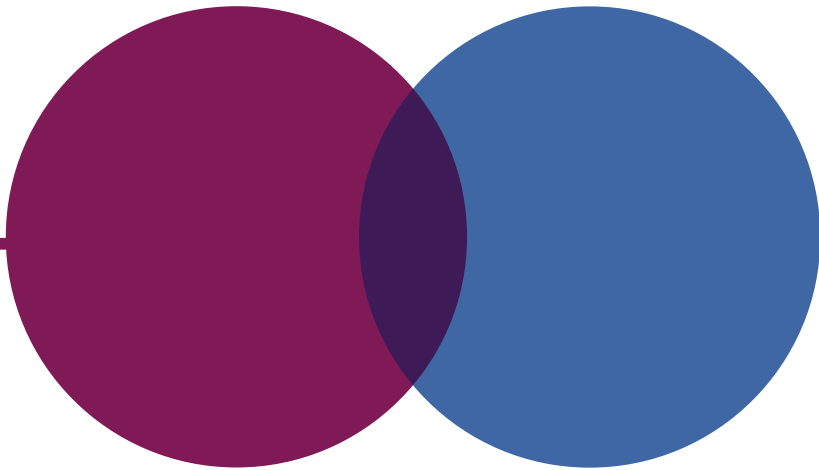




National Audit Office



# Evaluating government spending

HM Treasury, Cabinet Office


---

**REPORT**

**by the Comptroller  
and Auditor General**

---

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



We are the UK's independent public spending watchdog.

We support Parliament in holding government to account and we help improve public services through our high-quality audits.

The National Audit Office (NAO) scrutinises public spending for Parliament and is independent of government and the civil service. We help Parliament hold government to account and we use our insights to help people who manage and govern public bodies improve public services.

The Comptroller and Auditor General (C&AG), Gareth Davies, is an Officer of the House of Commons and leads the NAO. We audit the financial accounts of departments and other public bodies. We also examine and report on the value for money of how public money has been spent.

In 2020, the NAO's work led to a positive financial impact through reduced costs, improved service delivery, or other benefits to citizens, of £926 million.



National Audit Office

# Evaluating government spending

**HM Treasury, Cabinet Office**

---

## **Report by the Comptroller and Auditor General**

Ordered by the House of Commons  
to be printed on 30 November 2021

This report has been prepared under Section 6 of the  
National Audit Act 1983 for presentation to the House of  
Commons in accordance with Section 9 of the Act

---

**Gareth Davies**  
**Comptroller and Auditor General**  
**National Audit Office**

**29 November 2021**

## Value for money reports

**Our value for money reports examine government expenditure in order to form a judgement on whether value for money has been achieved. We also make recommendations to public bodies on how to improve public services.**

The material featured in this document is subject to National Audit Office (NAO) copyright. The material may be copied or reproduced for non-commercial purposes only, namely reproduction for research, private study or for limited internal circulation within an organisation for the purpose of review.

Copying for non-commercial purposes is subject to the material being accompanied by a sufficient acknowledgement, reproduced accurately, and not being used in a misleading context. To reproduce NAO copyright material for any other use, you must contact [copyright@nao.org.uk](mailto:copyright@nao.org.uk). Please tell us who you are, the organisation you represent (if any) and how and why you wish to use our material. Please include your full contact details: name, address, telephone number and email.

Please note that the material featured in this document may not be reproduced for commercial gain without the NAO's express and direct permission and that the NAO reserves its right to pursue copyright infringement proceedings against individuals or companies who reproduce material for commercial gain without our permission.

Links to external websites were valid at the time of publication of this report. The National Audit Office is not responsible for the future validity of the links.



# Contents

**Key facts** 4

**Summary** 5

## **Part One**

Background 13

## **Part Two**

Developing a more strategic approach to evaluation and use of evidence 19

## **Part Three**

Actions to strengthen provision and use of evaluation in practice 27

## **Appendix One**

Our audit approach 40

## **Appendix Two**

Our evidence base 42

This report can be found on the National Audit Office website at [www.nao.org.uk](http://www.nao.org.uk)


If you need a version of this report in an alternative format for accessibility reasons, or any of the figures in a different format, contact the NAO at [enquiries@nao.org.uk](mailto:enquiries@nao.org.uk)


The National Audit Office study team consisted of:


Vyara Apostolova, Simon Banner, Phil Bradburn, Araz Enayati Rad, Adam Halford, Elisabeth Moore, Eleanor Robinson, Ronan Smyth, and Jeremy Weingard, under the direction of Ruth Kelly.

For further information about the National Audit Office please contact:

National Audit Office  
Press Office  
157-197 Buckingham Palace Road  
Victoria  
London  
SW1W 9SP

 020 7798 7400

 [www.nao.org.uk](http://www.nao.org.uk)

 @NAOorguk

## 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.<sup>1</sup> 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).

<sup>1</sup> ‘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.**<sup>2</sup> 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).

<sup>2</sup> 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.<sup>3</sup> 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.<sup>4</sup> 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).

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

<sup>4</sup> 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.

# Part One

## Background

**1.1** Evaluation is important for learning whether government interventions are working and to demonstrate accountability for the use of public money.<sup>5</sup> Policy makers can learn from evaluations to help decide whether interventions should be continued, expanded, improved or stopped altogether. Using evaluation evidence in policy design can enable decision-makers to better target their interventions and maximise the chances of achieving desired objectives. From an accountability perspective, evaluations allow Parliament, scrutiny bodies and the public to examine how effectively public money has been spent.

**1.2** Evaluation is a systematic assessment of the design, implementation and outcomes of an intervention. Its purpose is to provide insights into how an intervention has been implemented and what effect it had, for whom, how and why. There are three different types of evaluation: process, impact and economic (also known as value-for-money evaluation):

- Process evaluations examine activities and implementation.
- Impact evaluations focus on the impact of an intervention and identify the change in outcomes directly attributable to an intervention, as well as its impact on different groups.
- Economic evaluations compare the benefits and costs of an intervention and assess whether an intervention was a good use of resources.

**1.3** Evidence-based decision-making is an ambition of the UK government.<sup>6</sup> While evaluation is crucial to understanding the effectiveness and efficiency of interventions, it is not the only type of evidence used to inform decisions. For example, performance monitoring provides real-time insights into the delivery of an intervention. It is different from evaluation, primarily because it does not assess the implementation, impact or value of an intervention.

5 HM Treasury, *Magenta Book: Central Government guidance on evaluation*, March 2020.

6 *Declaration on Government Reform* available at: [www.gov.uk/government/publications/declaration-on-government-reform](http://www.gov.uk/government/publications/declaration-on-government-reform); the Ditchley Annual Lecture on "*The privilege of public service*" available at: [www.gov.uk/government/speeches/the-privilege-of-public-service-given-as-the-ditchley-annual-lecture](http://www.gov.uk/government/speeches/the-privilege-of-public-service-given-as-the-ditchley-annual-lecture)

**1.4** In the March 2020 Budget, government announced plans for using the 2020 Spending Review to incentivise departments to improve evaluations of their work. In June 2021, government published its Declaration on Government Reform, in which it expressed its vision for more rigorous evaluation of policy and decisions. The declaration also emphasises the importance of transparency around the effectiveness of government interventions.

### **The evaluation landscape**

**1.5** Other countries also recognise the importance of evaluation. In 2020, the Organisation for Economic Co-operation and Development (OECD) reported that evaluation is incorporated into the budget cycle of half of the 42 countries it surveyed, and that two-thirds of the countries it surveyed had developed some form of legal framework for policy evaluation. Similarly, the UK government sets out requirements and expectations of departments around evaluation in a range of HM Treasury documents. Taken together, central guidance makes it clear that departments are expected to undertake comprehensive, robust and proportionate evaluations of their interventions:

- According to the central government guidance on handling public funds, *Managing Public Money*, “accounting officers should take personal responsibility for ensuring that their organisations’ procurement, projects and processes are systematically evaluated to provide confidence about sustainability, effectiveness, prudence, quality and good value for the Exchequer as a whole, not just for the accounting officer’s organisation”. Finance directors have responsibility for supporting their accounting officer in respecting these standards.
- The Green Book, central government guidance on appraisal and evaluation, states that “all proposals must as part of the proposal contain proportionate budgetary and management provisions for their own monitoring and evaluation”.
- The Magenta Book, central government guidance on evaluation, sets out detailed best practice on evaluation methods, use of evaluation across the policy cycle and disseminating evaluation evidence to policy-makers and the public.
- The Orange Book, central government guidance on risk management, states that “all strategies, policies, programmes and projects should be subject to comprehensive but proportionate evaluation, where practicable to do so”.
- The Government Social Research Publication Protocol states that “government social research and analysis should be published promptly, with the normal maximum being 12 weeks from agreeing the final output”, including evaluation reports.



**1.6** There is no single body with sole responsibility or oversight for the provision of evaluation across government. While departments are responsible for evaluating their policy interventions and learning from existing evaluation evidence, there are several central departments and cross-government functions involved in providing oversight and support of evaluation across government (**Figure 1** overleaf).

**1.7** While accounting officers are personally responsible for evaluating their departments' interventions, the provision and use of evaluation evidence requires effective collaboration between a department's analysis, policy and finance communities. They should consider relevant evaluation evidence to provide advice and support to ministers and senior civil servants so they can make informed decisions.

