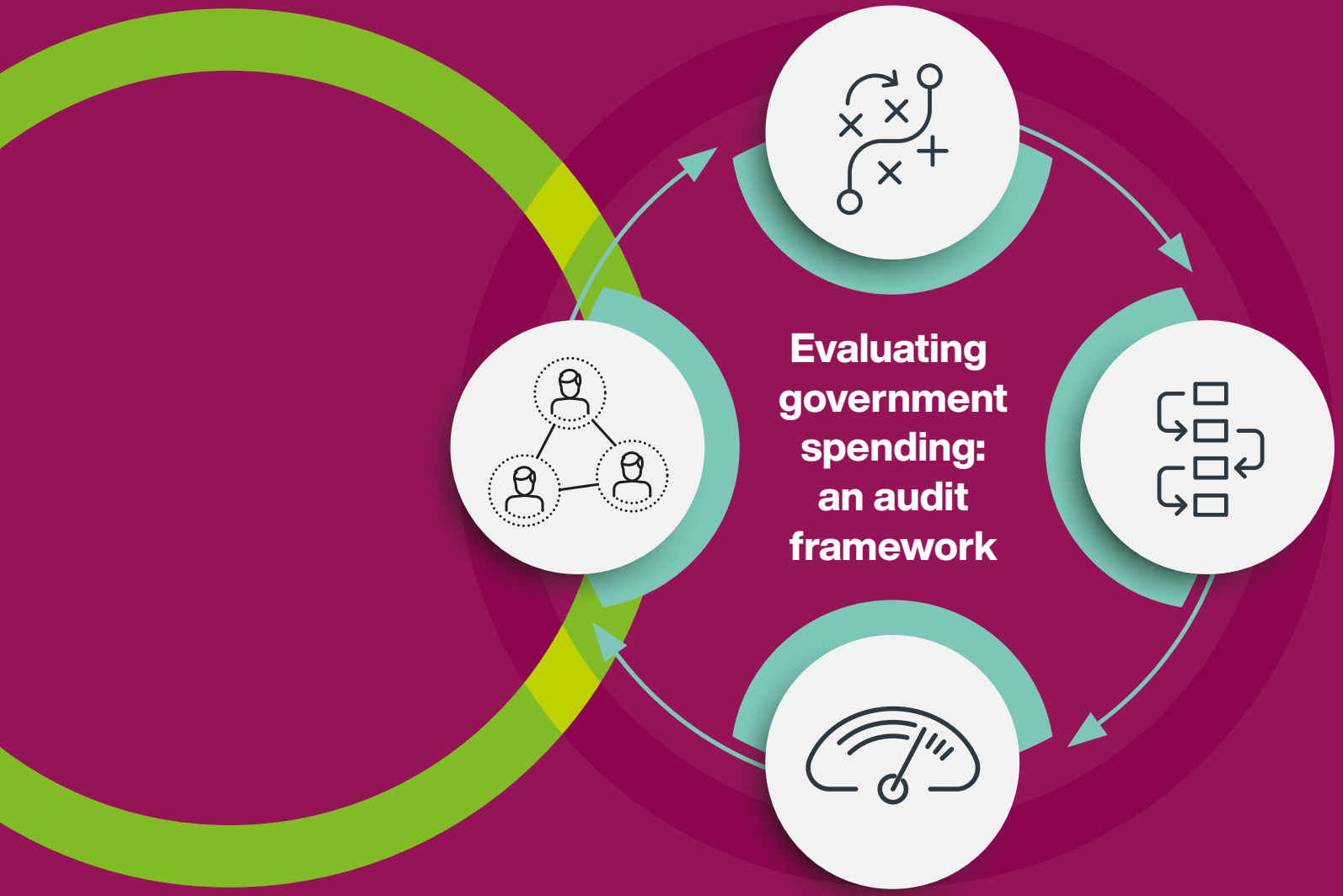
Good practice guidance

Evaluating government spending: an audit framework

National Audit Office



April 2022

Design & Production
by Communications Team
DP Ref: 011054

We are the UK's independent public spending watchdog

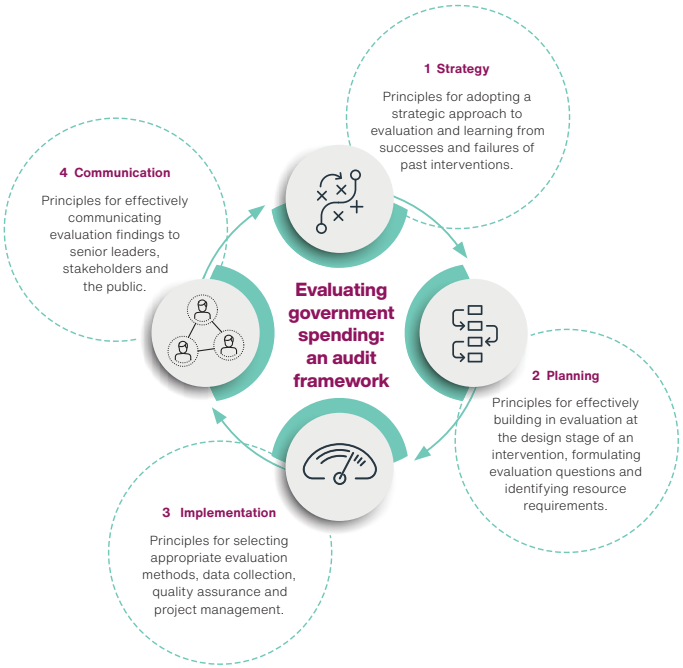
1 Strategy

Align evaluation strategy with organisational strategic objectives

- Is there a single evaluation strategy for the department that links to priority outcomes?
- Are the most pressing evaluation gaps prioritised and appropriately resourced?

