3.30).

22 Government does not capture and publicise how evaluations lead to improved outcomes. We identified examples of departments changing interventions based on evidence from evaluations. At present, however, the details of examples like these are not available outside the relevant department itself (paragraphs 1.12 and 3.31).

Conclusion on value for money

23 While individual departments have undertaken initiatives to improve evaluation, the use of evaluation continues to be variable and inconsistent, and government has been slow to address the known barriers to improvement. As a result government cannot have confidence its spending in many policy areas is making a difference. Government has recently committed to improve evaluations, included requirements relating to evaluation in some spending decisions, and strengthened capacity through the creation of the Analysis Function and a central Evaluation Task Force. These interventions will take time to mature. Nevertheless, this renewed focus on evaluation is a welcome step to using evidence better and improving value for money.

24 Government needs to clarify responsibilities, oversight and communication of evaluation evidence. Building on the reforms it has made, and the efforts of individual departments, government will have to do more to address the systemic barriers to effective evaluation and the application of evaluation evidence to policy-making. Otherwise it will not be able to ensure evaluations drive improved outcomes.

Recommendations

25 Ministers and accounting officers are accountable to Parliament for evaluation of departmental activities, with chief analysts and heads of policy playing key support roles. Our recommendations are directed toward the centre of government and aimed at securing long-lasting improvements for all government departments, building on progress and momentum to date.

26 To improve the way that the separate parts of the evaluation system work, individually and collectively, HM Treasury and Cabinet Office should:

- a** work with the Evaluation Task Force to publish the roles and responsibilities of government bodies with respect to evaluation. This should include but not be limited to HM Treasury, the Evaluation Task Force, the Analysis Function and the CGEG; and
- b** publish a plan for improvements to the evaluation system, including the outcomes they want to see and how they will achieve and measure them, including clear criteria for assessing whether the Evaluation Task Force is achieving its purpose.

27 The Analysis Function should:

- c** set out the appropriate governance structure for the ownership, maintenance, assurance and monitoring of evaluation standards as presented in its Analysis Functional Standard. It should agree with HM Treasury the funding and capacity implications for this governance structure; and
- d** work with Cabinet Office to develop an appropriate assessment framework, which will provide the Analysis Function with the necessary levers to monitor and support departments' implementation of the Analysis Functional Standard.

28 To promote transparency and strengthen incentives across government, HM Treasury should:

- e** write to departments asking them to publish an evaluation strategy covering their key evaluation evidence gaps, planned evaluations, lessons from recent evaluations and details of planned evaluation spend and staff resources. This should form part of future spending review settlement conditions and be updated in line with departments' Outcome Delivery Plans and no less frequently than every three years;

- f** work with Cabinet Office to reinforce the expectation of an 'open by default' transparency commitment relating to publication of evaluations when policies are planned, by recording departments' explanations of why they have not published evaluations; and
 - g** work with the Evaluation Task Force on a robust and documented system to follow up cases where programme funding is conditional on the department performing evaluation activities and intervene if departments fail to do so.
- 29** To raise standards and support departments in consistently meeting evaluation requirements, the Analysis Function and the Evaluation Task Force should work with others in the evaluation community of practice (including CGEG and government professions) to make available in a single place:
- h** good practice, toolkits and operational guidance including, for example:
 - on how evaluation approaches can be embedded into existing departmental information and processes including risk management arrangements, to identify evidence gaps and make use of evaluation findings;
 - ways to strengthen integration of evaluation and policy design;
 - practical examples of how agile evaluation approaches have been embedded within policy delivery; and
 - to support access to thematic knowledge of what is working, why and lessons learned from evaluation findings across government; and
 - i** information on which interventions are continued, changed or stopped as a result of evaluations, to demonstrate the practical impact of good evaluation evidence on decision-making and help inform assessments of whether the evaluation system is working as intended.
- 30** To deliver a step-change in the evaluation capacity and capability of analysts and policy staff, the Analysis Function should:
- j** work with the CGEG, departments and the Cabinet Office to assess government's specialist evaluation capacity and capability and agree a plan to address identified shortfalls; and
 - k** work with the Policy Profession to deliver plans to assess and improve evaluation literacy for policy professionals and analysts across government.