



Department  
for Education

# **DfE Evaluation Strategy 2022**

**June 2022**

# Contents

Forward from the Permanent Secretary	3
Introduction	4
Our evaluation strategy for 2022	6
Enabling greater oversight of evaluation activity	7
Ensuring evaluations are proportionate and well planned	8
Enhancing the robustness of our evaluations	9
Encouraging efficiency in evaluation	10
Next Steps	11
How will we know we're succeeding?	11

## Forward from the Permanent Secretary

Building a strong, robust, evidence-base to provide the department with up-to-date intelligence to deliver the best possible support to children, and for students and learners of all ages, has never mattered more. With the recent introduction of enhanced support within government to develop and boost the government's evaluation capability, we are taking the opportunity to renew our evaluation strategy. Our aim is to maintain and enhance the culture of evaluation within our department and ensure we continue to be at the cutting edge of evaluation practices.

This document sets out our strategy to ensure our future evaluation plans are proportionate (taking into account the amount being invested in the policy), comprehensive (answering the full set of questions required to take a decision on the future of the policy) and consistent (we are designing evaluations of the right standard *across* DfE). We also want to support the bodies we work with, such as schools and local authorities, to continue to learn and grow in their own evaluations through the data we, and they, collect.

Three elements will ensure the success of this strategy within the department. Firstly, the department's oversight of policies and their evaluation plans will be consistent. Secondly, we will ensure analytical and policy staff have the support, guidance, and training they need to succeed. This document sets out how we intend to meet the first two of these needs. The third element required is a shift towards a more evaluative culture across Whitehall. We will utilise the additional support provided to Departments to demonstrate the impact of policies that is being provided by the Cabinet Office through the Evaluation Task force, to ensure our department continues to be at the forefront of that change.

Within the department we are committed to demonstrating the impact of the services we support, understanding what we do well and what we can do better, and feeding that insight into policy development. We recognise the need for policies to be designed with robust evaluation in mind; to be appropriately resourced and have the time built in for benefits to be realised and measured. The publication of this document is a statement of intent by myself and my leadership team to take an active role in reinforcing our culture of robust evaluation. To achieve our ambition, we will commit to work closely with our partners and stakeholders.

I am therefore pleased to introduce this strategy which sets our direction to build on the good practice we have already put in place, and outlines where we aspire to be in the future.

[Signature]

Susan Acland-hood

## Introduction

The Department for Education is responsible for education and services for children and young people in England, as well as apprenticeships and wider skills. At our heart, we are the department for realising potential. We enable children and learners to thrive, by protecting the vulnerable and ensuring the delivery of excellent standards of education, training, and care. Our work helps realise everyone's potential – and that powers our economy, strengthens society, and increases fairness. We work to provide children's services, education and skills training that ensures opportunity is equal for all, no matter background, family circumstances, or need. We work with 17 agencies and public bodies to achieve that aim. In our 2021 Outcome Delivery Plan<sup>1</sup> we reported that we employed around 7,300 people (FTE) and our departmental expenditure limit (excluding depreciation) was around £77 billion.

Education will be the driving force to help unleash Britain's potential and rebuild after COVID-19. Now, more than ever, it is vital that we invest in having the best evidence-base possible to support the delivery of our priorities. HM Treasury has published guidance on how to undertake robust evaluation of our policies and create an evidence-base of 'what works'. The Green Book provides guidance on how to appraise and evaluate policies<sup>2</sup>, while the Magenta Book<sup>3</sup> provides guidance on what to consider when designing an evaluation. This strategy builds on both these documents.

Our vision is for a department which is able to access a strong evidence-base to support policy making and able to demonstrate the changes our policies are achieving. This document sets out DfE's strategy for future evaluation of departmental policies; and how the department will ensure that they are targeted, proportionate, of a high quality, and are contributing to our understanding of how best to support children's services and education and skills.