Understand the extent to which previous interventions have worked

- Do senior leaders use the lessons from previous evaluations?
- Is there a regularly updated knowledge base built on robust and systematic reviews of existing evaluations and their applicability to the relevant UK context?



2 Planning

Plan to incorporate evaluation throughout the policy cycle

- Is evaluation evidence used to help shape the design of policy options?
- Is there a plan for how emerging evaluation evidence from that intervention will be used to improve it?

Be clear about the questions to be answered by the evaluation

- Is there a clear link between evaluation plans for the intervention and the departmental evaluation strategy?
- Is the intervention designed with evaluation in mind?
- Will answering the evaluation questions provide the insights senior leaders need to make decisions?

Consider evaluation requirements at the business planning stage, including early identification of resources needed

- Are evaluation plans and resources scrutinised at the investment approval stage?
- Are there clearly defined processes, responsibilities and governance arrangements?
- Are evaluation plans proportionate to the size, complexity and risk of the intervention?

Engage with stakeholders in planning evaluations

- Have stakeholders been engaged in the design and planning of evaluations to ensure evaluation activities are relevant and useful?
- Have likely impacts on different segments of the population been considered?

3 Implementation

Pick the most appropriate evaluation method

- Will the method(s) provide insight into whether the intervention has worked as intended?
- Is there a clear, robust counterfactual to assess what would have happened in the absence of the intervention?
- Have the right outcome measures to quantify impact been chosen?
- Have sufficient analytical and other resources been allocated to deliver the chosen evaluation method?

Assess and collect data

- Are data required already available, or do new data collection systems need to be set up?
- Are pre-planned data quality checks built in?

Build in and carry out quality assurance throughout the evaluation

- Are there quality assurance steps at each stage of the evaluation?
- Is there peer review by experts independent of the intervention, and evidence of implementation of peer review recommendations?

Keep the evaluation on track using good project management

- Do governance and assurance structures provide oversight and challenge?
- Is there an evaluation leader who takes personal ownership of the project, and has the necessary authority and influence?
- Are evaluation plans realistic and achievable?
- Is there an active risk management plan that helps to mitigate emerging risks?

4 Communication

Communicate evaluation findings to senior leaders at key decision points

- Is there a process to share evaluation findings regularly with senior leaders?
- Are there opportunities to assess, with the support of relevant evaluation experts, whether an intervention needs to be amended, expanded or stopped as a result of evaluation findings?
- Are decisions taken in response to evaluation findings recorded?

Publish and share findings with a suitable range of stakeholders in a timely manner

- Are the evaluation findings being published in line with government requirements?
- Have the data, methodology and limitations of the evaluation been described adequately?
- Are findings communicated in ways that are useful to different stakeholders, and are the implications for good practice clear?

Collate evaluation findings and lessons learned so that they can be easily accessed and shared

- Have lessons been identified and captured on what works and why, and about how the evaluation was designed and implemented?
- Can the evaluation evidence base be accessed easily by those who would benefit?

Introduction

What is evaluation and why is it important? | What is the guide about and how do I use it? | What does the guide cover?

Why is evaluation important?

Evaluation is a systematic assessment of the design, implementation and outcomes of an intervention, according to the government Magenta Book guidance. Its purpose is to provide insights into how an intervention has been implemented, and its impacts – who has been affected, how and why.

Evaluation is a vital tool that can help government make informed decisions about whether to launch, continue, expand or stop an intervention. It can help government to learn what works from past interventions and improve the design and implementation of future ones, thereby improving the likelihood of successfully achieving objectives. Evaluation also supports accountability to Parliament, scrutiny bodies and the public by allowing them to understand the difference being made by public spending.

Central government guidance makes it clear that departments are expected to undertake proportionate and robust evaluations of their interventions. *Managing Public Money*, central government guidance on handling public funds, states that accounting officers should take personal responsibility for ensuring that their organisations' procurement, projects and processes are systematically evaluated. However, our value-for-money work continues to find many examples of evaluation not being carried out, as well as weaknesses in how it is used in decision-making. Our 2021 report *Evaluating government spending* has identified areas of good practice and common challenges for effectively providing and using evaluation evidence to improve outcomes.

What is evaluation?

There are three different types of evaluation: ***process, impact*** and ***economic*** (also known as value-for-money evaluation):

- Process evaluation examines activities and implementation ("What can we learn from how the intervention was delivered?").
- Impact evaluation focuses on the impact of an intervention and identifies the change in outcomes directly attributable to an intervention, as well as its impact on different groups ("What difference has the intervention made?").
- Economic evaluation (or value-for-money evaluation) compares the benefits and costs of an intervention and assesses whether an intervention was a good use of resources ("Was the intervention worth the cost?").

Introduction

What is evaluation and why is it important? | What is the guide about and how do I use it? | What does the guide cover?

What is the guide about and how do I use it?

This guide provides the National Audit Office's (NAO's) perspective on what we look for in terms of evaluation in the different stages of the policy cycle. It will help departments understand how we audit evaluation evidence and the overall evaluation system that departments operate.

The framework is not a technical guide for evaluation practitioners. It is intended to complement central government guidance on evaluation, such as the Magenta Book.

Who this guide is aimed at

We have prepared this guide to help accounting officers, other senior leaders and policymakers understand how the NAO assesses the robustness of the evaluation arrangements in departments. It should support departments in strengthening these arrangements to help government secure value for money from its interventions. This guide is also intended to help scrutiny bodies assess departments' evaluation arrangements. Our main audiences are:

- **senior leaders in departments** who are accountable to Parliament for evaluating the implementation and outcomes of interventions; and
- **members of Parliament, their staff and other stakeholders, such as departmental Audit and Risk Committees,** who wish to hold departments to account.

