Government Major Projects Evaluation Review

Evaluation Task Force

April 2025

Ministerial Foreword

The Government Major Projects Portfolio comprises the most complex and strategically significant projects and programmes delivered by the government. Because these projects are so large and complex, it's important to evaluate them properly to understand whether they are effective or

Georgia Gould OBE MP
Parliamentary Secretary for the Cabinet Office

not, learn lessons and ensure the best value for money. Evaluation is fundamental to an outcome-focussed and more accountable approach that is needed to deliver on our mission to rebuild Britain.

I therefore welcome this review, conducted by the Evaluation Task Force, into evaluation in the Government Major Projects Portfolio. The review shows that 63% of major projects now have some form of evaluation plan and 34% have a high-quality evaluation in place. This represents significant progress from the past, but there is still a need for considerable further improvement which will require a cross-government effort.

I am pleased therefore to publish this review alongside an action plan that will embed evaluation into all stages of the major project lifecycle, so that evaluation becomes a 'must do' rather than a 'nice to have'.

Georgia Gould MP
Parliamentary Secretary for the Cabinet Office

Executive Summary

The Government Major Projects Portfolio (GMPP) is a collection of the government's most complex and strategically significant projects. Major projects range from railways, roads, schools, hospitals and housing, to energy, telecommunications, defence, IT and transformation programmes. The 2023-24 GMPP comprises 227 projects with a total cost of £834 billion.

Because these projects are so large and complex, it is important to evaluate them properly to learn from them and hold the government accountable for their delivery. However, past evidence has shown that good evaluation does not happen enough for major projects. A review in 2019 found that only 8% of spending on the GMPP at that time had suitable impact evaluation plans in place.

Since then, the number and total cost of projects in the GMPP has grown significantly, and steps have been taken to improve evaluation standards across government, including the creation of the Evaluation Task Force (ETF) in 2021.

The ETF has now conducted a new review to see how well current GMPP projects are being evaluated and what needs to be done to keep improving.

Findings

The findings from this review show that while the coverage of major project evaluation has increased significantly since 2019, there is still considerable further improvement required. Of the 227 projects that are part of the GMPP in 2023-24:

- **A total of 144 projects (63% of the GMPP) provided evidence of some form of evaluation plans for the review.** These ranged from mature evaluations that are already being implemented to those at an early stage of development.

- **A third (34%) of all GMPP projects** (representing £378 billion in total cost) **were found to have good quality evaluation plans in place**.

- **The remaining two thirds (66%) of all GMPP projects** (representing £456 billion in total cost) **did not provide evidence of good quality evaluation plans**, meaning that either no evaluation plan was provided, or existing plans were not found to meet the criteria for quality applied in this review.

This means that there is still a significant gap in the coverage of suitable evaluation plans across the GMPP. The review identified several reasons for this gap, which fall into three categories:

- **Operational challenges**: Some projects did not leave enough time to design evaluations early on, making it hard to estimate their impact later. Others had trouble accessing or generating the data they needed.

- **Cultural challenges**: Some senior staff and decision-makers view evaluation as a 'luxury' or lower priority, rather than a necessary and valuable part of the delivery of major projects.

- **Resourcing challenges**: Some projects did not secure the necessary resources - like staff, funding, and expertise - to carry out good quality evaluations.

The findings from this review have informed an action plan to overcome these barriers and continue improving major project evaluation over time. The action plan will be taken forward by the ETF, HM Treasury (HMT) and the National Infrastructure and Service Transformation Authority (NISTA), working with government departments and major project teams. The action plan has four components:

- **Evaluation governance:** improve governance to make having an appropriate evaluation plan a requirement for HMT spending approval in the Treasury Approval Process, and to monitor the ongoing design and delivery of the evaluation through the major project gateway process. The ETF will also work with departments to ensure these requirements are mirrored in internal governance processes within departments.

- **Evaluation capability**: improve evaluation skills across GMPP project teams and NISTA assurance teams through developing and sharing examples of best practice and tailored training resources.

- **Embedding evaluation into project delivery:** set out evaluation requirements for project delivery teams in the Teal Book, and monitor compliance with requirements by using the Evaluation Registry, which is an online repository of all planned, ongoing and complete UK government evaluations.

- **Promoting the value of evaluation**: engage major project stakeholders across government, including evaluation and project delivery teams in departments, to take forward opportunities to promote evaluation and its value throughout the project lifecycle.

Contents

Introduction

This section provides context on evaluation in the Government Major Projects Portfolio (GMPP), the aims of this review and the approach taken.