### **Coverage of evaluation across government**

**1.8** Despite government's commitment to evidence-based decision-making, much government activity is either not evaluated robustly or not evaluated at all. Our 2013 report, *Evaluation in government*, which examined the coverage, quality, use and resource costs of evaluation activity conducted or commissioned by government departments found gaps in coverage. In 2019, the Prime Minister's Implementation Unit (PMIU) found that coverage gaps continue to exist. It examined the scale and quality of impact evaluations for projects within the Government Major Projects Portfolio (GMPP)<sup>7</sup> and across a sample of departmental innovation funds.<sup>8</sup> It found only nine out of 108 major projects – representing £35 billion (8%) of £432 billion in expenditure – had plans for robust impact evaluation in place, while 77 (64% of spend) had no evaluation arrangements (**Figure 2** on page 17). For innovation funds, which are explicitly set up to establish evidence, the analysis showed that only 19% of £19 billion estimated expenditure had robust impact evaluation to enable departments to identify what works.<sup>9</sup> These findings led PMIU to conclude that “government has little information in most policy areas on whether billions of spend are making a difference”. The analysis excluded government projects not on the GMPP, and government does not hold data on the coverage and quality of evaluation of smaller-scale interventions and other ‘business as usual’ activities. As a result of the gaps in evaluation coverage, government cannot have confidence it is spending public money well.

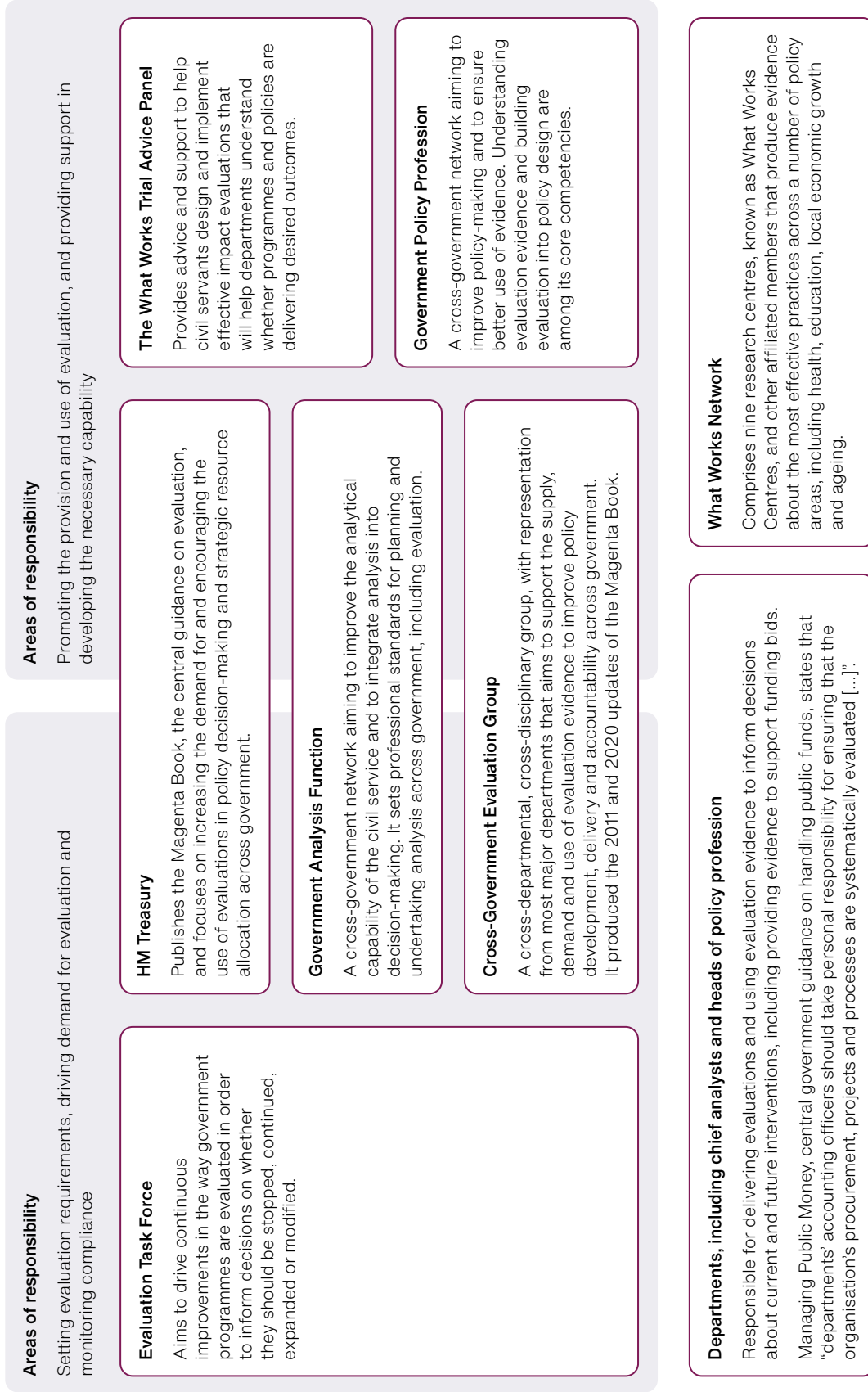
7 The Government Major Projects Portfolio (GMPP) comprises the largest, most innovative and highest risk projects and programmes delivered by government. More detail on GMPP is available at: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/1002310/IPA\\_AR2021\\_final\\_14Jul.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1002310/IPA_AR2021_final_14Jul.pdf)

8 Projects on the Government Major Projects Portfolio are the ones that are more likely to have evaluations in place due to their scale and delivery risk. There is no available data on the coverage and quality of evaluation of government projects that are smaller in scale.

9 Government-backed innovation funds provide support to UK-based businesses or research organisations to: research and develop a process, product or service; test innovative ideas; and collaborate with other organisations.

**Figure 1** The responsibilities of government departments and cross-government functions, relating to provision and use of evaluations

There is no single body with sole responsibility or oversight for the provision and use of evaluation across government

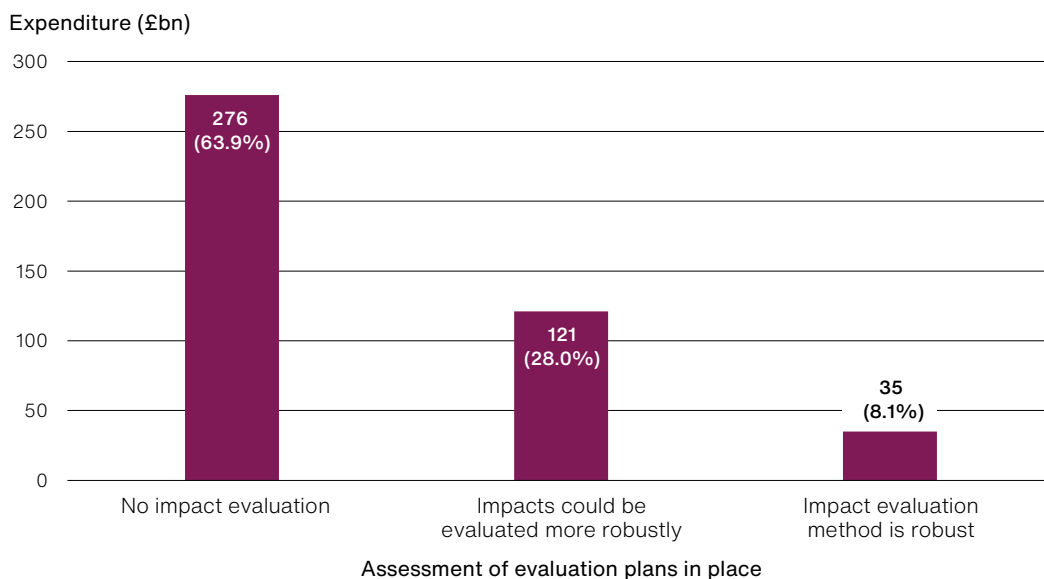


Source: National Audit Office analysis of departmental documents

**Figure 2**

Evaluation gap across the Government Major Projects Portfolio, December 2019

In 2019, the Prime Minister's Implementation Unit's assessment found that only 8% of £432 billion planned expenditure on the Government Major Projects Portfolio had plans for a sufficiently robust impact evaluation

**Notes**

- 1 The Government Major Projects Portfolio (GMPP) was created to improve the delivery of government's biggest and riskiest projects. The Prime Minister's Implementation Unit (PMIU) assessed the scale and quality of impact evaluations of 108 projects on the GMPP, corresponding to £432 billion of planned expenditure.
- 2 PMIU's assessment excluded government projects not on the GMPP. There are no comparable data on the coverage and quality of evaluations of smaller-scale government interventions.

Source: National Audit Office analysis of Prime Minister's Implementation Unit December 2019 review of the scale and quality of impact evaluations in government

**1.9** In our 2013 report, we found that few departments had plans to evaluate all their major projects; plans to evaluate impact or value for money related to only £90 billion of £156 billion in major projects expenditure. Our 2013 assessment of the fitness for purpose of 34 evaluations from four departments found significant variation.<sup>10</sup> Only 14 evaluations were of a sufficient standard to give confidence in the effects attributed to policy because they had a robust counterfactual. There was little systematic information from government on how it had used the evaluation evidence that it had commissioned or produced. However, departments were able to point to a few examples (such as the discontinuation of the stamp duty holiday for first-time buyers in November 2011) where evaluation evidence had clearly informed policy decisions.

<sup>10</sup> In our 2013 report we assessed fitness for purpose using the Maryland Scale, a five-point scale designed by the University of Maryland to classify the strength of evidence.

**1.10** Our latest review of National Audit Office and Committee of Public Accounts reports published since 2013 found persistent weaknesses in relation to evaluation across a range of departments and policy areas. These reports found that many high-profile interventions had poor-quality evaluations, or were not evaluated at all, undermining departments' ability to demonstrate value for money.<sup>11</sup> Previous reports have also criticised departments for insufficient use of evaluation evidence in decision making, as well as for failing to publish results of evaluations in a timely manner.

**1.11** Despite the overall negative picture on evaluation, our reports have highlighted examples of good practice in the provision and use of evaluation evidence. For example, in 2018 the Department for Business, Energy & Industrial Strategy (BEIS) introduced a central analysis, monitoring and evaluation database to refer to when setting up new schemes and to share learning across policies. The Department for Work & Pensions (DWP) demonstrated a commitment to improving the evidence base of its programmes and allocated £100 million to pilot initiatives and evaluation.