Evidence base

This guide builds on findings and recommendations from NAO reports, information from official government guidance on evaluation and international comparisons of evaluation systems. Sources include:

- our 2013 report *Evaluation in government*, which focused on the coverage and quality of impact and economic evaluations across government and identified barriers to the production and use of evaluation evidence;
- our 2021 report *Evaluating government spending*, which examined the state of evaluation across government and progress in addressing systemic barriers to the provision of evaluations and use of evaluation evidence;
- previous NAO reports which commented on the provision and quality of evaluations undertaken across a range of departments and policy areas;
- The Magenta Book, HM Treasury's guidance on evaluation;
- wider government guidance documents which relate to evaluation, such as Managing Public Money, The Green Book and the Government Social Research: Publication protocol; and
- Organisation for Economic Co-operation and Development (OECD) guidance documents and cross-national studies on using evaluation evidence in policy-making.

Introduction

What is evaluation and why is it important? | What is the guide about and how do I use it? | What does the guide cover?

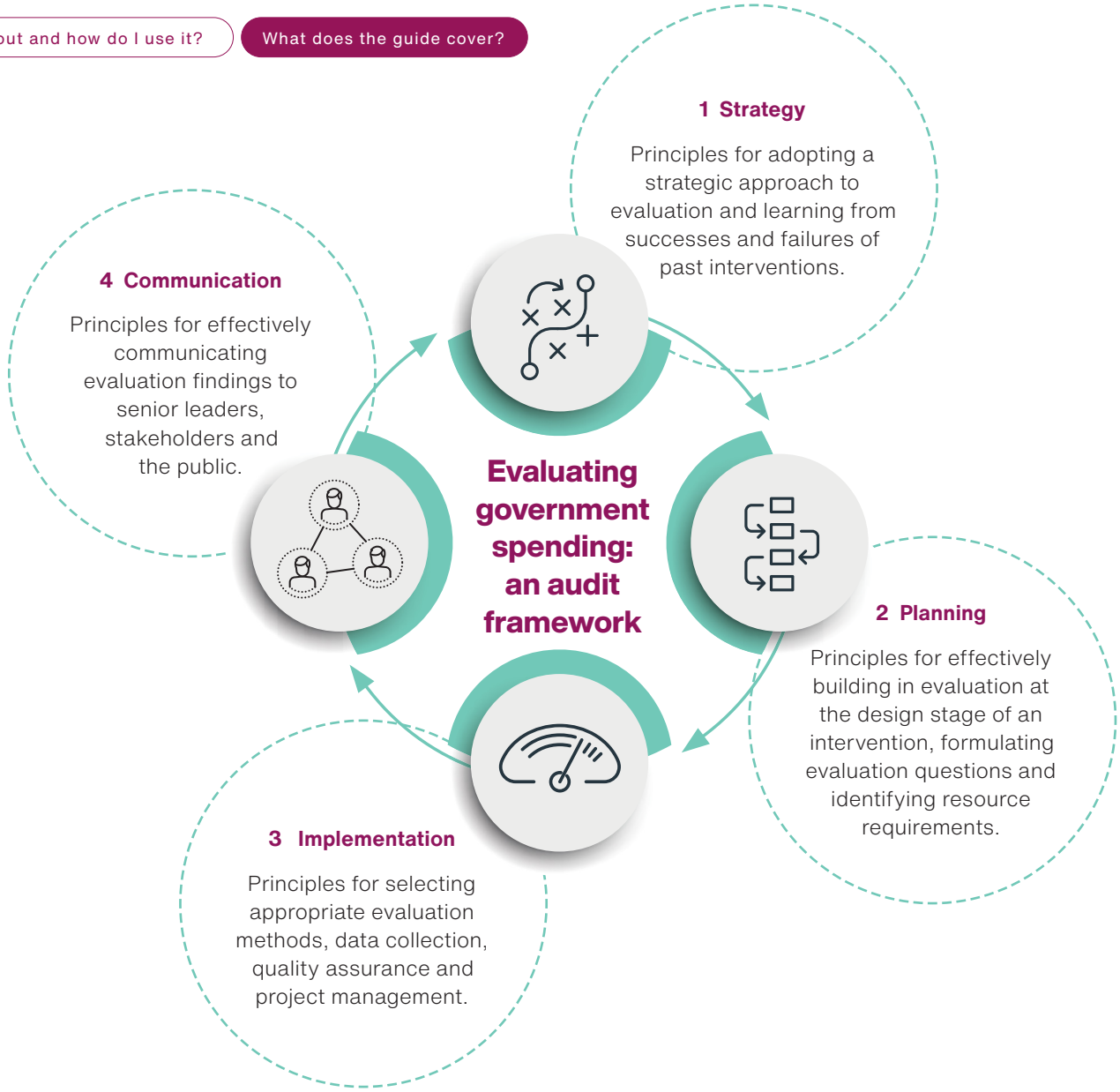
What does the guide cover?

This guide sets out a framework for how the NAO reviews the use of evaluation in policy-making and the robustness of the evaluation system that departments operate. It consists of four stages in the evaluation process: Strategy; Planning; Implementation; and Communication. The four stages form a cycle, as existing evaluation evidence is used to inform new interventions.

We set out key principles at each of these four stages in the evaluation process along with:

- a description of what good looks like and questions to consider; and
- case study examples.

The principles are grouped according to where they primarily apply. Some principles, such as quality assurance, are cross-cutting and are relevant to several stages of the evaluation process.



