



Government
CAMPUS



An evidence-based and impact-led Government Campus



**Our evaluation
strategy**

**November 2021
Updated June 2023**





What's changed in this update?

- We have clarified our delivery commitments, to reflect that our third commitment now encompasses a range of enabling projects, rather than the one project initially highlighted.
- We have updated the contextual information and have moved some information into our accompanying delivery update, where it relates to a specific period of time rather than the lifetime of this strategy.
- We have updated our design principles for evaluations to reflect lessons learned since we started implementing this strategy. We have removed these from this document and published them as a separate annex, to make these more accessible and transparent.
- We have provided more information on the criteria we take into consideration when deciding which 'tier' a programme falls into for evaluation purposes.

Executive summary

This document outlines how we are using evidence and evaluation to make sure the Government Campus is evidence based and impact led. It sets out both a long-term ambition to transform the way that we use evaluation and research to inform our decision-making, and the steps we have been taking in the period up to 2025 to make measurable progress against that vision. There is a shorter, summary version of this document.

[Read the summary](#)

Our strategy focuses on:

- a) Increasing the **quantity** of evaluation.
- b) Increasing the **quality** of evaluation we undertake.
- c) Ensuring evidence and evaluation is as joined up as possible, so we are creating a strategic picture of what works for our learners, in our context, under what conditions and why.

It will comprise three main areas of activity:

1. Theory-of-change driven evaluation of the Government Campus as a complex intervention.
2. A tiered approach to evaluating specific learning and development programmes and activities within the Campus.
3. Enabling projects that create the conditions for success against the above two goals.

1. The Evaluation Opportunity

1.1 The Government Campus as a learning and development intervention

The Government Campus was established in September 2020. It is an intervention to build Government capability using learning and development (L&D). It is the strategic centre and the single location, for civil servants to develop the knowledge, skills and networks which they need to develop their careers in public service and deliver the capability Government needs now and in the future to solve the most complex policy challenges. The Government Skills and Curriculum Unit delivers the Government Campus, and exercises leadership in the areas which are best led from the 'centre' and/or where partnership working across the civil service is needed, whilst recognising and supporting local variation where there are legitimate reasons for this.

The Government Campus comprises a number of interrelated components, including:

1. A new **curriculum** for working in Government which specifies relevant knowledge, skills and networks and how they can be developed.
2. A range of **L&D interventions** mapped to the curriculum - this draws together existing L&D interventions, newly commissioned learning, and the refocusing, growth and development of some parts of the existing offer. The interventions include an extensive externally-commissioned offer, via central contractual learning frameworks, and internally designed and delivered activities (these are variously designed and/or delivered by GSCU, functions, professions, departments and agencies).
3. The L&D interventions and contexts that GSCU maps and coordinates particularly focuses on '**formal**' learning offer, which comprises everything from one-off, asynchronous self-serve courses, to extended, selective development schemes; as well as some '**semi-formal**' activities like coaching and mentoring, and is intended to increasingly take account of 'learning in the flow of work' and informal learning.
4. A new **digital Government Skills Campus** will be the primary interface of civil servants with the L&D offer and curriculum.
5. GSCU is responsible for policy and/or best practice around some key **systemic levers and mechanisms available to us to generate the behavioural and cultural change necessary** to develop civil servants' capability against the curriculum, including mandation, accreditation and assurance; communications etc., working with other actors in this space, such as leaders of functions, professions and departments.
6. The **GSCU** was established in September 2020 and is based in the Cabinet Office and located in multiple sites across the UK. The unit brings together the central teams focused on training and development, including cross-public sector leadership and talent schemes.

1.2 How did we arrive at our evaluation strategy for the Government Campus?

Our evaluation strategy, and our theory of change for the Government Campus was developed based on six months' scoping and systematic social research activity, led by a social researcher in partnership with the Campus and Curriculum Strategy and Partnerships team, and other GSCU teams via the Senior Leadership Team (SLT). This has ensured a robust evidence base for the decisions we have taken around our approach to evaluation and goes beyond ensuring that our approach meets stakeholder needs and expectations - by engaging stakeholders in co-design we have established a co-owned and consensual strategy. Further details on the approach taken to developing the strategy, and the social research process, can be found in [Appendix 1](#).