**1.12** During our fieldwork for this report we identified examples of government making changes as a result of evaluation findings. In preparing this report we did not audit the robustness of individual evaluations:

- **More funding for programmes where evaluations show they deliver intended outcomes:** The impact evaluation for the Troubled Families programme helped the Department for Levelling Up, Housing & Communities (DLUHC) to secure further funding in the 2020 Spending Review for its Supporting Families programmes.<sup>12</sup>
- **Changes to programmes where evaluations show they could be improved:** The Department for International Trade (DIT) is redesigning its export promotion interventions in response to evaluation evidence on what works. DWP started carrying out weekly rather than fortnightly job-search reviews for claimants of Jobseeker's Allowance after a randomised control trial provided evidence of the positive impact of such a change.
- **Identifying what works:** In 2019, a DLUHC impact evaluation demonstrated that the Rough Sleeping Initiative (RSI) was having a significant effect on reducing street homelessness in the areas that had received RSI funding compared with a counterfactual group.

During our fieldwork for this report, we found few examples of interventions ending where evaluations have shown they do not work as intended. BEIS told us it cancelled the Green Homes Grant after monitoring information the department was collecting for its planned evaluation showed the intervention was not delivering as intended.

11 Examples include: Committee of Public Accounts, *Training new teachers*, HC 73, June 2016; Committee of Public Accounts, *Better Regulation*, HC 487, October 2016; Committee of Public Accounts, *National Citizen Service*, HC 955, March 2017; Committee of Public Accounts, *Transforming Courts and Tribunals: progress review*, HC 27, November 2019.

12 The Department for Levelling Up, Housing and Communities was formerly called the Ministry of Housing, Communities & Local Government.

## Part Two

### Developing a more strategic approach to evaluation and use of evidence

**2.1** This part examines government's actions to provide a more strategic approach to evaluation and the use of evaluation evidence; the requirements and expectations that the centre of government sets for departments; and how departments translate them into their own evaluation activities.

#### **Why a strategic approach is important**

**2.2** A strategic approach to evaluation evidence is important at the centre of government and at a departmental level. This involves being clear about what government seeks to achieve and setting clear arrangements for pursuing its aims and monitoring progress toward them.

**2.3** Our criteria for assessing government's strategic approach are:

- the degree to which the centre of government is active in providing strategic leadership, direction and oversight of activities and resources, including setting clear responsibilities and objectives to assess whether the framework is delivering as intended;
- the clarity of requirements and expectations that the centre of government sets for departments. There should be delivery plans with appropriate capacity and capability and arrangements for monitoring progress in implementing plans; and
- the extent to which departments adopt a strategic approach to evaluation, including planning and prioritising activity to fill significant gaps in their evidence base.

## **Actions by the centre of government to provide strategic direction**

### Actions between 2013 and 2019

**2.4** Between our 2013 report and 2019, the centre of government took some positive but limited steps to improve government evaluation activity:

- At the time of our 2013 report, government had recently set up a network of What Works Centres, responsible for synthesising evaluation evidence on the effectiveness of policy in a range of fields.
- In 2015, the Cabinet Office and the Economic and Social Research Council established the cross-government What Works Trials Advice Panel of around 50 experts within and outside government to support the increased use of controlled experiments in public policy.
- In 2017, the Government Analysis Function was established as a network bringing together analysts from across government. Its role is to lead the analytical community, improve analytical capability and share best practice, including in relation to evaluation.

**2.5** During this time, there were no arrangements at the centre of government to provide leadership and strategic direction for government evaluation activity. There was no clear oversight and view on whether departments were complying with requirements around evaluation, or systematic arrangements to disseminate examples of good practice within departments more widely across government. This contributed to considerable variation in quality of evaluation across departments.

**2.6** There were some levers available to HM Treasury. Our 2013 report recommended that HM Treasury should ask departments to provide evaluation evidence in the context of strategic resourcing decisions such as spending reviews. But HM Treasury did not make use of these levers before the 2020 Spending Review.

### Actions since 2020

**2.7** In 2020, HM Treasury began to look for opportunities to examine the quality of evaluation evidence as part of its budget processes and funding decisions. In the March 2020 Budget, government committed to using the 2020 Spending Review to “require every department to produce plans to improve evaluation of its work” and ensure “all programmes are supported by robust implementation and evaluation plans.” This was supported by the refresh of the Magenta Book by the Cross Government Evaluation Group (CGEG), which provided updated evaluation guidance for departments.

**2.8** HM Treasury used information on evaluation that it had collected from departments in early 2020, to inform the design of the Spending Review process. Departments were asked to provide evaluation evidence for each line of capital budget spend, and each line of new resource budget spend. A joint HM Treasury and Cabinet Office team used an assessment framework to examine departments' arrangements and plans for evaluating each policy area and intended outcomes over the Spending Review period.

**2.9** HM Treasury's assessment of departmental information confirmed varying evaluation coverage and quality across departments and reporting of those arrangements. HM Treasury gave departments additional funding for some programmes where it assessed evaluation arrangements as being particularly good. Following its review of departments' evidence bases and evaluations, and in concluding the Spending Review, HM Treasury sent departments financial settlement letters setting conditions to improve the quality of evaluation arrangements. It asked each department to agree at least five priority evidence gaps to be addressed with robust evaluation and to appoint its Director of Analysis or equivalent to be accountable for delivery, robustness, and use of evaluation. HM Treasury intended a similar approach to setting requirements for evaluations at the 2021 Spending Review.

**2.10** There are further examples of HM Treasury and Cabinet Office being more proactive on evaluation since 2020 (**Figure 3** overleaf), and stakeholders we interviewed have welcomed this change in momentum.

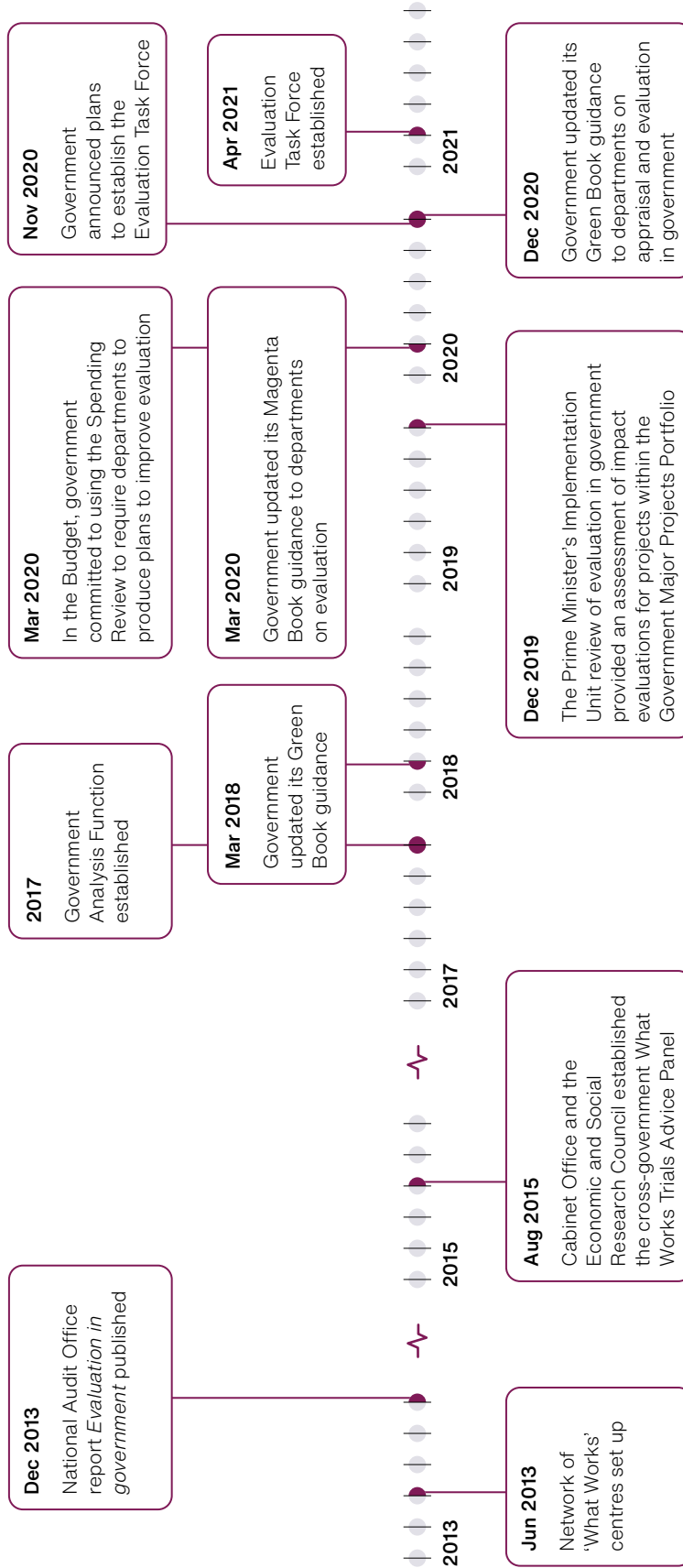
### Clarity of ownership and requirements

**2.11** The overall system of evaluation across government involves many different public bodies and networks. Without a common understanding of roles and responsibilities for different aspects of evaluation, there is potential for inefficiency and reduced effectiveness.

**2.12** The different bodies involved in evaluation in government have a degree of understanding of their roles and responsibilities. But this has not been set out and communicated clearly. In broad terms, the Evaluation Task Force is intended to drive improvements in government evaluations so that robust evidence informs spending and operational decisions. The Analysis Function's focus is on raising skills and capability among the analytical community, and supporting consistent professional standards, including in relation to evaluation. In September 2021, the Analysis Function set up a new Strategy and Delivery Division to strengthen the support to analysts across government. The Policy Profession focuses on improving the policy capability of heads of policy profession and departments, recognising evaluation as a key skill.