1 Strategy

Definition

When government considers a new intervention or a business-as-usual activity, it will consider what the intervention is intended to deliver and how it fits with wider government policies and objectives. Evidence from previous evaluations can inform the rationale behind an intervention, and evaluation of the intervention can help the department decide whether an intervention is delivering its objectives. Taking a strategic approach to evaluation can help departments to identify, prioritise and plan all evaluation activities across the department, including the allocation of limited evaluation resources.

Key principles

Align evaluation strategy with organisational strategic objectives

Understand the extent to which previous interventions have worked

Strategy

Key principles

Align evaluation strategy with organisational strategic objectives

Understand the extent to which previous interventions have worked

Align evaluation strategy with organisational strategic objectives

Departments need to set out their evaluation priorities and decide how to allocate evaluation resources between competing interventions. Considerations include the link to departmental priority outcomes, the risk or materiality of an intervention and the size of any gaps in the evidence base.

In general, we expect to see the following:

- A single evaluation strategy for the whole department. This approach can provide an overview of evaluation activities in the department and so help make sure that the most pressing evaluation gaps are prioritised, planned for and resourced.
  - ☑ Will the suite of planned evaluations provide sufficient, timely evidence of whether, why and how the department has achieved its priority outcomes?
  - ☑ Does the department plan to measure the impacts from both new interventions and business-as-usual activities?
  - ☑ Have any material gaps in analytical and evaluation capability and coverage been identified and addressed?
  - ☑ Is there an effective strategy to manage evaluations which are delivered in collaboration with other organisations?
  - ☑ Is the evaluation strategy reviewed regularly and progress of planned evaluation activities clearly communicated?
- Evidence that the department has considered whether its evaluation plans are proportionate, capable of providing robust evidence, and likely to produce findings within a useful timescale.
  - ☑ Is it clear why and how resources will be allocated to individual evaluations?
  - ☑ Are the analyst community and senior leaders clear about what must be done, why, when and by whom to deliver the strategy?
  - ☑ Does the department have the analytical capacity, capability, resources and support to carry out the planned programme of evaluation work?

Case study

Alignment of evaluation plans with departmental priorities

To improve its strategic approach to evaluation, in 2020 the Department for Levelling Up, Housing & Communities (DLUHC) set up a Monitoring and Evaluation Strategy Group which meets quarterly. The group's aims include developing evaluation plans and driving high standards of policy and programme evaluation across all policy areas. DLUHC identified its evaluation priorities and used this process to inform its planned evaluation activity, described in its Outcome Delivery Plan for 2021-22. DLUHC also set out further evaluation plans for individual policy areas such as housing.

Source: Comptroller and Auditor General, *Evaluating government spending,* Session 2021-22, HC 860, National Audit Office, December 2021

Understand the extent to which previous interventions have worked

The rationale behind an intervention should be informed by outputs and learnings from previous evaluations of what interventions have worked and why. Existing evaluation evidence will help in understanding what interventions can be used to achieve the intended outcome in the relevant UK context.

In general, we expect to see the following:

- The department maintains and regularly updates the knowledge base on the evidence that underpins interventions in key policy areas, collected through systematic reviews of existing evaluations or other robust methods.
  - ☑ Does the knowledge base draw on evidence generated by internal evaluations as well as external experts such as academics, other government departments, or practitioners in other countries?
  - ☑ Has the department documented its decisions where evidence from different sources is contradictory?
  - ☑ Does the department understand the extent to which the evidence is applicable to the relevant UK context?
- Senior leaders use lessons from previous evaluations to inform the evaluation strategy.
  - ☑ Does the evaluation strategy draw on evaluation evidence?
  - ☑ Does the department act on the results of evaluations whether findings were favourable or otherwise; for example, when making funding decisions?
- An assessment of the evidence from previous interventions that acknowledges negative findings or gaps in understanding.
  - ☑ Does the department have an effective process for identifying and acting on gaps in the evidence base that underpin its interventions?

Case study

Building the evidence base in a particular sector

In 2018, the Centre for Homelessness Impact, working with the Campbell Collaboration, created the Evidence and Gap Maps for Homelessness.[1] Evidence and Gap Maps are intended to provide quick and efficient tools to highlight what evidence exists for specific interventions and outcomes. An Evidence and Gap Map is a presentation of the available, relevant evidence for a particular sector. These maps direct users to this evidence to help inform strategy and programme development and provide the basis for evidence-based decision-making. Online interactive versions of the homelessness map are available at: EPPI-Mapper.

1 The Centre for Homelessness Impact is an independent What Works Centre, part of the What Works Network, which focuses on homelessness. More detail on the What Works Network is available at: www.gov.uk/guidance/what-works-network

2 Planning

Definition

The Green Book stipulates that monitoring and evaluation should be part of the development and planning of an intervention from the start. A clear evaluation planning process helps to make sure that all involved in the intervention understand their roles and responsibilities, that any major gaps are addressed, and that the practical arrangements for implementation of an evaluation are considered at an early stage in the policy cycle.

Key principles

Plan to incorporate evaluation throughout the policy cycle

Be clear about the questions to be answered

Consider evaluation requirements at the business planning stage, including the identification of resources needed

Engage with stakeholders in planning evaluations

Planning

Key principles

Plan to incorporate evaluation throughout the policy cycle

Be clear about the questions to be answered

Consider evaluation requirements at the business planning stage, including the identification of resources needed

Engage with stakeholders in planning evaluations

Plan to incorporate evaluation throughout the policy cycle

Evaluation is most useful when it informs thinking throughout the policy development cycle – before, during and after implementation. Early in the process, evaluation evidence may be used to help shape the design of an intervention and how it will be implemented. For example, piloting might be used to identify practical barriers or risks that might prevent the intervention from working as intended. Integrating evaluation throughout the policy-making process will help departments identify how the intervention might be modified to maximise its impact.