The Government Major Projects Portfolio

The GMPP is a collection of the government's most complex and strategically significant projects. Projects are typically considered 'major projects' if they require primary legislation or approval from HM Treasury (HMT), either because the budget exceeds a department's delegated authority level and/or because the project is novel or complex. While the GMPP spans many of the government's most high profile projects, it represents only a portion of the projects delivered across government.[1]

Projects on the GMPP are independently scrutinised by the National Infrastructure and Service Transformation Authority (NISTA) who provide advice and monitor delivery progress.[2] Major projects range from railways, roads, schools, hospitals and housing, to energy, telecommunications, defence, IT and transformation programmes. The 2023-24 snapshot of the GMPP comprises 227 projects with a total cost of £834 billion.[3]

Evaluation is research to understand the outcomes of a project, including whether it worked as intended, how it was delivered in practice and whether it represented good value-for-money. Evaluation enables *learning* by providing the evidence on what is and is not working well, and why. This evidence allows projects to refine their delivery and inform the design of future projects addressing similar objectives. Evaluation also provides *accountability,* to show whether or not projects are delivering against what they set out to achieve.

The nature of GMPP projects means that it is essential that these projects are evaluated proportionately and that learning is shared. Evaluation is needed to inform the ongoing delivery of these projects, and future decisions about whether projects should be continued, modified or stopped. Major projects are therefore required to implement robust and proportionate evaluation in line with the government's Magenta Book evaluation guidance.[4]

However, past evidence has shown that high quality evaluation does not happen enough for major projects. A review of major project evaluation carried out by the Prime Minister's

---

[1] Infrastructure and Projects Authority Annual Report 2023-24 (PDF, 1.91 MB).

[2] At the time this review was carried out, NISTA had not yet been created. NISTA was established in April 2025 to bring together the Infrastructure and Projects Authority (IPA) and National Infrastructure Commission (NIC). This review will refer to the IPA when discussing the review methods and findings that were completed before NISTA was formed. The final section of this report (the Action Plan) discusses what will happen in the future, and will therefore refer to NISTA rather than the IPA.

[3] Infrastructure and Projects Authority Annual Report 2023-24 (PDF, 1.91 MB). By total cost, we refer here to the 'whole life cost' of major projects, which is defined as the total cost of each project over its entire life, from planning and design through to delivery, ongoing maintenance and eventual closure.

[4] Magenta Book (link last accessed 7th February 2025)

Implementation Unit (PMIU) in 2019 found that only 8% of the spend on major projects in the GMPP at that time had high quality impact evaluation plans in place.[5]

The aims of this review

This review set out to answer three questions:

1. What is the coverage and quality of evaluation for the current GMPP?

2. What are the barriers to evaluating major projects?

3. What actions are needed to drive continuous improvement in evaluation of major projects?

This review has been led by the Evaluation Task Force (ETF), a joint HMT and Cabinet Office unit set up to put evaluation evidence at the heart of government decisions. The review has been undertaken in partnership with the Infrastructure and Projects Authority (IPA - now part of NISTA) and HMT. The ETF engaged regularly with central evaluation teams in the relevant government departments throughout the review.

Approach and evidence used

This review has been undertaken over two phases as set out below.

The first phase ('Phase 1') was conducted with Ipsos and Ecorys in 2023. This involved a survey of major projects to understand their evaluation plans, a review of submitted evaluation plans, and qualitative interviews with a sample of major project teams. A methodology was developed to assess the quality of major project evaluation plans, with interim findings and recommendations provided.

In April 2024, the ETF launched a second phase of the review ('Phase 2'). This was undertaken to get an updated picture of evaluation coverage and quality for existing major projects and to give new projects a chance to share their plans. Senior officials in HMT and IPA wrote to relevant Senior Responsible Owners (SROs) asking for evaluation plans to be submitted. The major projects in scope were either projects assessed as not having robust impact evaluation plans in Phase 1 of the review, or projects which joined the GMPP after the cut-off point for inclusion in the Phase 1 review (after March 2022). These requests provided two deadlines to submit evaluation plans: an initial deadline in the Spring of 2024, and a final deadline of December 2024. The ETF reviewed all plans submitted during this phase, following the same assessment framework as during Phase 1 with some minor refinements. Full detail on the Phase 2 methodology is outlined in Annex A.

Across both phases of this review, 'robust' plans are defined as those with a suitable evaluation approach to address the evaluation aims and questions in the context of the project, taking into consideration relevant process, impact and value for money evaluation methods.[6]

---

[5] Report - Evaluating government spending (.pdf, 470 KB)

[6] For impact evaluations, this requires evidence of experimental, quasi-experimental or theory-based methods that are appropriate and proportionate in the context of the project. In contrast to the 2019 review, for this review we included projects with a well-designed theory-based evaluation in the

Findings in this report are based on a combination of Phase 1 and Phase 2 assessments of the 2023-24 GMPP, which comprises 227 projects. This means that these results apply to a snapshot of the GMPP at a particular point in time.

Structure of this report

This report is divided into the following three sections:

1. **Coverage and quality of evaluation in GMPP:** This section reports on the overall proportion of the GMPP with robust evaluation plans. It also presents figures showing the coverage and quality of evaluation across different categories of projects and types of evaluation.