**Figure 3** Timeline of important evaluation-related developments since 2013

Since 2020, the centre of government has been increasingly proactive in providing leadership and strategic direction for evaluation activities across government



**Notes**

- 1 The update of the Magenta Book by the Cross-Government Evaluation Group, published in March 2020, included new supplementary guidance on a range of issues, including handling complexity in policy evaluation.
- 2 The Evaluation Task Force aims to: increase the scale and quality of impact evaluation in priority areas; ensure evaluation results inform spending and operational decisions; increase decision makers' access to evaluation results, and third party access to government data for evaluation purposes, by improving transparency; and ensure that evaluation practice in departments is supported by better governance, processes, and resources.
- 3 The Green Book update, in December 2020, strengthened requirements around evaluation, saying "all proposals must as part of the proposal contain proportionate budgetary and management provisions for their own monitoring and evaluation."

Source: National Audit Office analysis of departmental documents



**2.13** There are, however, areas of overlap that need careful management.

The system is reliant on good and frequent communication between key individuals and across organisations and networks. There are gaps in some areas: for example, HM Treasury has not set out clear arrangements for maintaining and promoting the Magenta Book. It was updated by the CGEG as a cross-departmental group of evaluation practitioners in 2011 and 2020. However, it is not clear who is responsible for the governance of the Magenta Book, including commissioning of future updates. There is also no body responsible for collating and communicating good practice and operational guidance.

### Objectives and measuring performance

**2.14** Before the Evaluation Task Force, government did not have a stated objective for the system of evaluation, as distinct from its guidance on how departments should undertake evaluations. The purpose of the Evaluation Task Force is “to drive continuous improvements in the way government programmes are evaluated in order to inform decisions on whether they should be stopped, continued, expanded or modified and ensure robust evidence on the effectiveness of policies and programmes sits at the heart of spending and operational decisions”. The vision of the Analysis Framework is “analysis, evidence and research [that] helps make better decisions to improve outcomes for the UK”.

**2.15** An Oversight Board advises the Evaluation Task Force, approves its strategy and programme of work, scrutinises its performance and offers an escalation route to unblock barriers that the Evaluation Task Force faces. Government has also set out Key Performance Indicators (KPIs) for the Evaluation Task Force to measure progress towards meeting its goals. However, the KPIs do not provide specific targets or other criteria to help assess whether it has delivered its purpose.

**2.16** The government plans a ‘sunset clause’ whereby ministers will decide on the merits of continued funding for the Evaluation Task Force. This decision will be based primarily on reviewing Evaluation Task Force’s performance against its objectives, as assessed by the Oversight Board, and comparing savings and additional revenues that the Evaluation Task Force secures, against the costs of running the team.

## The requirements and expectations that the centre of government sets for departments

**2.17** Across government, the large majority of chief analysts (15 out of 16) and heads of policy profession (13 out of 16) in core departments told us that the goal of evaluation is primarily about learning rather than accountability. It is helpful for both communities to agree on the primary purpose of evaluation, and give sufficient attention to both learning and accountability.

**2.18** Chief analysts told us about the extent that they could fulfil central government requirements. Most were confident that they could meet the technical requirements (including robustness, proportionality and quality assurance), but fewer were confident about other requirements such as timely publication of evaluations (**Figure 4**). Among our case study departments, Department for Levelling Up, Housing & Communities (DLUHC) officials need ministerial approval to publish outputs from evaluations, while Home Office officials told us that there is a general expectation of transparency but that not all evaluations enter the public domain. Department for Education officials told us that they operate with an expectation that evaluations of new interventions are published.

## Departments' strategies for evaluation

### Identifying, prioritising and planning evaluation activities

**2.19** Departments publish information on research questions in Areas of Research Interest documents. They also publish Outcome Delivery Plans, which include a summary of intentions to provide evaluative evidence on priority outcomes. However, these documents do not provide information on how departments prioritise areas for evaluation and how risks or importance of those areas are aligned to priority outcomes.<sup>13</sup> Departments plan their evaluation activities in different ways; some take a more strategic approach than others. Six departments had a single departmental-wide evaluation strategy, seven had only individual strategies in specific policy areas, while three had no evaluation strategy at all. There are benefits to having a single strategy; for example, it can help ensure that the most pressing evaluation gaps are prioritised, planned for and resourced. It also provides an overview of evaluation activity in the department.

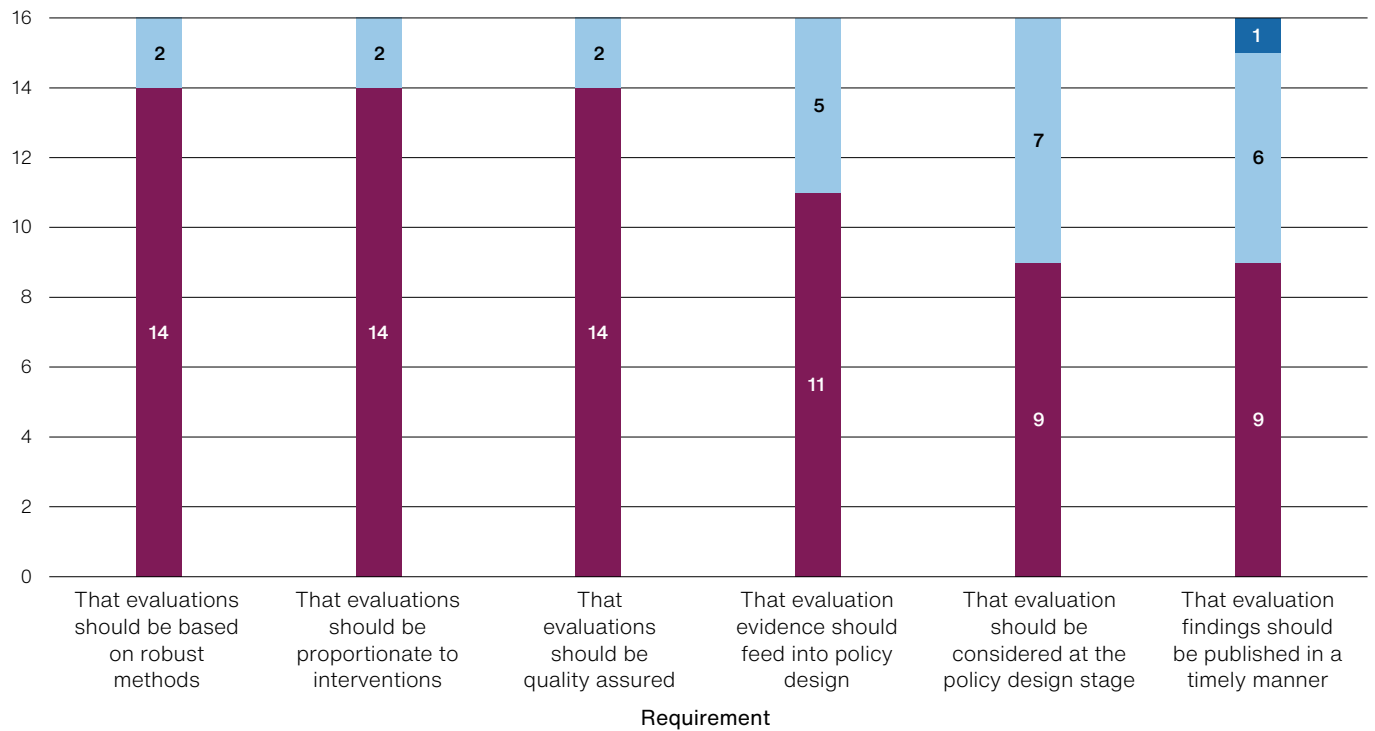
<sup>13</sup> Areas of Research Interest (ARI) give details about the main research questions facing government departments. Outcome Delivery Plans (ODPs) set out each government department's priority outcomes and the department's strategy for achieving them.

**Figure 4**

Chief analysts' views on the extent to which their departments are able to fulfil existing central requirements for evaluation

The majority of chief analysts said that their departments were able to fulfil central requirements for robustness and proportionality of evaluations. Nine out of 16 chief analysts were confident that their departments were able to publish evaluation findings in a timely manner

Count of chief analysts' views



- Unsure
- In some or limited number of cases
- In all or most cases

**Notes**

- 1 We surveyed chief analysts in all 16 core government departments between 28 June and 19 July 2021. The obtained response rate was 100%.
- 2 Survey question: "To what extent is your department able to fulfil each of the following central government requirements for evaluation in relation to your department's policy interventions?"

Source: National Audit Office survey of chief analysts in core government departments

**2.20** This variability was reflected in our case study departments. To improve its strategic approach to evaluation, DLUHC has recently set up a Monitoring and Evaluation Strategy Group whose aims include developing evaluation plans and driving high standards of policy and programme evaluation across all policy areas. DLUHC has recently identified its evaluation priorities across the whole range of its policy areas and is using this process to inform its planned evaluation activity, which is described in its Outcome Delivery Plan for 2021-22.<sup>14</sup> The Department has also set out evaluation plans for different policy areas such as housing.<sup>15</sup> The Home Office allocates its analytical resources across different business areas which decide on evaluation priorities, and it does not have a single evaluation strategy for the whole department.