---

<sup>1</sup> <https://www.gov.uk/government/publications/department-for-education-outcome-delivery-plan/dfе-outcome-delivery-plan-2021-to-2022#a-executive-summary>

<sup>2</sup> <https://www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-government>

<sup>3</sup> <https://www.gov.uk/government/publications/the-magenta-book>

## Expiry date

This guidance will expire June 2032

## Who is this publication for?

This strategy document is for:

- DfE staff
- Partner agencies and government bodies
- Local authorities
- School leaders, school staff and governing bodies in all maintained schools, academies and free schools

Our aim is that this document starts and/or continues conversations around evaluation.

## Main points

The department has renewed its evaluation strategy. The four main aims of the strategy are:

- Enabling greater oversight of our evaluations
- Ensuring evaluations are proportionate and well planned
- Enhancing the robustness of our evaluations
- Encouraging efficiency in evaluation

This document sets out how we know if we are achieving our aims.

## Our evaluation strategy for 2022

Project evaluations allow us to assess the efficacy of the design, implementation, and impact of our projects. By enhancing our evaluation capacity, capability, and oversight, we are ensuring that we can monitor the performance (outputs and outcomes) of our policy interventions, make evidence-based policy decisions and maximise the impact of our work.

In 2011, we reviewed our current evaluation practices and developed an initial strategy to improve the evidence-base that was feeding into schools. As a result of this work, we set up the Education Endowment Foundation to increase the use of randomised controlled trials in schools. We have also set up the Children’s Social Care Innovation Programme to support robust evaluation in local authorities. We embedded a behavioural insights unit, engaged in cross-government evaluations, and also are active participants on the government Trial Advice Panel. This strategy aims to build on that progress, with a focus on our evaluation of central DfE policy design and implementation. We propose the following four key areas of our strategy to enhance our evaluation work, which this paper will explore in detail (Table 1):

**Table 1: Summary of evaluation strategy 2022-2025**

Aim	Detail
<b>Enabling greater oversight</b>	Continue the improvement of our management information, to produce detailed internal information on our ongoing evaluation projects. This will deliver insight to DfE on which parts of our work are not currently being evaluated. Providing a full internal picture of the level of evaluation for all projects to get a picture of the coverage and different types and scopes of ongoing evaluations
<b>Ensuring evaluations are proportionate and well planned</b>	Be strategic in our overall evaluation activity, ensuring evaluations are only in place when needed, and that adequate support is given where evaluations are complex. A strategic approach to evaluations will allow for early decision making and the collection of baseline data where appropriate. In this way, risks are considered and mitigated, ethical and GDPR requirements are met, and evaluations are aligned to business planning.
<b>Enhancing the robustness of our evaluations</b>	Provide leadership and team support on increasing the quality of our evaluations and ensuring skills and advice are in place to produce robust evaluations of our policies and programmes. Ensuring all DfE evaluations are performed to a high standard. This allows DfE to make evidence-based decisions when targeting resources and respond robustly to external scrutiny. It also allows for findings to be easily compared in the future and studies to be replicated.

<b>Encouraging efficiency in evaluation</b>	Support efficient practice, providing advice and guidance to teams on existing resources so that external and in-house resources (such as previous evaluations or existing data) are utilised where appropriate. This will reduce duplication of effort and allow for easier synthesis of existing evidence. Also encouraging the use of internal and external longitudinal data and making data easier to use in evaluation by making it easier to track evaluation participants in central datasets.
---	--

## Enabling greater oversight of evaluation activity

As is set out in the government Magenta book<sup>4</sup>, evaluations can be broken down into three broad groups:

- **Process evaluation**- evaluations with a focus on how a policy is implemented and operated.
- **Impact evaluation**- evaluations with a focus on the causal effect of a policy.
- **Value for money evaluation**- evaluations which compare the benefits of a policy with its costs.

We will undertake a mapping exercise to create an internal filterable database to monitor live evaluations, which can be automatically added to by our systems. This will allow our central evaluation team to identify gaps in project evaluation and provide targeted support.

Historically, the DfE has predominantly run process evaluations of policies. These may, for example, check if stakeholders are happy with the delivery of a policy through interviews or surveys and may collate reported positive changes those stakeholders feel have occurred due to the policy. There is more we can do to make a causal link between a policy and impact. For example, a mentoring scheme may be run within a school where grades increase, and staff and students have fed back that the scheme was helpful. Ideally, we need to be sure that grades increased because of that scheme and not another reason. Although process and value for money evaluations are important, we are aiming to better target and increase our impact evaluations, where appropriate. We will increase our capability to undertake such evaluations, so that we can link our activities to more quantifiable results.