2. **Barriers to evaluating major projects:** This section describes challenges that were identified affecting the scale and quality of major project evaluation.

3. **Action plan:** This section sets out the joint action plan being put in place by the ETF, HMT and NISTA to drive continuous improvement in the evaluation of major projects. This includes steps to improve governance, enhance evaluation capabilities and better embed evaluation into project delivery.

The ETF would like to thank all the major project teams and stakeholders across government that have contributed evidence and insights to this review.

---

definition of robustness, recognising that these approaches are often suitable to the context of major projects. Further details are provided in Annex A.

Section 1   Coverage and quality of evaluation across the GMPP

This section outlines the overall proportion of the GMPP with robust evaluation plans. It also presents figures showing the coverage and quality of evaluation across different categories of projects and types of evaluation.

The proportion of the GMPP with robust evaluation plans in place

Since the previous review in 2019, there has been significant improvement in the coverage and quality of evaluation for the GMPP but further improvement is still required.

A total of 144 projects (63% of the GMPP) provided evidence of some form of evaluation plans for the review. These ranged from mature evaluations that are already being implemented to those at an early stage of development.

Overall, a third (34%) of all GMPP projects were assessed as having robust evaluation plans in 2024, and two thirds (66%) did not.

The gap in evaluation for major projects reflects issues of both coverage and quality. Of the 66% of projects categorised as not having evidence of robust evaluation:

- 45% of these projects did have an evaluation in place but it did not meet the criteria to be judged as robust. In some cases, this includes projects that were at an early stage at the time of this review and still developing their plans for evaluation. This means that these projects may still develop robust evaluation plans as they progress with their work.

- 55% of these projects did not provide any evidence that an evaluation plan is currently in place.

Evaluation coverage is typically better for the higher cost projects. The 34% of projects with robust evaluation plans represent £378 billion in total cost, 45% of the total cost of the GMPP. The 66% of projects that did not provide evidence of robust evaluation plans represent a total cost of £456 billion.

Figure 1: A third (34%) of GMPP projects have robust evaluation plans in place

Percentage and total cost of GMPP projects, by evaluation plan status, UK, 2023 to 2024
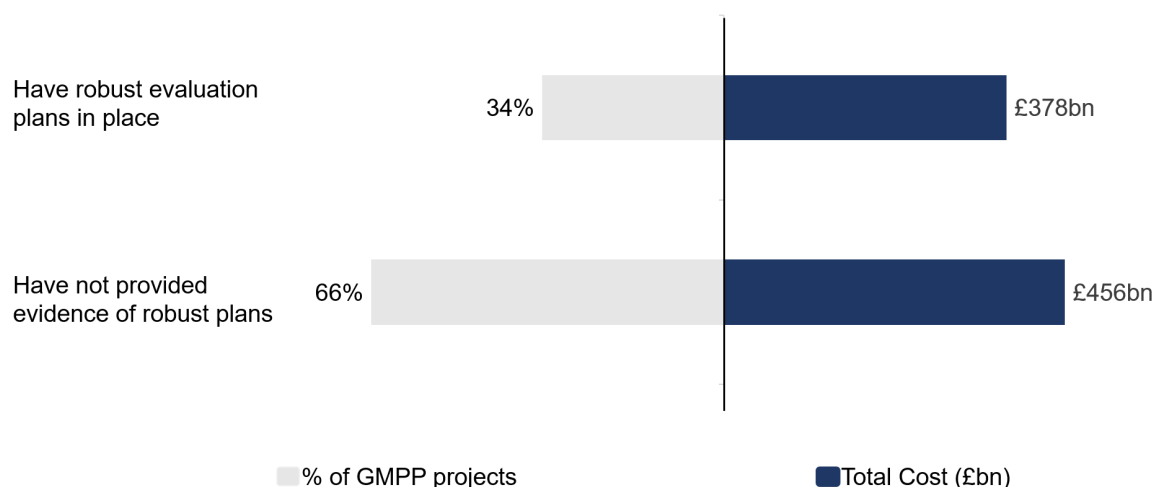
Source: ETF assessment of evaluation plan quality; cost data from IPA annual report 2023-24

Notes:
1. The total number of projects is 227.
2. How to interpret this figure: This figure shows the proportion of major projects that were assessed to have robust evaluation plans in place or not. For example, 34% of the GMPP were assessed to have robust evaluation plans in place, representing £378bn of spend.

Variation in evaluation coverage and quality between projects

Differences in evaluation coverage and quality were identified depending on the type of project and type of evaluation.