In general, we expect to see the following:

- Departments use evaluation evidence to help shape the design of policy options.
  - ☑ Are analysts and policy professionals sufficiently integrated to make sure that evaluation is incorporated into the policy-making process?
  - ☑ Do senior leaders have good access to lessons learned from previous evaluations?
- Departments use emerging evaluation evidence on the delivery of a new intervention or business-as-usual activity to inform decisions on improving, expanding or stopping the intervention.
  - ☑ Are evaluations aligned to key points in the policy cycle?
  - ☑ Has the department demonstrated how it plans to adjust policies in light of new evidence?
- Departments use evaluations to understand whether interventions work as intended, then use this evidence in future development of policy.
  - ☑ Have sufficient analytical resources been deployed to help senior leaders understand evaluation results?
  - ☑ Has the department conducted evaluation in cases where it is a mandatory requirement, such as regulatory policies subject to Post-Implementation Review (PIR); regulations containing a sunset or review clause; and requirements of the International Development Assistance Act 2015?

Case study

Building evaluation at the policy design stage

Our report on *Evaluating government spending* found that almost half of departments could fulfil the 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 in the 16 departments we surveyed agreed that the opportunity to learn was not being built into policy design and delivery.

Source: Comptroller and Auditor General, *Evaluating government spending*, Session 2021-22, HC 860, National Audit Office, December 2021

Planning

Key principles

Plan to incorporate evaluation throughout the policy cycle

Be clear about the questions to be answered

Consider evaluation requirements at the business planning stage, including the identification of resources needed

Engage with stakeholders in planning evaluations

Be clear about the questions to be answered

There are three main types of evaluation: process, impact and economic evaluation. For a full understanding of whether an intervention worked, how, why and for whom, and at what cost, all three types of evaluation will be needed. Clear evaluation questions can help make sure that the evaluation is designed to provide useful insights.

In general, we expect to see the following:

- Clear links between evaluation plans for an intervention and the overall departmental evaluation strategy.
  - ☑ Does the department have a good understanding of what it aims to achieve from its evaluation?
  - ☑ Is it clear why the evaluation approach was chosen?
- Clear evaluation questions informed by the intended outcomes of the intervention and the theory of change for how the intervention is expected to impact outcomes.
- Interventions designed with evaluation in mind (such as the use of pilot schemes or a phased roll-out).
  - ☑ Does the design of the intervention provide opportunities to learn? For example, is it feasible to pilot the intervention, and can such a pilot be effectively evaluated?
- Fit-for-purpose evaluation that is proportionate and reflects the needs of senior leaders and scrutiny bodies.
  - ☑ Is the evaluation designed so that findings will be available to support the planned decision points for the intervention?
  - ☑ Do senior leaders understand the limitations of using the evaluation methods chosen?

Case study

Failure to identify evaluation outcomes

In 2020, we noted that HM Treasury, when designing tax expenditures, had not given enough consideration to how it will measure impact. We found that government had failed since 2013 to plan for the evaluation of tax expenditure measures either at design stage, or where costs or benefits differed significantly from forecasts. Consequently, Parliament was not provided with the information necessary to assess the value for money of tax expenditures. We found that other countries had more comprehensive evaluation and reporting despite most having comparatively lower levels of tax expenditures.

Source: Comptroller and Auditor General, *The management of tax expenditures,* Session 2019-20, HC 46, National Audit Office, February 2020

Planning

Key principles

Plan to incorporate evaluation throughout the policy cycle

Be clear about the questions to be answered

Consider evaluation requirements at the business planning stage, including the identification of resources needed

Engage with stakeholders in planning evaluations

Consider evaluation requirements at the business planning stage, including the identification of resources needed

Early consideration of evaluation requirements such as resourcing and skills will help make sure that planned evaluations can be completed to the appropriate standard.

In general, we expect to see the following:

- Scrutiny of evaluation plans at the investment approval stage.
  - ☑ Is the evaluation plan appropriately funded and scrutinised as part of internal approval processes?
  - ☑ Does the department take appropriate action, including seeking a ministerial direction where necessary, where evaluation is not feasible or where there is a risk of not securing value for money?
- Clearly defined processes, responsibilities and governance arrangements for evaluations.
  - ☑ Do the department's arrangements for the approval and quality assurance of the evaluation help promote public trust and accountability?
- When deciding to undertake an evaluation, departments consider how to balance the robustness, scale and cost of evaluation activities with the size, complexity and timeframe of the intervention.
  - ☑ Are evaluation plans proportionate to the size, complexity and risk of the intervention?
- Departments identifying early the resources they need to carry out proportionate evaluations of interventions.
  - ☑ Are there sufficient analytical and other skills and resources available to carry out a proportionate evaluation of the intervention?

Case study

Evaluation planning at an early stage in policy life-cycle

HM Treasury used the 2020 Spending Review to strengthen the link between funding decisions and the strength of evidence and robustness of evaluation plans. It required departments to set out the evidence base and evaluation plans for their interventions and awarded additional funding to programmes with thorough plans for evaluation. HM Treasury decided to continue this approach in the 2021 Spending Review to further incentivise departments to increase the use and provision of evaluation of their interventions.

Source: Comptroller and Auditor General, *Evaluating government spending,* Session 2021-22, HC 860, National Audit Office, December 2021

Engage with stakeholders in planning evaluations

Good policy development is likely to include extensive collaboration with a wide range of stakeholders, including designers, implementers, beneficiaries and interest groups, to understand how the intervention will work from a range of perspectives. External stakeholder engagement helps in tailoring the evaluation to meet user needs.