---

<sup>4</sup> <https://www.gov.uk/government/publications/the-magenta-book>

## Ensuring evaluations are proportionate and well planned

Planning how the success of a policy, activity or other decision will be measured at design stage is key to a robust evaluation. Collecting baseline data at the start is essential to ensure feasibility and whole team understanding of success criteria. It also allows teams to work in decision points and respond to issues in real time.

Using control groups (preferably randomly comprised of hypothetically potential participants) allows for the strongest causal links between a policy and impact, as the evaluators can say with certainty that any measurable difference between the treatment and control groups is not due to some other factor.

The department has governance structures in place at the business case and research procurement stages, where teams who require support to ensure evaluation is built in at the planning stage will be identified. Where evaluations are planned, the appropriateness of that evaluation within the wider DfE picture can be examined. Activities which are a) high cost, b) novel, or not based on a strong evidence-base or c) high risk are prioritised for impact evaluation.

Such evaluations can be time and resource intensive. Setting up a control group therefore requires early planning. Measuring policy impacts on a control group, as well as a treatment group, can be particularly resource intensive and, even when feasible, may not be appropriate for low risk, low impact projects. It is therefore important to be strategic in evaluation planning.

Our governance structures provide challenge to teams to ensure evaluation is considered early and objectively. This ensures we are prioritising evaluations of policies which are high cost or high risk and do not have a strong evidence-base. When submitting research proposals, teams are also asked to include any ethical concerns they may have about the projects.

The department will provide useful resources for teams to help with strategic decision making. This includes guidance for policy professionals to help them understand the key principles of evaluation planning, and for analysts to help them advise on the appropriateness of evaluation plans (including 'red flags' such as evaluations which place inappropriate burdens on delivery partners, or which measure against inappropriate outcome measures).

## Enhancing the robustness of our evaluations

After agreeing what level of evaluation is needed, the next step is to ensure that sufficient skills exist within the department to conduct the agreed evaluation. Evaluations are only worth doing if they are done well. An evaluation that has the wrong scope, is measured against the wrong outcomes, or isn't robust, will result in us not being able to confidently link our policies to impact.

At a minimum, policy teams (with help from Government Social Research (GSR) analysts) will need to be able to articulate the Theory of Change of their policy; their hypothesis of what the policy is aiming to achieve, ideally these will be set out in a logic model which links activities to specific outcome measures. A core element of a Theory of Change is that it encourages consideration of long-term impacts, and planning for data collection, as well as encouraging teams to test the assumptions behind their policies. A Theory of Change is also a useful tool in checking those policies and their accompanying evaluations are properly resourced, both in terms of time, budget, and staff. Theory of Changes are routinely used in DfE, but more can be done to ensure that they are drawn up early in the policy planning process to ensure that can inform evaluation planning.

The main point of contact for evaluation in DfE is the department's evaluation team. Their role is to provide central support to embedded analysts in providing advice to policy teams and organise training and events to help analysts build and maintain their skills. They are DfE's representative for central and cross-government evaluation colleagues, such as the Evaluation Task Force. The team draws together the evaluation expertise within the department to provide advice and workshop difficult evaluations. It also provides guidance to policy teams on involving analysts early in policy planning and critical assessment of evidence to help ensure planned policies are based on the most robust evidence possible.

We also want to increase the number of impact evaluations we are conducting, where appropriate. The 'gold standard' for this kind of evaluation is randomised controlled trials, as it creates a counterfactual group to say robustly what outcomes would have been without the policy. Experimental evaluations also allow impact to be quantified. The evaluation team will provide enhanced support, including bringing together expertise from across the department and government, to help teams decide if this approach is appropriate and advise on the planning and delivery of these kinds of evaluation methods.

## Encouraging efficiency in evaluation

Another aim of our strategy is to bring together methods and subject experts with teams who require help to ensure opportunities for efficiency are not being missed, for example collecting survey data to evaluate a policy where the same data is being collected elsewhere, or commissioning advice when we have the knowledge in-house. We also have governance structures in place to ensure teams check for learning outcomes of previous work done that maybe relevant.