### Oversight of evaluation activity and resources used

**2.21** Government does not know how much it spends on evaluation activities or how many civil servants work on evaluations. It is therefore difficult for government and departments to assess whether evaluation resources are appropriate to deliver their plans and how to manage resource gaps. Only four out of 16 departments had a readily available estimate of spending on evaluation activity, and only five had a readily available estimate of the number of staff working on evaluation. Lack of visibility of evaluation spend across multiple budgets within departments made it difficult for them to provide evaluation spend figures. In general, departments do not use central budgets to fund evaluation activity, and funding is allocated to evaluation through separate policy areas and as part of specific programme budgets. We identified from publicly available data that the core government departments awarded contracts worth at least £84 million and £67 million in cash terms to conduct or support evaluation activity in 2019-20 and 2020-21 respectively.<sup>16</sup> This is in the context of government spending of £885 billion in 2019-20.

**2.22** The large majority of chief analysts (13 out of 16) told us that there was someone tasked with overseeing the provision, quality and coverage of evaluations in their departments, but only in around half of those cases were they senior civil servants. In three cases, there was no one with specific responsibility. Weaknesses in these arrangements create a risk that departments have no central oversight of evaluation activities and cannot track their progress or identify gaps in evaluation arrangements. Our case studies illustrated the benefits of having this oversight. The Home Office told us it analysed the mismatch between evaluation demand and resources and used this to support a successful bid for more analysts in the 2020 Spending Review. The Department for Education told us that its research and evaluation budget for externally commissioned work is held centrally, with the central research division. There is a clear expectation that policy teams should identify, early on in the policy design process, new areas of evaluation work.

<sup>14</sup> Ministry of Housing, Communities & Local Government, *MHCLG Outcome Delivery Plan: 2021 to 2022*, 15 July 2021.

<sup>15</sup> Ministry of Housing, Communities & Local Government, *Housing Monitoring and Evaluation Strategy*, January 2019.

<sup>16</sup> These estimates do not include the costs associated with evaluations produced by departments' in-house analysts.

## Part Three

### Actions to strengthen provision and use of evaluation in practice

**3.1** This part examines the barriers to effective provision and use of evaluation evidence and the actions that government has taken to address them.

**3.2** Below is an outline of our criteria for an effective evaluation system:

- The centre of government should be clear on the barriers to evaluation and their relative importance. It should develop, implement and monitor the effectiveness of plans for addressing those barriers.
- The centre of government should provide suitable guidance that sets out the principles and standards expected of high-quality evaluations undertaken by departments. It should be clearly communicated and understood by those who need to act on it.
- Government should have a range of effective levers, and enough adverse consequences for non-compliance, to incentivise departments to follow the requirements.
- Departments should be able to routinely access up-to-date thematic learning from relevant evaluations to answer questions or guide future policies.

#### **Understanding barriers to good evaluation and use of evaluation evidence**

**3.3** In December 2019, government reviewed the scale and quality of impact evaluations across government focused on large, high-risk or innovative spend and examined the Government Major Projects Portfolio (GMPP) and departmental innovation funds. It identified many barriers to evaluation, including lack of political interest, limited access to administrative data, and time limits on spending that forced delivery of policies to advance without evaluation planning. These barriers are broadly the same as the factors we identified in our 2013 report, which included a lack of incentives for departments to generate and use evaluation evidence, with few adverse consequences for failing to do so. We set out in **Figure 5** on pages 28 and 29 the supply- and demand-side barriers identified by government, and from our work.

**Figure 5**

## Barriers to better provision and use of evaluation by government departments

Barriers to the production and use of evaluation evidence, on both the demand and supply sides, create risks to the effectiveness of decisions

Area of risk	Barriers to the supply of evaluations	Barriers to the demand for evaluations
<b>Political factors</b>		
Lack of political support	Insufficient input from analysts into decision-making.	Lack of political interest in, or support for, evaluation (mismatch between political timetables and the timelines of evaluation producers; electoral cycle; high civil service turnover; political decisions being driven by values rather than outcomes); departments sceptical that ministers have the appetite for results; project owners disinclined to evaluate due to negative exposure if harmful or neutral impact is detected.
Cultural buy-in	Project owners seeing evaluations as an imposed process or hurdle rather than as necessary Research & Development (R&D) to ensure good value for money and optimal design of future interventions.	
Transparency	Little transparency around results of evaluation, and publication bias against negative findings; central approval process for publication of findings.	Lack of transparency and openness to feedback (fear of uncovering programme failure).
<b>Strategic factors</b>		
Leadership and support from centre	Insufficient or ineffective leadership and support; little central oversight of results of evaluation; centre of government and departments may not always share the same view on the purpose of evaluations.  Lack of HM Treasury scrutiny of whether promised evaluation is delivered once a business case has been approved and variable extent to which it challenges business cases that lack evaluation plans. Departments experience few consequences for not undertaking evaluations.	
Departmental leadership and strategies	No strategic approach to evaluation and spending on evaluation – activity not linked systematically to priority evidence gaps or risk levels.	
Resources	Lack of resources to manage evaluation effectively.	
<b>Technical factors</b>		
Integrating policy and analysis	Weak integration of people and processes to join policy and analysis effectively. Policy not designed with testing and evaluation in mind and analysts involved too late in the policy-making process. Evaluation plans not built into standard project approval processes at a sufficiently early stage.	
Short-termism/behaviour driven by for example spending reviews		Short-termism (spending review time periods); time limits on spending force delivery to proceed without evaluation planning.
Skills	Shortage of evaluation specialists (systemic capacity and skills issues across government); lack of strategic leadership on skills and capability.	Skills gap within policy profession and perceived divisions between policy and analysts.

**Figure 5** *continued*

## Barriers to better provision and use of evaluation by government departments

Area of risk	Barriers to the supply of evaluations	Barriers to the demand for evaluations
<b>Technical factors</b> <i>continued</i>		
Shared outcomes/data	Weak data management and exploitation; limited opportunities to undertake low-cost rapid impact evaluation via secure linked administrative data.	
Learning mechanisms	Limited or non-existent mechanisms for capturing, managing and accessing lessons from evaluations across government. Not sharing lessons learned more widely across government.	

Source: National Audit Office analysis of December 2019 Prime Minister's Implementation Unit review of evaluation in government, and interviews with stakeholders

**3.4** These barriers are consistent with international experience. The Organisation for Economic Co-operation and Development (OECD) reported in 2020 that the main, interdependent, barriers countries face when promoting policy evaluation across government include:

- political interest in, and demand for, evaluation;
- absence of a strategy for policy evaluation that promotes a whole-of-government approach;
- limited availability of human resources (capacities and capabilities) for policy evaluation;
- quality of evidence; and
- limited use of evaluation results in policy-making.<sup>17</sup>

## Views on the relative importance of barriers

**3.5** In our 2013 report, chief analysts and their evaluation staff considered mismatches in timing between production of evaluation evidence and policy decisions, plus a lack of demand from policy colleagues, to be the key barriers to better-quality evaluations and use of evaluation evidence.<sup>18</sup>

<sup>17</sup> Organisation for Economic Co-operation and Development, *Improving Governance with Policy Evaluation: Lessons from Country Experiences*, June 2020.

<sup>18</sup> Comptroller and Auditor General, *Evaluation in government*, National Audit Office, December 2013.

**3.6** The 2019 Prime Minister's Implementation Unit (PMIU) analysis found that many programmes could not be evaluated as robustly as possible because they did not prioritise evaluation at the project design phase. The Magenta Book emphasises the importance of planning an evaluation early, in maximising learning opportunities and reducing costs of data collection. However, we found that almost half of departments could fulfil this requirement to build in evaluation at the policy design stage in only some or a limited number of cases. Ten chief analysts and eight heads of policy profession agreed that the opportunity to learn is not being built into policy design and delivery. Five out of those 10 chief analysts received support from the centre of government on integrating evaluation plans into policy design, but only two were satisfied. The remaining five chief analysts were unaware support is available. Five chief analysts and seven heads of policy profession said that another barrier was the insufficient capacity of analysts to help policy-makers understand evaluation evidence.

**3.7** Chief analysts and heads of policy profession see the barriers differently (**Figure 6**). Effective collaboration and coordination between the various communities will be important for ensuring there are the right strategies in place to address these barriers.

**3.8** The barriers most commonly selected by chief analysts were:

- opportunities to learn are not adequately built into policy design and delivery;
- evaluation evidence is not understood enough by policy-makers; and
- the lack of pressure and demand from senior policy colleagues on evaluation.

**3.9** The barriers most commonly selected by heads of policy profession were:

- evaluation evidence is not available when needed;
- opportunities to learn are not adequately built into policy design and delivery; and
- the knowledge base of evaluation evidence is difficult to access.



**Figure 6** Barriers to using evaluation evidence to inform decision-making in government departments  
 Our surveys of chief analysts and heads of policy profession found that they had different views on the most commonly selected barriers to using evaluation to inform decision-making



**Potential barriers to evaluation**

- Chief analysts
- Heads of policy profession

**Notes**

- 1 We surveyed chief analysts and heads of policy profession in all 16 core government departments between 28 June and 19 July 2021. We obtained a response rate of 100% in both surveys.
- 2 Survey question: "For each of the following factors, to what extent do you agree or disagree that it is a barrier to using evaluation evidence to inform decisions on your department's interventions?"

Source: National Audit Office surveys of chief analysts and heads of policy profession in core government departments

## **Actions to address barriers, and remaining gaps**

### Actions on political factors

#### **Political and cultural**

**3.10** Political interest and support are key factors in promoting effective use of evaluation and, since 2020, there has been increased interest in evaluation from HM Treasury and the Cabinet Office. Conversely, lack of political interest and support can be a barrier which influences the appetite for commissioning evaluations and how the results are used. Short electoral cycles and high turnover of ministers can create inconsistent demand for evaluation. Achieving recognition of the value of evaluation evidence can be a challenge, particularly among senior policy colleagues in departments. Around half of chief analysts (seven out of 16) thought that there was a lack of pressure and demand from senior policy colleagues in support of evaluation. The Cabinet Office told us that it is undertaking work to understand how people behave in their interactions with others across the civil service and that it plans to draw lessons from this wider work to inform its actions on evaluations.