Coverage and quality of evaluation plans by IPA annual report category

GMPP projects fall into one of four IPA categories, according to their purpose and the nature of delivery. These categories are:
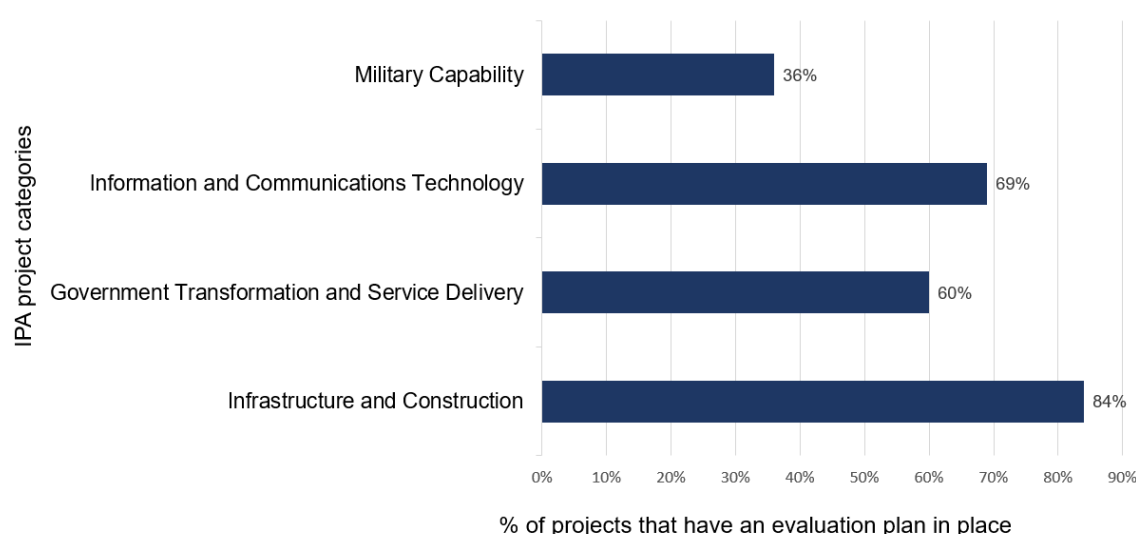
- Infrastructure and Construction
- Government Transformation and Service Delivery
- Information and Communication Technology
- Military Capability[7]

---

[7]The four categories are defined as follows: (1) Infrastructure and Construction projects involve constructing new buildings or improvements to the UK's energy, environment, transport, telecommunications, sewage and water system. (2) Government Transformation and Service Delivery Projects are those that aim to modernise government operations and improve how services are delivered to the public. (3) Military Capability projects deliver new Government military equipment. (4) Information and Communication Technology projects involve modernising technology to reduce costs and provide better access to services. For more information please see the Infrastructure and Projects Authority Annual Report 2023-24 (PDF, 1.91 MB).

Infrastructure and Construction projects are more likely to have an evaluation plan in place than other types of project. As shown in Figure 2 below, 84% of Infrastructure and Construction projects have an evaluation plan in place compared to 69% of Information, Communication and Technology projects; 60% of Government Transformation and Service Delivery projects; and 36% of Military Capability projects.

Figure 2: A higher proportion of Infrastructure and Construction major projects have any evaluation plan in place at all, compared to other types of major project

Percentage of GMPP projects with any evaluation plan in place at all, by IPA category, UK, 2023 to 2024

% of projects that have an evaluation plan in place

Source: Existence of an evaluation plan from ETF data; IPA annual report categories from IPA annual report (2023-24)
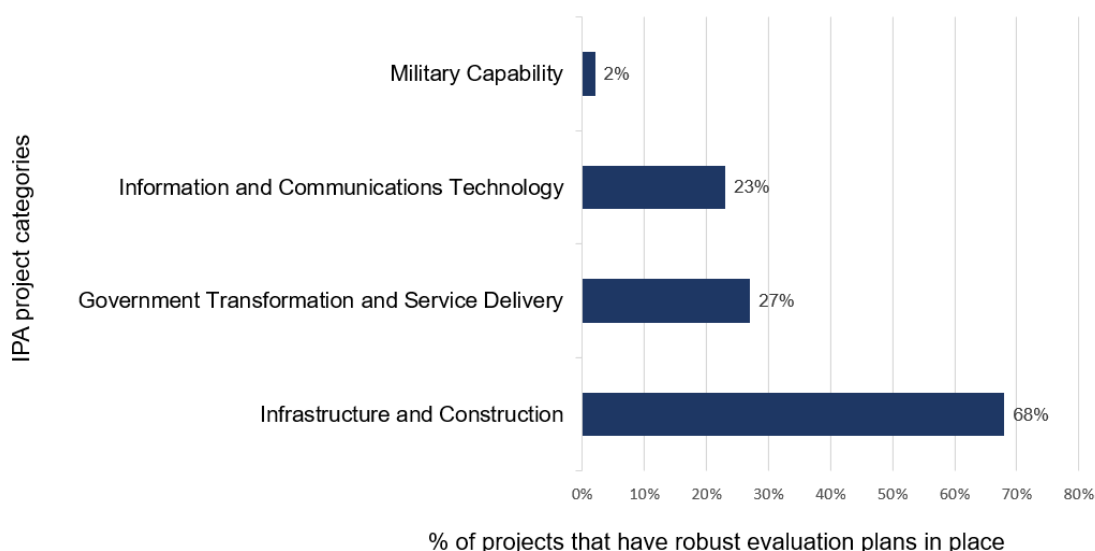
Notes:
1. The total number of projects in each category is: Government Transformation and Service Delivery (89); Information and Communications Technology (26); Infrastructure and Construction (68); Military Capability (44).
2. How to interpret this figure: This figure shows the proportion of major projects within each IPA annual report category that were found to have any evaluation plan in place at all. For example, among Infrastructure and Construction projects, 84% had an evaluation plan.