In general, we expect to see the following:

- Comprehensive stakeholder engagement plans for the whole duration of evaluation activities, including for bodies that sit outside government.
  - ☑ Has there been meaningful engagement with key stakeholders to help ensure evaluation activities are relevant and useful?
  - ☑ Are stakeholders involved in the design and planning of evaluation activities, including any external evaluators that will carry out the evaluation?
  - ☑ Are there sufficient safeguards in place to prevent stakeholders from exerting undue influence on the planning and implementation of evaluation activities, thereby compromising the integrity of the evaluations?
- Departments considering the likely impacts of interventions on different segments of the population.
  - ☑ Do evaluations measure how interventions and business-as-usual activities affect different demographic groups or geographic areas?

Case study

Standards for stakeholder engagement

The OECD's 2020 report on *Improving governance with policy evaluation* said that governments were increasingly eager to engage a wide range of internal and external stakeholders in the decision-making process to generate a broader consensus and increase the legitimacy of public-policy decisions. In the Netherlands, for example, the Ministry of Finance's *Regulations for periodic evaluation research* (March 2018) laid down rules for the participation of stakeholders in periodic evaluations. With each policy evaluation, at least one independent expert must give an opinion on the quality of the evaluation. Similarly, the European Commission's *Better Regulation Guidelines* (November 2021) contains a chapter that describes standards for stakeholder engagement. According to these guidelines, views from stakeholders should be included in the evaluation of all programmes and policies issued by the Commission, as well as initiatives with impact assessments.

Source: Organisation for Economic Co-operation and Development, *Improving governance with policy evaluation: lessons from country experiences*, June 2020

3 Implementation

Definition

Evaluation design and implementation involves prioritisation based on timelines, resources and the feasibility of methods, and setting up good project management arrangements. Robust baseline measures should be identified at an early stage, ensuring that intervention progress can be tracked from a suitable starting point.

Key principles

- Pick the most appropriate evaluation method
- Assess and collect data
- Build in and carry out quality assurance throughout the evaluation
- Keep the evaluation on track using good project management

Pick the most appropriate evaluation method

Many methods can be used to explore whether an intervention has the desired impact. The selection of methods should be informed by the theory of change and the uncertainties and assumptions identified. A robust impact evaluation will need a clear counterfactual to assess what would have happened in the absence of the intervention. It may be important to consider if the intervention has differential impacts across the population. These considerations may influence the evaluation methods chosen.

In general, we expect to see the following:

- A choice of evaluation method(s) that reflects the theory of change on which the intervention is based (in the case of new interventions).
  - ☑ Will the chosen evaluation methods provide good insights into whether the intervention works as expected?
- Resources appropriate for the methods chosen.
  - ☑ Have sufficient analytical and other resources (time, money, or skills) been allocated to deliver the methods chosen for the evaluation?
- Choice of high-quality and appropriate outcome measures to quantify impact.
  - ☑ Does the department consider which outcome measures are likely to best answer the evaluation questions?

Case study

The importance of a counterfactual

The NAO's 2020 report on electricity networks looked at support that Ofgem gave electricity network companies to engage in technological and commercial innovation. An Ofgem-commissioned evaluation of one such scheme found that it delivered significant cost savings for consumers and reductions in CO_2 emissions. However, this evaluation was limited. Ofgem did not ask the evaluators to provide a robust estimate of how much innovation might have taken place in the absence of the innovation support (in other words, the counterfactual). This type of evaluation is always difficult, but evaluators could have been required to take a more advanced approach to estimation than they did.

Source: Comptroller and Auditor General, *Electricity networks,* Session 2019-20, HC 42, National Audit Office, January 2020

Successful use of randomised control trials

The NAO's 2021 report on *Managing tax debt through the pandemic* found that HM Revenue & Customs (HMRC) had been strong in using and learning from behavioural insights. It has been conducting randomised control trials on debt management since 2012 to improve the impact of its interventions.

Source: Comptroller and Auditor General, *Managing tax debt through the pandemic,* Session 2021-22, HC 799, National Audit Office, November 2021

Assess and collect data

An evaluation may be impossible, limited or more costly in the absence of good plans for data collection. Data are critical to measuring changes in outcomes over time due to an intervention.

In general, we expect to see the following:

- Consideration of: the evaluation questions to be answered; which bodies can provide relevant data; and data access constraints.
  - ☑ Have existing data from administrative and monitoring systems, or large-scale, long-term surveys been considered as sources?
  - ☑ Are the data management arrangements sufficiently robust and effective for collection and recording of data and their secure storage?
- Building in pre-planned data quality checks.
  - ☑ Does the department have reliable and robust data quality assurance arrangements to provide assurance over the findings of its evaluations?
  - ☑ Has the department considered any bias in its data collection and taken steps to minimise it?
- Linking of different data, where possible, to create richer datasets, improve the quality of data and avoid duplication of data collection.
  - ☑ Is it possible to combine qualitative and quantitative data to balance the limitations of one type of data with the strengths of another?

Case study

Using common metrics across organisations

In 2013, we found that the Low Carbon Innovation Coordination Group, a group that coordinated the activities of departments and bodies with low carbon innovation interests, piloted and evaluated a set of output and outcome metrics. The Group developed seven common metrics on outputs and outcomes linked to capacity-building, energy policy goals and economic benefits. The Group also piloted a template to capture qualitative information on design and delivery of projects to help share good practice and lessons learned between members of the Group and to inform future activities.

Source: National Audit Office briefing for the Energy and Climate Change Select Committee, *Public funding for innovation in low carbon technologies in the UK*, 2013

Monitoring data used for decision-making

In 2021, the Department for Business, Energy & Industrial Strategy told us it cancelled the Green Homes Grant after monitoring information that the Department was collecting for its planned evaluation showed the intervention was not delivering as intended.