**3.11** The 2019 PMIU review found evidence that it was hard to embed a culture of open enquiry and overcome the temptation to seek to justify chosen policies. Concerns that evaluations will produce ‘unhelpful findings’ about government initiatives may mean evaluation findings are not used or even that evaluations are not undertaken. However, only three out of 16 chief analysts and four out of 16 heads of policy profession thought that evaluation evidence which does not support the political direction or commitments was a barrier to it being used.

#### **Transparency**

**3.12** Government recognises that the principle of transparency is essential to good evaluation, but there are few mechanisms in place to ensure or monitor that departments publish evaluation reports and share findings. Government guidance is clear on the need for transparency. The Magenta Book says departments must inform the public about the outcomes from their work and be accountable for their spending, and that publishing the communications plan, so external observers are aware of what will be published when, is good practice. It also highlights the importance of transparency in supporting credibility and that the presumption should be for maximum openness. The Green Book emphasises that evaluation reports and underlying research should be published in line with government transparency standards and Government Social Research: Publication Protocol.<sup>19</sup>

<sup>19</sup> HM Treasury, The Green Book, Central Government Guidance on Appraisal and Evaluation, 2020, paragraph 8.17. Government Social Research: Publication Protocol available at: [www.gov.uk/government/publications/government-social-research-publication-protocols](http://www.gov.uk/government/publications/government-social-research-publication-protocols)

**3.13** Many countries adopt a transparent approach to reporting evaluation results. In 2020, the OECD reported that 18 of the 42 countries that it surveyed make evaluation findings and recommendations available to the public by default.<sup>20</sup> In Norway, for example, an online evaluation portal provides public access to evaluations carried out on behalf of ministries and other state enterprises.<sup>21</sup>

**3.14** Since 2015, the Committee of Public Accounts has been highlighting departments' poor timeliness in publishing evaluation findings. There is no centrally held information on how many evaluations are completed but not published by departments. 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 of 16) told us that they could in only some or a limited number of cases publish evaluation findings in a timely manner. The Evaluation Task Force has a core responsibility to maintain a public register of evaluation plan summaries and trial protocols and to encourage departments on the timely publication of evaluation results on GOV.UK to support action on the results and to improve accessibility for decision-makers. This is to manage 'publication bias' – the risk that evaluations which have negative findings may not get published. The Evaluation Task Force told us it is planning to improve transparency of the evaluations that are commissioned and published, by collating the information to maintain the register.

**3.15** In our 2013 report, we noted that "independent evaluators outside of the government experience difficulties accessing a range of official and administrative data that can be used to evaluate the impact of government interventions". The Office for National Statistics (ONS) Secure Research Service makes some data available to external researchers, and the ONS aims to enhance this service as part of its Integrated Data Programme. The Evaluation Task Force has a responsibility to encourage an 'open by default' approach to publishing programme data so that programmes can be evaluated by third parties, but there is not yet a mechanism to ensure this happens.

## Actions on strategic factors

### Support from the centre

**3.16** The requirements on departments to evaluate are set out in several guidance documents. Government's guidance to departments is largely principles-based, with little practical guidance such as on good practice, toolkits and operational guidance.

**3.17** As well as setting requirements for evaluation, HM Treasury and the Cabinet Office, together with cross-government functions, offer support to departments on evaluation. This includes help with the design and delivery of proportionate evaluation and scrutiny of evaluation plans.

<sup>20</sup> OECD, *Improving Governance with Policy Evaluation: Lessons from Country Experiences*, 1 June 2020.

<sup>21</sup> Available at: [www.evalueringportal.no/om-evalueringportalen](http://www.evalueringportal.no/om-evalueringportalen)

**3.18** We found that there was substantial variation in chief analysts' awareness of the support that the centre of government provides (**Figure 7**). This means that opportunities may have been missed to improve the amount and the robustness of evaluation, and the learning that can take place across government.

**3.19** Our survey showed differences in how satisfied chief analysts were with evaluation support from the centre of government (**Figure 8** on page 36). Those who received support 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.

### Actions on technical factors

#### **Integrating policy and analysis**

**3.20** Effective use of evaluation evidence requires that evaluation evidence feeds into policy design. However, around one third (five out of 16) of chief analysts told us that this was possible in only some or a limited number of cases.

**3.21** Plans to evaluate and learn should also be integrated into policy design and implementation. However, there are limitations and challenges (paragraph 3.6). Good links between policy and analysis communities help facilitate integration of evaluation into policy design.

**3.22** We saw examples where departments have made progress on linking their analysis functions and the evidence they generate to policy-making. For example, the Department for Levelling up, Housing & Communities (DLUHC) has recently updated its business case template to include a requirement to provide details of any ring-fenced costs for impact evaluation of all investments of more than £100 million, and for smaller investments which are innovative, contentious or untested. Such behavioural nudges encourage close working of analysts with policy-makers at the early stages of interventions. **Figure 9** on page 37 shows other examples.

#### **Skills**

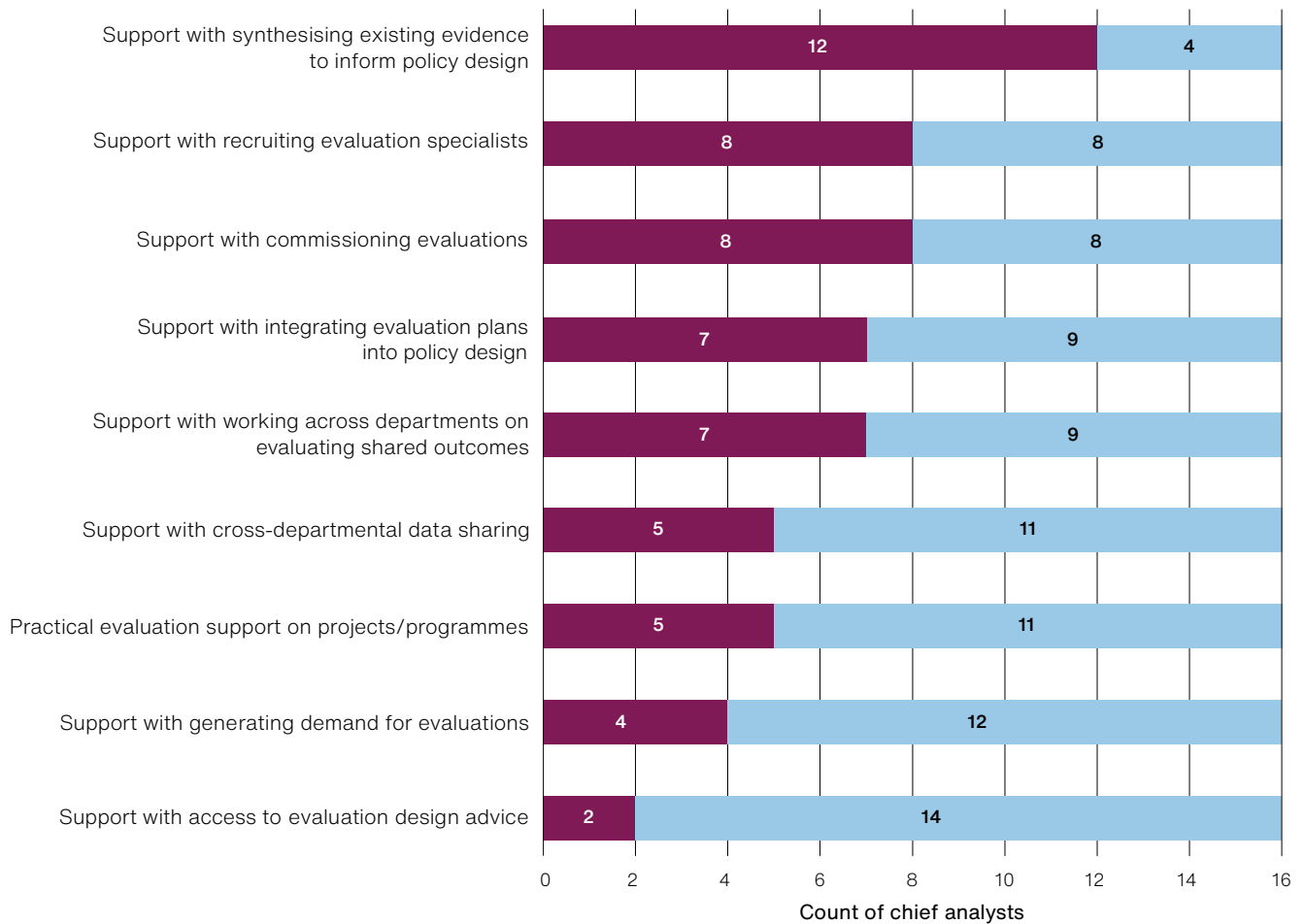
**3.23** Departmental chief analysts told us about challenges in recruiting and retaining skilled evaluators and senior analytical capability. The updated version of the Magenta Book published in March 2020 was accompanied by a framework setting out the knowledge and skills needed by analysts to deliver quality evaluations. The Analysis Function Career Framework sets out the skills and experience needed for analytical roles across government – including those involved in evaluation.