Infrastructure and Construction major projects are also considerably more likely to have robust evaluation plans than other categories. As shown in Figure 3 below, 68% of Infrastructure and Construction projects submitted evaluation plans that met the robust criteria, twice the average for all GMPPs. This compares to 23% of Information and

Communications Technology and 27% of Government Transformation projects, while Military Capability projects were the least likely (2%) to have robust evaluation plans.[8]

Figure 3: A higher proportion of Infrastructure and Construction major projects have high-quality evaluation plans compared to other types of major project

Percentage of GMPP projects with robust evaluation plans by IPA category, UK, 2023 to 2024

% of projects that have robust evaluation plans in place

Source: ETF assessment of evaluation plan quality; IPA annual report category from IPA annual report (2023-24)

Notes:
1. The total number of projects in each category is: Government Transformation and Service Delivery (89); Information and Communications Technology (26); Infrastructure and Construction (68); Military Capability (44).
2. How to interpret this figure: This figure shows the proportion of major projects within each IPA annual report category that were assessed as having robust evaluation plans in place overall. For example, among Infrastructure and Construction projects, 68% had robust evaluation plans.

Coverage and quality by evaluation type

There are three main types of evaluation: *process evaluation* examines how a project has been delivered in practice; *impact evaluation* aims to establish whether a project worked or did not work to change intended outcomes; and *value-for-money* evaluation explores whether the benefits of the project were worth the cost required. This review looked at the quality of plans for each of these three types of evaluation individually, alongside an overall assessment of the quality of evaluation plans as a whole.

---

[8] Particular challenges with developing evaluations for Military Capability projects were identified during the review, such as outcomes being harder to define or measure than other types of projects. Whilst relatively few of these projects met the threshold used for defining high-quality evaluations in this review, the review team observed improvements in both coverage and quality of evaluation for this project category throughout the review.

There was slightly more coverage of process and impact evaluation plans than there was of value-for-money evaluation plans. 54% of major projects had an impact evaluation plan, 55% had a process evaluation plan, and 46% had a value-for-money evaluation plan in place. 41% of projects had all three types of evaluation: process, impact and value-for-money.

In terms of quality, the incidence of a robust assessment was highest among impact evaluation plans and lowest among value-for-money evaluation plans. Nearly one third (31%) of major projects had a robust impact evaluation plan, compared to 28% with robust process evaluation plans and 17% with robust value-for-money evaluation plans.[9]

Figure 4: Nearly one third (31%) of major projects had a robust impact evaluation plan

Percentage and total cost of GMPP projects with robust evaluation plans by evaluation type, UK, 2023 to 2024
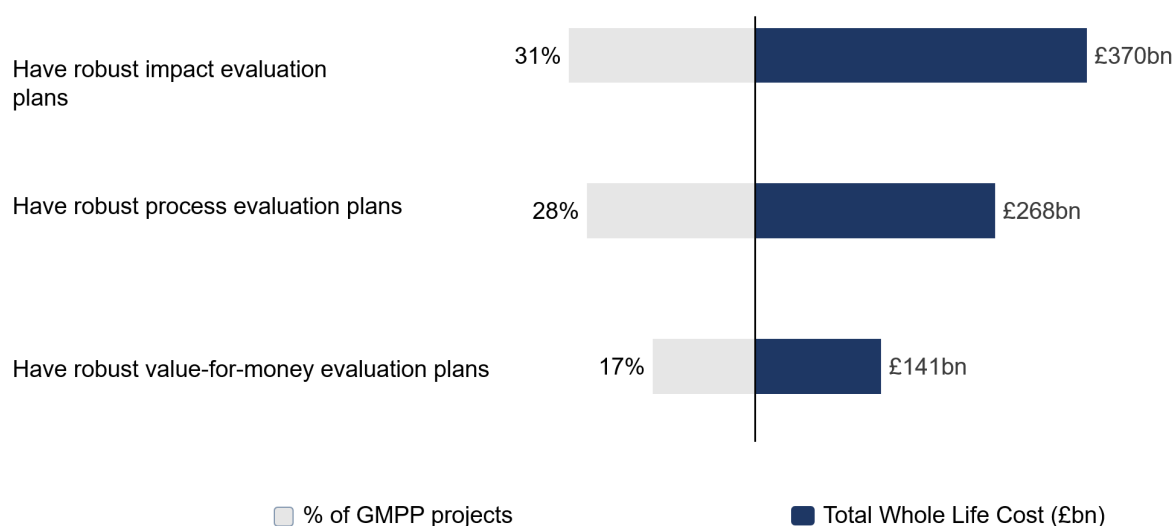
Source: ETF assessment of evaluation plan quality; IPA annual report category from IPA annual report (2023-24)