Source: Comptroller and Auditor General, *Evaluating government spending,* Session 2021-22, HC 860, National Audit Office, December 2021

Implementation

Key principles

Pick the most appropriate evaluation method

Assess and collect data

Build in and carry out quality assurance throughout the evaluation

Keep the evaluation on track using good project management

Build in and carry out quality assurance throughout the evaluation

Quality assurance of all aspects of an evaluation, including design, method(s) selection, data collection and reporting will help make sure that evaluation outputs are robust and fit for purpose. Feedback on quality by independent reviewers can improve the quality and impact of the evaluation and make sure that benefits found by the evaluation are valid.

In general, we expect to see the following:

- Quality assurance steps at each stage of the evaluation.
  - ☑ Are the quality assurance arrangements sufficient to provide confidence in the quality of the evaluation and do they comply with relevant guidance and standards?
  - ☑ Are there material gaps in the quality assurance process for the evaluation?
- Peer reviews by experts independent of the intervention and its evaluation to assess the appropriateness of evaluation questions, outcome measures and the robustness of the chosen methods and findings.
  - ☑ Have the evaluation questions, design, methods, data collection methods and findings been quality-assured by an independent expert?
- Implementation of peer review recommendations to improve the quality and relevance of the evaluation.
  - ☑ Are there sufficient opportunities for recommendations to be integrated into the evaluation?

Case study

Challenge and scrutiny provided by internal experts

The Department for Levelling Up, Housing & Communities (DLUHC) requires all proposals for commissioned research, above a minimum monetary threshold, to go through a quality gateway, comprising a panel of experts from across the department, to provide challenge and scrutiny. The gateway panel comprises the chief analyst, senior analysts from each of the professions and representatives from commercial, digital, data protection and central finance teams. The panel assesses bids in line with a number of criteria, such as whether projects have the necessary ministerial and financial approvals, and the research methodology is robust and follows relevant ethical procedures.
The panel also looks at whether the proposed work is in line with departmental priorities, procurement is as efficient and cost-effective as possible, and alternative sources of funding have been considered, such as co-funding between departments. The gateway team can also ask for regular updates on progress, and feedback of interim findings, which acts as a further quality check alongside the project management and steering group arrangements set up for individual projects.

Source: Comptroller and Auditor General, *Evaluating government spending*, Session 2021-22, HC 860, National Audit Office, December 2021

Implementation

Key principles

Pick the most appropriate evaluation method

Assess and collect data

Build in and carry out quality assurance throughout the evaluation

Keep the evaluation on track using good project management

Keep the evaluation on track using good project management

The skills and principles for managing an evaluation are the same that would apply for other projects of a similar size. Planning is critical to get the programme management right: our experience shows that project success is strongly linked to the quality of project initiation. Additional skills are needed if an evaluation is commissioned from an external provider. Our Good practice guidance: managing the commercial lifecycle provides advice on good practice in contract management.

In general, we expect to see the following:

- Clear and proportionate governance and assurance structures that provide effective oversight and challenge.
  - ☑ Does the department have a strong oversight system that can challenge the design and delivery of evaluation projects?
- An evaluation leader who takes personal ownership of the project, provides clear direction and has the necessary authority and influence.
  - ☑ Is a named individual responsible for successful delivery of the evaluation? Are other roles and responsibilities defined?
- A project plan that identifies the enablers required to achieve the objective (for example, people, funding, data access, partners).
  - ☑ Are the evaluation plans realistic and achievable with the resources currently available, and are they in place? Are monitoring arrangements in place?
- A risk management strategy which highlights challenges to the success of the project, suggests ways to mitigate the risks and monitors residual risks.
  - ☑ Is there an active risk management plan that helps to anticipate and mitigate emerging risks?
  - ☑ Does the department act effectively to address risks and barriers to a successful evaluation?

Case study

Key questions to ask when reviewing major programmes

The NAO Framework to review programmes sets out how we consider major programmes and shares our learning from recent work. It draws from our experience of around 200 studies reviewing public sector programmes since 2010.

It sets out the 18 key questions we ask when we review major programmes – grouped into the following four elements:

Purpose – Is there a strategic need for the programme and is this the right programme to meet the business need?

Value – Does the programme provide value for money?

Set-up – Is the programme set up in accordance with good practice and are risks being well managed?

Delivery and variation management – Are mechanisms in place to deliver the intended outcomes and respond to change, and is the programme progressing according to plan?

Source: National Audit Office, *Framework to review programmes*, April 2021

4 Communication

Definition

To maximise the impact of an evaluation, findings need to be clearly communicated to relevant stakeholders. Keeping senior leaders well informed of evaluation findings makes it more likely they will understand the impact of an intervention and consider potential improvements. Incorporating the latest evaluation evidence into the existing stock of knowledge can help inform the design of future interventions. Departments should be transparent about the outcomes of their interventions and publish evaluation outputs in a timely manner to support accountability.

Key principles

Communicate evaluation findings to senior leaders at key decision points

Publish and share findings with a suitable range of stakeholders in a timely manner

Collate evaluation findings and lessons learned so that they can be easily accessed and shared

Communication

Key principles

- Communicate evaluation findings to senior leaders at key decision points
- Publish and share findings with a suitable range of stakeholders in a timely manner
- Collate evaluation findings and lessons learned so they can be easily accessed and shared

Communicate evaluation findings to senior leaders at key decision points

Communicating evaluation findings to senior leaders in a timely manner will help inform decisions about whether to continue, amend or stop an intervention. The centre of government expects that departments will report and act on the results of evaluations.