**Figure 7**

## Chief analysts' awareness of evaluation support provided by the centre of government

Our survey of chief analysts found that there was substantial variation in awareness of different types of evaluation support

Type of support



■ Unaware this support is available

■ Aware this support is available

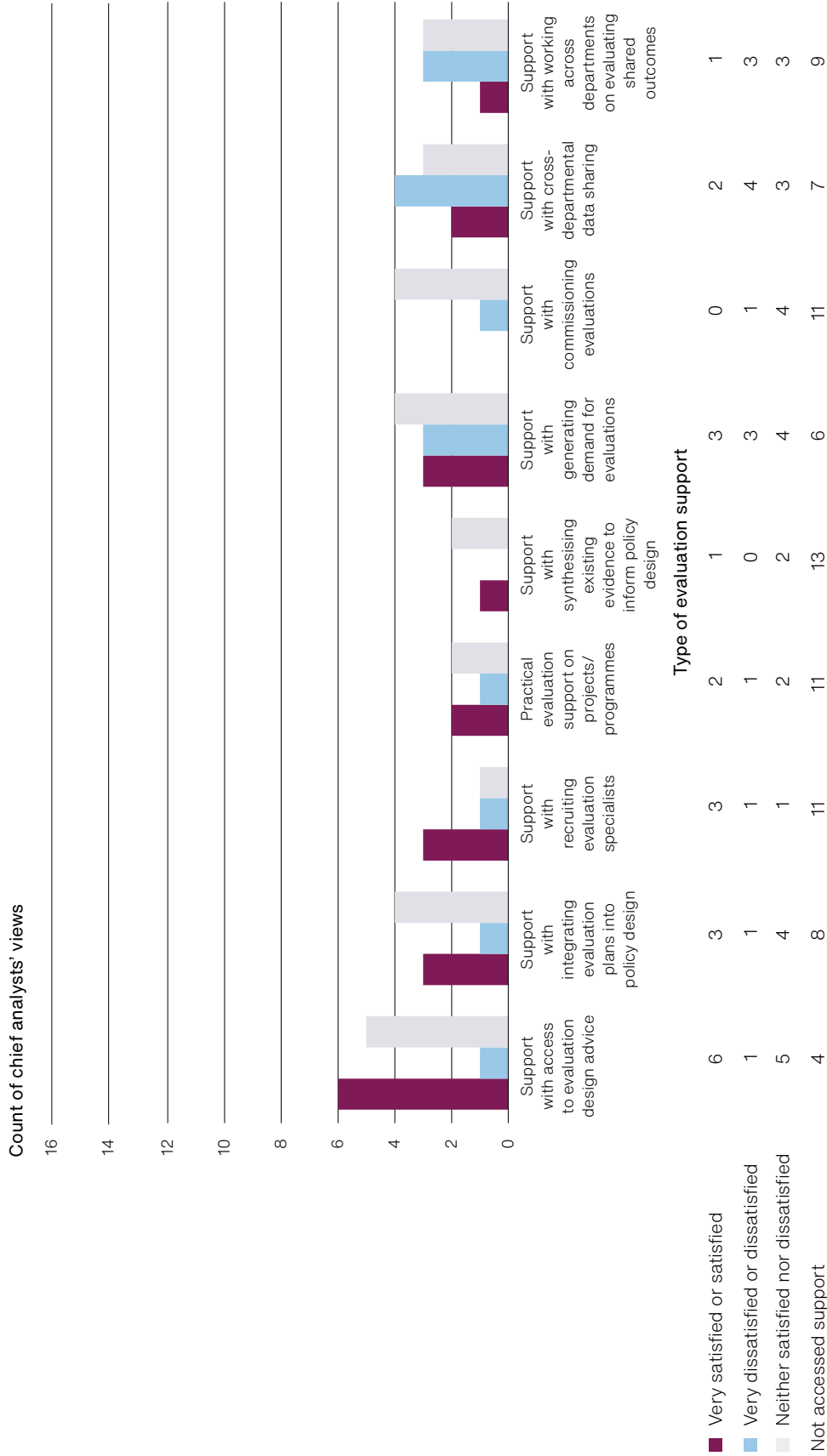
**Notes**

- 1 We surveyed chief analysts at all 16 core government departments between 28 June and 19 July 2021. The response rate was 100%.
- 2 Survey question: "How satisfied or dissatisfied are you with the support provided by the centre of government (i.e. HM Treasury and Cabinet Office) with the following aspects of evaluation activity?" Note that the survey question included an 'Unaware this support is available' response option.

Source: National Audit Office survey of chief analysts in core government departments

**Figure 8** Chief analysts' satisfaction with the evaluation support accessed by their departments

Our survey of chief analysts found a mixed picture of their satisfaction with evaluation support provided by the centre of government. They were most satisfied with support on access to evaluation design and least satisfied with support on cross-departmental data-sharing



**Notes**

- 1 We surveyed chief analysts in all 16 core government departments between 28 June and 19 July 2021. The response rate was 100%.
- 2 Survey question: "How satisfied or dissatisfied are you with the support provided by the centre of government (i.e. HM Treasury and Cabinet Office) with the following aspects of evaluation activity?"

**Figure 9**

## Examples of actions to bridge the gap between policy and evaluation

**Our review of evaluation arrangements in our case study departments, the Home Office, Department for Education and Department for Levelling Up, Housing & Communities (DLUHC) found evidence of actions taken to better embed evaluation in policy-making**

Action area	Example
Evaluation awareness training for policy staff	The Home Office has prepared a training pack for policy professionals that explains concisely the purposes of evaluation, why evaluation should be embedded in the policy process and the implications of not doing so.
Policy tests that incorporate evaluation	<p>The Home Office's internal <i>Policy Tests</i> guide prompts policymakers to ask themselves a series of fundamental questions to test the robustness of a policy. When considering "Will it work and how will you know?", policymakers are urged to plan how to measure and evaluate success, including impact and cost-effectiveness.</p> <p>The Department for Education's five 'Policy Tests', which include prompts to consider whether policy advice is based on the very latest expert thinking and whether ideas are feasible in practice, encourage policymakers to review delivery processes and implementation plans within early policy formulation.</p>
Close working with policy and operational staff	When evaluating the Troubled Families Programme, DLUHC analysts worked jointly with an evaluation and evidence lead in the policy team throughout the evaluation. There were regular meetings with policy staff to make joint decisions about the evaluation. Analysts shared findings with senior and working-level officials at regular intervals throughout the evaluation. The DLUHC provided reports and held discussions with local authorities on the findings.
Evidence packs and presentations	DLUHC shares an evidence pack on homelessness and rough sleeping with new senior policy officials and ministers when they join the Department. Analysts attend board and strategy meetings to provide updates on evaluations, hold teach-ins for policy-makers on the findings, and produce outputs that are tailored to the interests of operational staff.

Source: National Audit Office analysis of Home Office, Department for Education and Department for Levelling Up, Housing & Communities documents

**3.24** The Analysis Function recognises that policy officials need to have sufficient 'evaluation literacy' to understand the evidence before making decisions. The lack of understanding of the evaluation evidence was an issue expressed in different ways by both chief analysts and heads of policy profession. Our case study departments are working to help policy and operational staff understand and make more use of evaluation (Figure 9).

**3.25** In 2018, the Policy Profession introduced policy profession standards that include evaluation as a key skill. In 2021, the Analysis Function launched a Capability Framework, which aims to help senior civil service officials consider their existing skills and actions to develop their capabilities further. The Analysis Function told us that discussions are under way to embed learning on evaluation within the Civil Service Fast Stream programme and the Government Skills Curriculum. The Analysis Function also told us it expects to complete an audit of analytical skills among policy officials in March 2022, and produce an action plan.

### **Shared outcomes and data**

**3.26** Access to a range of data is important in evaluating whether policy interventions are being implemented and delivering outcomes as intended. The Magenta Book explains how important it is to plan early what data and evidence should be collected before and during the policy intervention.

**3.27** Most of the barriers to evaluating shared outcomes that cut across departmental boundaries involve issues such as data access, knowledge of sources, data linking, manipulation and quality. Few departments operate their analysis systems across departmental boundaries, and while some departments share data bilaterally, there is still a lot of friction in the system. In 2019, our *Challenges in using data across government* report found a lack of common data models and standards, poor data quality and difficulties in extracting and sharing data.

**3.28** Our survey of chief analysts found that out of the nine departments that had accessed central support with cross-departmental data sharing, only two reported that they were satisfied with the help provided. As part of the Integrated Data Programme, government is developing an Integrated Data Service (IDS) to make sharing data easier across government. The programme aims to improve evidence-based decision-making and the speed and effectiveness of cross-government analysis.

### **Learning mechanisms**

**3.29** Government is not managing its knowledge base effectively. This means that it is missing opportunities to improve outcomes and risking duplication of effort. Almost all chief analysts and heads of policy profession said that the purpose of evaluation was primarily about learning rather than accountability. Despite this, heads of policy profession told us they thought that evaluation evidence was not available when needed (eight out of 16); was hard to access and understand (eight out of 16); and was not robust or sufficient to be useful (six out of 16). One quarter of chief analysts agreed that there were difficulties in accessing the knowledge base on what works.



**3.30** The What Works Centres aim, among other things, to collate existing evidence on the effectiveness of programmes, and produce high-quality synthesis reports and systematic reviews.<sup>22</sup> Government is not taking full advantage of opportunities to bring together evaluation findings, extract the learning and apply lessons across different departments. The Home Office produces evidence packs for internal use that cover policy areas such as crime and policing. These packs provide a succinct overview of trends, evidence of what works and remaining evidence gaps. We also found some evidence of sharing specific findings with interested parties across government. The Evaluation Task Force told us it plans to publish a range of material including information about evaluations planned and published.

**3.31** Government is not capturing and publicising how evaluations lead to improved outcomes. At present information on examples like these is only held at departmental level.

### **Scrutiny**

**3.32** HM Treasury enhanced its level of scrutiny of evaluation arrangements at the 2020 Spending Review. It set out conditions in the settlements it agreed with departments, requiring actions in relation to evaluation arrangements at departmental level and for specific programmes. Departments told us that, while these conditions are helpful, HM Treasury has not put in place a formal process to monitor whether departments have implemented them.