The detailed data that the department holds (for example the National Pupil Dataset) represents an exceptional opportunity for creating intelligence on the impact of our work, e.g. by tracking the outcomes of participants. Opportunities also exist to work with data from other parts of government. To utilise this effectively, permissions to access the data required need to be obtained at the start of the project, including the consent of participants for future data matching where appropriate.

Thinking about data requirements early also keeps the option for quasi-experimental evaluations, where randomisation wasn't possible at the delivery stage, but potentially confounding variables can still be controlled between a treatment and control group within the data. We are also working to add data flags so that participants of trial interventions can be identified in our centrally collected data and our evaluation team is supporting the process to make this easier.

Another element of encouraging efficiency is knowing what hasn't worked in the past, so mistakes aren't repeated. With an increase of impact evaluations, we will have increased information on what doesn't work, as well as what has worked. Teams will be encouraged to share what hasn't worked so that teams can learn from the projects of others. and to publish all findings.

The department also recognises that there is a lot of good work going on within our remit by delivery partners and across government. We will continue to work with the government Trial Advice Panel and the Central Government Evaluation Group to build networks and share advice and findings across government. We will also continue to conduct joint evaluations with departments and work with What Works Centres to ensure our partner agencies are getting the evidence they need to support their work, ensuring expertise and data are being properly and appropriately utilised. As well as sharing findings, the department will also continue to work with delivery partners, such as education institutions and local authorities, to promote evidence-base practices. We will work together to build an evidence-base in their areas, for example through the provision of innovation funds and supporting partners in building appropriate evaluation infrastructure.

## Next Steps

### How will we know we're succeeding?

To track the implementation of this strategy the department has identified the following success criteria:

**Table 2: Evaluation Strategy 2022-2025 success measures**

<b>Enabling greater oversight</b>	<ul style="list-style-type: none"> <li>• We have an internal live database with details of all evaluations of policy interventions.</li> <li>• Policy professionals are incorporating evaluation early in their delivery plans.</li> </ul>
<b>Ensuring evaluations are proportionate and well planned</b>	<ul style="list-style-type: none"> <li>• Policy business cases are interrogated through DfE governance structures as to whether they have proportionate evaluation plans. Theories of Change and benefits realisation are signed off at business case stage.</li> <li>• Updated guidance and resources are available on the evaluation hub.</li> </ul>
<b>Enhancing the robustness of our evaluations</b>	<ul style="list-style-type: none"> <li>• All new evaluations have their robustness reviewed at research proposal stage</li> <li>• Evaluation is recorded in a way that allows for future replicability. (e.g. through trial protocols and publishing technical reports)</li> <li>• Staff evaluation confidence and capability is maintained. Links are maintained with external stakeholders to keep up to date with latest techniques.</li> </ul>
<b>Encouraging efficiency in evaluation</b>	<ul style="list-style-type: none"> <li>• We are publishing what works as well as what doesn't work in different contexts.</li> <li>• All new proposals can show to internal review boards that they have worked collaboratively by reviewing and considering relevant past projects and stakeholder work in research proposal submissions.</li> <li>• We have a searchable directory of previous project evaluations</li> <li>• Where appropriate, central data collections are used to support evaluations.</li> </ul>



Department  
for Education

© Crown copyright 2022

This publication (not including logos) is licensed under the terms of the Open Government Licence v3.0 except where otherwise stated. Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.

To view this licence:

visit [www.nationalarchives.gov.uk/doc/open-government-licence/version/3](http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3)

email [psi@nationalarchives.gsi.gov.uk](mailto:psi@nationalarchives.gsi.gov.uk)

write to Information Policy Team, The National Archives, Kew, London, TW9 4DU

About this publication:

enquiries [www.education.gov.uk/contactus](http://www.education.gov.uk/contactus)

download [www.gov.uk/government/publications](http://www.gov.uk/government/publications)



Follow us on Twitter:  
[@educationgovuk](https://twitter.com/educationgovuk)



Like us on Facebook:  
[facebook.com/educationgovuk](https://facebook.com/educationgovuk)