Notes:
1. The total number of projects is 227.
2. How to interpret this figure: This figure shows the proportion of major projects that have robust evaluation plans in place according to each of the three main evaluation types, and the total cost represented by those projects. For example, 31% of the GMPP have a robust impact evaluation plan in place, representing £370bn of spend.

---

[9] Note that it was possible for projects to be assessed as having overall robust evaluation plans, even if not every evaluation type (process, impact and value-for-money) was separately assessed as robust. This could be the case if the package of evaluation activity as a whole was considered appropriate and proportionate, given its stage at the time of review.

Section 2 Barriers to evaluating major projects

This section summarises the barriers and challenges to evaluation of major projects. These were explored to understand the reasons for the persistent gaps in evaluation coverage and quality in the GMPP (outlined in Section 1) and to inform the steps required to address these (outlined in Section 3). The barriers were identified by Ipsos and Ecorys as part of their analysis, principally from research interviews with a sample of major project teams, and are summarised below. Further detail is provided in the supplementary interim findings report (published separately).

The review identified several barriers and challenges to the systematic application of evaluation across the GMPP. These can be classified into three categories: operational barriers, cultural barriers and resourcing barriers.

Operational barriers

Some projects were not able to access or generate the data they needed for evaluation. There are several reasons why this happened.

Evaluation was often not planned early enough to allow teams to identify suitable data sources, set up the necessary data sharing permissions to access data, and/or collect new data. In some cases this was due to a shortage of people with the right skills in departments able to support major project teams with evaluation planning. In other cases, it was linked to evaluation not being appropriately prioritised at senior levels (discussed further below).

Many project teams also reported a lack of time to plan evaluation early enough to develop high-quality designs, due to the fast-moving nature of implementing some major projects.

In other cases, there were specific complications linked to the nature of the project that created operational challenges for planning and delivering high-quality evaluation.

Cultural barriers

Some projects reported encountering resistance to evaluation among more senior staff and decision makers, with evaluation being perceived as a 'luxury' or lower priority in the context of immediate delivery pressures. As a result, insufficient resources were allocated.

In other cases, project staff reported a perceived lack of flexibility in designing evaluations - feeling that the three main types of evaluation (impact, process and value-for-money) were not always relevant to their projects. Some project staff felt that they had developed alternative approaches that worked well for their needs, such as benefits management methods, and called for greater acknowledgement of the flexibility needed in evaluation design. These insights point to a need for further clarification on the relationship between evaluation and other practices such as benefits realisation, to ensure the merits of each

approach is well understood and that they are applied proportionally to the context of each major project.

Resourcing barriers

Some major project staff reported having insufficient resources (staffing, funding, teams, and systems) dedicated to evaluation. While this review found evidence that this is improving within several departments, there is still some way to go for some projects to have the staff and budget allocations needed for high quality evaluation.

Section 3   Action plan

Based on the findings of this review, we have identified a set of actions to improve evaluation of major projects:

1. Improving evaluation governance in HMT, NISTA, and across government.
2. Improving evaluation capability in project delivery, NISTA and HMT.
3. Embedding evaluation into project delivery.
4. Promoting the value of evaluation to key stakeholders.

These are summarised in Figure 5 below, and build on the recommendations identified by Ipsos in their interim findings (see the supplementary interim findings report, published separately).

Figure 5: Action plan to improve evaluation of major projects

The action plan represents an initial set of actions to improve major project evaluation, which will be completed by the end of 2025/26. Implementing these actions will require joint working between the ETF and other parts of government including HMT, NISTA and government departments.[10] We therefore list the main government stakeholder for each action in the tables below. These actions build on the broader work of the ETF to promote evaluation across government as part of the ETF's Strategy.[11]

The action plan will be governed and monitored by a joint working group with representatives from the ETF, HMT, NISTA and government departments. The working group will review progress following the completion of the action plan, invite feedback and identify what further steps are required to continually improve evaluation of major projects over the longer-term.

---

[10] NISTA is part of HMT but is listed as a separate stakeholder given NISTA's specific focus on major projects.
[11] https://www.gov.uk/government/organisations/evaluation-task-force

Improving Evaluation Governance

Major projects are subject to a set of central and departmental governance and assurance processes. The following steps will be taken to strengthen the emphasis on evaluation within these processes and embed proportionate evaluation as a requirement for major projects.