**3.33** Outside of Spending Reviews, HM Treasury has few levers and exercises little scrutiny to ensure that departments comply with evaluation requirements and expectations. There are no systematic processes at the centre of government for following up whether departments are complying with the requirements set out in *Managing Public Money*, the *Magenta Book* or the *Analysis Functional Standard*. This means that examples of departments not meeting requirements are likely to persist, and no one is tracking and acting on this at a cross-departmental level. Within departments, we found some examples of quality assurance structures. DLUHC, for example, requires all proposals above a minimum monetary threshold for commissioned research to go through a quality gateway, comprising a panel of experts from across the department, to provide challenge and scrutiny.

<sup>22</sup> The What Works Network comprises nine independent What Works Centres, three affiliate members and one associate member. The What Works Centres cover policy areas such as crime reduction, homelessness and children's social care. A full list of What Works Centres can be found at [www.gov.uk/guidance/what-works-network](http://www.gov.uk/guidance/what-works-network).

# Appendix One

## Our audit approach

**1** This report examines government’s progress in developing the provision and use of evaluation evidence across government, since our 2013 report. We considered:

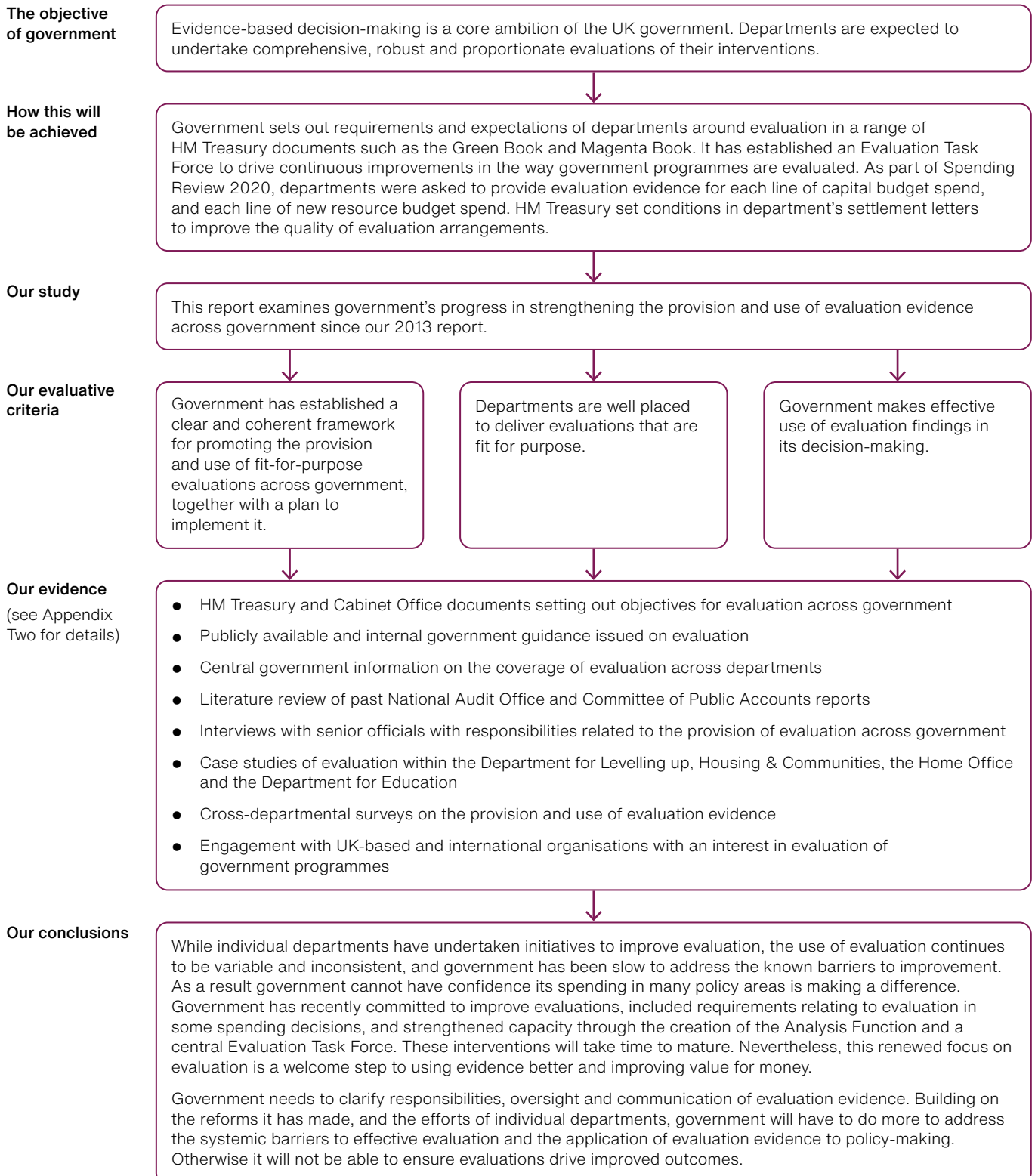
- how government has acted on our 2013 report recommendations;
- the extent to which government has adopted a strategic approach to evaluation; and
- progress in addressing barriers to good evaluation and good use of evaluation evidence.

We did not consider how policy decision-makers draw on other types of evidence, nor did we assess the robustness of individual evaluations.

**2** We applied an analytical framework with evaluative criteria that considered whether government manages its evaluation system effectively to support improved outcomes for the public.

**3** Our audit approach is summarised in **Figure 10** and our evidence base is described in Appendix Two.

**Figure 10**  
Our audit approach



## Appendix Two

### Our evidence base

- 1** We reached our conclusion on government's progress in developing the provision and use of evaluation evidence across government since our 2013 report *Evaluation in government* by analysing evidence collected between June and November 2021.
- 2** We applied our analytical framework to assess the provision and use of evaluation evidence across government. Our audit approach is outlined in Appendix One.
- 3** We examined government's progress in developing the provision and use of evaluation evidence by:
  - a** reviewing and analysing documents from HM Treasury, Cabinet Office, Analysis Function and Policy Profession setting out objectives for evaluation across government, including those related to the establishment and functioning of the Evaluation Task Force. We also reviewed publicly available and internal guidance issued in relation to evaluation. We also examined monitoring information on the coverage of evaluation across departments collected by the centre of government, including outputs of its assessment of departments' arrangements for evaluation;
  - b** reviewing and analysing past National Audit Office (NAO) and Committee of Public Accounts reports commenting on the provision and use of evaluation evidence as part of their assessment of government programmes. We reviewed NAO reports since 2013 and Committee of Public Accounts reports since 2015 and identified common weaknesses and good practice examples in relation to the provision and use of evaluation evidence. The findings of this analysis fed into our interviews with senior officials from the centre of government about barriers to the provision of evaluation;
  - c** conducting interviews with senior officials with responsibilities related to the provision of evaluation across government. We held meetings with staff from the Evaluation Task Force, Cabinet Office, HM Treasury, Analysis Function and Policy Profession to understand the progress made in incentivising and overseeing the provision and use of evaluations across departments. Interview notes were triangulated with document review findings and our analysis of past NAO and Committee of Public Accounts reports;

- d** examining structures and processes in relation to evaluation in three case study departments: Department for Levelling Up, Housing & Communities, Home Office and the Department for Education. We selected these departments based on selection criteria that considered cost and scale of interventions, public and Committee of Public Accounts interest and other aspects leading to variation in evaluation coverage. We requested and analysed the same collection of documents from each of the three departments, covering areas such as their understanding of central requirements, their internal structures and processes for delivering fit for purpose evaluations, arrangements for sharing evaluation evidence with decision-makers and actions taken in response to evaluation findings. We also carried out in-depth interviews with senior officials with responsibilities related to the provision of evaluations in each department;
- e** carrying out two cross-departmental surveys on the provision and use of evaluation evidence, one targeted at chief analysts and one at heads of policy profession. Both surveys covered all 16 core government departments. They were intended to provide a cross-government picture of the use of evaluation evidence across government. Both surveys obtained a 100% response rate, with every department submitting a survey response. The survey of chief analysts sought their views on the purpose of evaluation, barriers to using evaluation evidence, clarity of central requirements and the support provided by the centre of government. It also asked for information about departments' spending on evaluations and use of evaluation strategies. The survey of heads of policy profession sought their views on the purpose of evaluation and barriers to using evaluation evidence;
- f** engaging with UK-based and international organisations with an interest in evaluation of government programmes. We analysed responses from Supreme Audit Institutions in response to our request for reports examining the provision of evaluation in their respective governments. We received responses from Denmark, Finland, France, New Zealand, Norway, Sweden and the US. We also conducted interviews with several third-party organisations within the UK to understand their perspective on the state of evaluation in government;

- g** using government's publicly available database of contracts, Contracts Finder, to estimate the amount that core government departments have spent on contracts to support their evaluation activity between 2019-20 and 2020-21. Our approach to identifying contracts related to evaluation consisted of searching the database using the search terms 'evaluation' and 'evaluate' and cleaning the resulting list of contracts to exclude false matches. Our final estimates include contracts seeking external providers to undertake an evaluation of departments' programmes, as well as contracts seeking expert advice or training on evaluation methods. Note that Contracts Finder only covers contracts awarded by government worth more than £10,000. Low-cost evaluation contracts are thus excluded from our estimate. Furthermore, our estimate only covers spend by the core government departments and excludes contracts awarded by local government and arm's-length bodies. For these reasons our estimate of spend on contracts to support departments' evaluation activity should not be interpreted as the amount the whole of government spends on external evaluations.

This report has been printed on Pro Digital Silk and contains material sourced from responsibly managed and sustainable forests certified in accordance with the FSC (Forest Stewardship Council).

The wood pulp is totally recyclable and acid-free. Our printers also have full ISO 14001 environmental accreditation, which ensures that they have effective procedures in place to manage waste and practices that may affect the environment.



National Audit Office

Design and Production by NAO Communications Team  
DP Ref: 009292-001

£10.00

ISBN 978-1-78604-400-6



9 781786 044006