In general, we expect to see the following:

- A process that makes sure that emerging evaluation findings are shared with senior leaders on a regular basis.
  - ☑ Is it clear when evaluation evidence will be available and the implications for the intervention's design and implementation?
  - ☑ Are there sufficient opportunities to consider evaluation findings and assess whether an intervention needs to be amended, expanded or stopped?
  - ☑ Is there close collaboration between analytical and policy teams to help make sure that evaluation evidence is well understood by senior leaders?
- A governance structure for an intervention that includes a senior official who will oversee the evaluation.
  - ☑ Is the senior official responsible for the delivery of the evaluation part of the governance structure of the intervention?
- A record of the decisions and actions taken in response to evaluation findings.
  - ☑ Is there a comprehensive record of decisions taken in response to evaluation evidence, regardless of whether the findings were favourable or not?

Case study

Regular reporting of evaluation findings to senior leaders

The Department for Levelling up, Housing & Communities held regular meetings with senior leaders for Supporting Families (previously Troubled Families) to ensure that the evaluation of the programme met their needs.
The Department presented findings at regular intervals throughout the evaluation with the view that they would feed into policy decisions. Given that the implementation of the programme relied on local authorities, analysts also worked closely with policy teams to ensure that key findings from the evaluation were fed back to local authorities.

Source: Comptroller and Auditor General, *Evaluating government spending,* Session 2021-22, HC 860, National Audit Office, December 2021

Publish and share findings with a suitable range of stakeholders in a timely manner

The Magenta Book and the Government Social Research Publication Protocol make it clear that government is expected to make evaluation outputs publicly available. Publishing evaluation outputs demonstrates a commitment to transparency and enables external scrutiny and learning. Sharing evaluation findings with stakeholders will help raise the profile of the findings, and can help to improve outcomes.

In general, we expect to see the following:

- Timely publication of an objective and comprehensive evaluation report in line with the requirements set out in the Government Social Research Publication Protocol and any conditions set by HM Treasury.
  - ☑ Does the department publish high-quality evaluation reports in a timely manner?
  - ☑ Is the department's published evaluation in line with the evaluation plan; for example, as set out as part of the business case for the intervention?
- Inclusion of technical appendices containing underlying data and a detailed description of the methodology supporting the findings.
  - ☑ Does the evaluation report clearly communicate the methodology and limitations of the evaluation?
- Departments share evaluation findings across different communication channels to reach a wide range of stakeholders.
  - ☑ Are evaluation findings communicated in ways that are useful to the various different types of end-users, and are the implications for good practice clear?

Case study

Public repository of evaluations in Norway

The Norwegian government maintains an evaluation portal that collects a selection of evaluations carried out across a range of departments and state enterprises since 2005. The portal can be accessed by government officials as well as the general public. Users can search evaluations by policy area, department and type of evaluation. The portal is available online at: https://kudos.dfo.no/

Collate evaluation findings and lessons learnt so they can be easily accessed and shared

Findings from new evaluations expand departments' existing body of evidence. This helps them apply the learning to subsequent policies and can help avoid repeating past mistakes. Easy access to information on what works can save time and help ensure that the strategic case for future interventions is stronger. Where applicable, departments may carry out or commission (for example, from What Works Centres) systematic reviews on the effectiveness of interventions in their policy areas to develop good understanding of what has and has not worked.

In general, we expect to see the following:

- ‘Lessons learned’ exercises where lessons about an intervention are identified and captured, including for interventions that were not successful in achieving their objectives.
  - ☑ Has the department carried out a comprehensive ‘lessons learned’ exercise to identify what has worked well and what has not in relation to the intervention and the evaluation?
- An assessment of the strengths and weaknesses of the evaluation design and challenges in implementing the evaluation.
- A central repository of evaluation evidence and similar types of evidence produced by the department and by third parties commissioned to perform evaluations.
  - ☑ Is the department's evidence base updated as new evidence emerges?
  - ☑ Is knowledge about interventions managed in a way that enables the sharing of learning across policy areas?
  - ☑ Is the evidence base in the department accessible to analysts and policy professionals?

Case study

The value of a central evaluation repository

In 2018, the Department for Business, Energy & Industrial Strategy set up a central analysis, monitoring and evaluation database to share learning across the department. The database brings together impact assessments, post-implementation reviews, business cases and evaluations. This allows for comparisons across its policies and improved implementation of new schemes by applying lessons from past interventions. The database is also intended to provide more accurate estimates of benefits and costs when setting up new schemes.

Source: Comptroller and Auditor General, *Evaluating government spending,* Session 2021-22, HC 860, National Audit Office, December 2021

Evidence synthesis and What Works Centres

What Works Centres collate existing evidence on the effectiveness of programmes, and produce synthesis reports and systematic reviews across several policy areas. 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.

Glossary

Business-as-usual activities or activity	The normal, expected operations of government in contrast to any projects and interventions associated with change.
Evaluation capability	The analytical and other resources available to carry out high-quality evaluations.
Evaluation strategy	A high-level plan that typically shows which activities the organisation will evaluate, how it will do so, and with what resources.
Intervention(s)	A programme, policy or activity to achieve a desired outcome.
Outcome measure	A standardised way of measuring the positive or negative effect of an intervention.
Randomised control trials	An experiment which aims to reduce bias when testing a new intervention. The people participating in the trial are randomly allocated to either i) the group receiving the new intervention being tested or ii) a control group that does not receive the new intervention.
Scrutiny bodies	Parliamentary committees and other bodies outside the department who are interested in how it performs.
Theory of change	Captures the theory of how the intervention is expected to work (setting out all the steps expected to be involved in achieving the desired outcomes), the assumptions made, the quality and strength of the evidence supporting them, and wider contextual factors.