Stakeholder	Action	Progress	Deadline
HMT	1.1 Strengthen the emphasis on evaluation requirements in the Treasury Approval Process guidance.	Complete	April 2024
	1.2 Scrutinise compliance by major projects with the strengthened evaluation requirements for the Treasury Approval Process.	Ongoing	Ongoing
	1.3 Review and update evaluation requirements in Managing Public Money and identify updates for the Green Book for implementation following the wider Green Book review.	In progress	March 2026
	1.4 Add statements on evaluation compliance to the requirements for Accounting Officer's Assessments and Annual Report and Accounts.	In progress	
NISTA	1.5 Include questions on evaluation in NISTA's quarterly reporting system (GRIP).	Complete	June 2024
	1.6 Develop NISTA data systems to collect and disseminate monitoring data on evaluation performance to NISTA staff and departments	In progress	September 2025
	1.7 Update materials for NISTA's onboarding processes and assurance gateways to reflect the strengthened evaluation requirements in the Treasury Approval Process guidance.	In progress	
Departments	1.8 Identify cases where improvements may be needed within departmental evaluation governance, such as investment committees or approval boards.	In progress	March 2026
	1.9 Monitor the implementation of departmental evaluation strategies for progress in improving major project evaluations.	Ongoing	Ongoing

Stakeholder	Action	Progress	Deadline
NAO	1.10 ETF to provide insights on performance against evaluation requirements where appropriate to inform NAO reviews.	Ongoing	Ongoing

Improving Evaluation Capability

Improving evaluation capability is necessary in many areas to ensure that major project teams have access to the skills that are needed to develop and deliver good quality evaluations. The following steps will be taken to enhance capabilities in evaluation for those working in and with major projects across the GMPP system.

Stakeholder	Action	Progress	Deadline
GMPP project teams	2.1 ETF to create the Evaluation Registry as a tool for teams to log and browse evaluations across government, allowing them to share good practice and findings.	Complete	April 2024
	2.2 ETF to develop five 'evaluation demonstrators' to show good practice for major projects with characteristics identified as posing challenges for evaluation.	In progress	September 2025
	2.3 ETF to prioritise major projects for its evaluation advice and access to the Evaluation and Trials Advice Panel.	Ongoing	Ongoing
	2.4 ETF to develop tailored evaluation training and resources for major project teams.	In progress	March 2026
	2.5 ETF to develop a commercial solution for faster access to evaluation suppliers outside of government when required.	In progress	
NISTA	2.6 ETF to deliver training for NISTA assurance reviewers to help them to scrutinise evaluation in the GMPP assurance pathway.	Complete	March 2025
	2.7 ETF to develop evaluation content for relevant NISTA training and resources, including those for Senior Responsible Owners, such as the Major Projects Leadership Academy.	In progress	March 2026

	2.8 Increase NISTA's in-house evaluation capabilities over time, so that it is better equipped to assure compliance with evaluation requirements for major projects.	Ongoing	Ongoing
Departments	2.9 Ensure that evaluation capability building work in departments is suitably aligned to major projects.	In progress	March 2026

Embedding evaluation in project delivery

Evaluation needs to be better embedded throughout the project delivery cycle, so that evidence is available to enable continuous improvement throughout delivery and lessons learned inform the design of future projects. The following steps will be taken to do this.

Stakeholder	Action	Progress	Deadline
Major project teams	3.1 Major project teams are required to log their evaluation plans on the Evaluation Registry and upload published findings as these become available. ETF will monitor compliance on an ongoing basis.	Complete	April 2024
NISTA	3.2 Ensure evaluation requirements are clearly set out in the Teal Book guidance for project delivery.	Complete	April 2025
	3.3 The ETF will work with NISTA to review the Project Delivery Capability Framework and outline evaluation content to consider for the next refresh.	In progress	March 2026
	3.4 Include evaluation requirements in appointment letters and guidance for Senior Responsible Owners of major projects.	In progress	September 2025

Promoting the value of evaluation to key stakeholders

All of the steps above require raising awareness about the value and importance of evaluation across the GMPP. The following steps will be taken to promote the value of evaluation.

Stakeholder	Action	Progress	Deadline
HMT and NISTA	4.1 ETF to establish a working group to identify opportunities to promote evaluation for major projects and oversee the delivery of this action plan.	Complete	February 2025
Departments	4.2 ETF to disseminate the findings of this evaluation review with relevant departments and in cross-government forums.	In progress	June 2025
	4.3 ETF to raise awareness about the importance of evaluation and the quality expected, through meetings with relevant major project teams, portfolio offices and evaluation teams.	Ongoing	Ongoing

Annex A   Evidence used for this review

Gathering evidence

Ipsos and Ecorys carried out Phase 1 of this review in 2023. During this phase, a survey was sent out to departments' central evaluation teams to ask for information about the evaluation plans associated with all major projects in the department. Ipsos and/or Ecorys reviewed all evaluation plans that were submitted to assess whether they were robust and proportionate (the criteria for assessing robustness is described in the next subsection). During Phase 1, Ipsos also carried out some qualitative interviews with major project teams to help understand any barriers and challenges around evaluation.

The ETF carried out Phase 2 of the review during 2024-25. This phase was undertaken to ensure the final results reflected ongoing developments in evaluation plans since the review started and the latest composition of the GMPP. Phase 2 therefore focused on major projects falling into one of the following categories:

- Projects that did not have any evaluation plans during Phase 1.

- Projects where evaluation plans reviewed in Phase 1 were not found to be robust.

- Projects that were new to the GMPP since Phase 1.

Senior officials in HMT and IPA wrote to relevant SROs of projects meeting any one of those criteria in Spring 2024, to ask them to share evaluation plans. If projects did not yet have evaluation plans in place (for example, in cases where plans were still being developed), they were asked if they could commit to sharing a robust plan by December 2024. Members of the ETF reviewed the evaluation plans submitted in Spring 2024, and again in December 2024 when the final set of evaluation plans was shared.

The ETF provided some limited support to major project teams during the process of gathering documents, for example, meeting with project teams to discuss expectations and answer questions. The ETF also consulted regularly with departments' central evaluation teams.

The process of undertaking this review and the phased approach appears to have helped to support increases in evaluation coverage during the review period. At the end of phase 1 22% of major projects were assessed as having robust evaluation plans, compared to 34% at the end of phase 2.

Reviewing evaluation plans

In both phases, a bespoke quality assessment framework was used to assess the quality of evaluation plans. This framework was developed with Ipsos and Ecorys to capture the following dimensions:

- **Aims**: Clarity of the evaluation's rationale, key questions and target audience.

- **Methods**: Quality of the overall evaluation design, and individual components (process evaluation, impact evaluation and value-for-money evaluation).

- **Management**: Suitability of the management plans in place around the evaluation (for example, availability of staff and budget to deliver it).

Within these key areas, the review framework included a mix of simple Yes-No questions (for example, "Is there a process evaluation?") with questions to assess the quality of different aspects using a scoring of 0-2 (for example, "If there is a process evaluation, how robust is the design?). The scoring approach was as follows:

- 0 = There is substantial missing information to be able to judge the quality of the respective criteria;

- 1 = There is limited information supplied to be able to assess quality, and/or it requires significant improvement;

- 2 = There is substantial and satisfactory information for the criterion. In some cases, a margin for improvement is possible;

- N/A = This criterion does not apply to this evaluation.

There was one difference in the scoring approach between the two phases. For projects assessed during Phase 1, the ETF developed an overall summary assessment of the robustness of the evaluation by taking the average score awarded to the process, impact and value-for-money components. In Phase 2, the ETF review team awarded an overall score directly (without using an average).[12]

In Phase 1, assessments were carried out by Ipsos and Ecorys. In Phase 2, assessments were carried out by ETF evaluation experts: each plan was reviewed independently by two reviewers, who then held a moderation meeting to discuss their findings and agree on final scores.

The final assessment of robustness used in this review is based on a combination of Phase 1 and Phase 2 assessments of the 2023-24 GMPP, which comprises 227 projects.

Limitations of this review

The review only captures a snapshot of evaluation coverage. This report sets out findings against the list of major projects captured in IPA's Annual Report for 2023-24. This provides a static snapshot of the coverage of robust evaluation across the GMPP at a particular point in time. The findings from this review will become outdated over time as new major projects

---

[12] This refinement was made to allow reviewers greater discretion to determine if an overall package of evaluation was robust and proportionate in the specific context of the project reviewed, even if some components were not present. For example, in a small number of cases, such as the renewal of some digital service provisions, reviewers judged that a process evaluation alone was sufficient and appropriate for a major project evaluation plan to be considered robust overall, when taken alongside existing benefits realisation plans.

are onboarded, and others are offboarded, from the GMPP. The evaluation plans of individual projects will also evolve over time, as they move from planning to delivery of their evaluations.

Key information needed to make an accurate assessment may have been missing for some projects. We assessed the quality of evaluation plans based on documentation provided. Reviewers may not always have had access to important context or accompanying information to inform the scores awarded. This meant that it was sometimes difficult for reviewers to determine the appropriate score due to a lack of evidence (particularly for questions in the management and governance section).

There are challenges with assessing the quality of major project evaluations in a standardised way. This review has developed a quality framework to assess each evaluation plan in a structured way while taking into account the context of each major project. However, major projects vary considerably in their scale, timeframes and what they deliver. This makes reflecting the specific context of individual major projects challenging and requires a degree of professional judgement from the evaluation reviewers in each case.