1.3 Evaluation priorities and goals

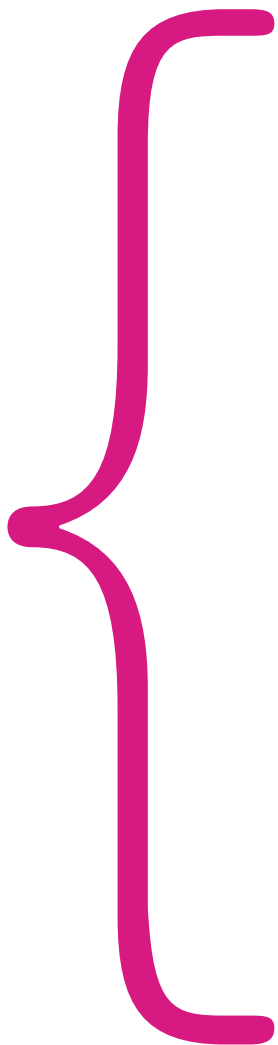
Our analysis of Campus stakeholders' views and experiences captured in the workshop data identified three main opportunities we must capitalise on:

- (i) To increase the **quantity** of evaluation, so that we find out more about the impact, implementation and return on investment (ROI) of our activities, so we can be even more targeted in our work and more confident in meeting the expectations of our senior stakeholders and the public.
- (ii) To increase the overall **quality** of evaluation (notwithstanding existing pockets of excellence), so we, and our stakeholders can have greater confidence in our findings, and we can answer some of the pressing questions we currently struggle to address.
- (iii) To **join-up** our L&D evaluation activities across GSCU and the wider civil service system so that our evaluation work is more **cumulative**, allowing us to reach more definitive answers to shared questions about impact, implementation and VFM; to have a better awareness of relevant similarities and differences across context; and to make better use of resources by not 'reinventing the wheel' each time.

1.4 Three interconnected areas of evaluation activity, and their intended impact

Based on the profile and purpose of the Government Campus, and the three priorities identified above, there will be three main interconnected areas within our Campus evaluation activity:

The three interconnected areas of Campus evaluation:



Area 1: evaluation of the Government Campus as an integrated intervention

- Campus evaluation carefully prioritised based on our **theory of change**, which lays out what outcomes we expect to see, when, and the theorised causal relationships we will need to test.

Area 2: evaluation of specific Campus components and programmes

- Evaluation of programmes, projects, schemes and components of the Campus using a **tiered approach**, guided by design principles for evaluations.

Area 3: enabling projects

- Ensuring we have the networks of internal and external expertise needed to ensure up-to-date knowledge of the evidence base, methods and measures.
- Building capacity, capability and resources within and beyond the Campus team.
- Optimising our commercial relationships as a lever for our evaluation goals.
- Working collectively with the L&D and people analytics communities across Government to share and define best practice and ensure evidence is available to support key decisions.

Area 1 is discussed first in this document, followed by Area 2, and finally Area 3.

Our evaluation activity has two primary purposes, and equal consideration has been given to each in planning our approach to evaluation:

1. To account for the impact and return-on-investment of Campus spend to key stakeholders and the public (**summative assessment**).
2. To generate purposeful and usable findings that help us refine programme design to create even greater benefits for participants and the system (**formative assessment**).

We regularly review our portfolio of evidence and impact work to check it reflects current policy and delivery priorities; meets our design principles and quality standards and is aligned with key civil service analytical priorities; and is being delivered in an efficient way which offers maximum value and benefits.

2. Our approach to evaluation of the Government Campus as an integrated intervention

2.1 The Government Campus is a complex intervention:

1. It is not a single intervention, but is comprised of multiple varied, nested and interconnected interventions, each of which is substantial in its own right. Its impact is not intended to be merely aggregative (i.e. adding together the impact of each component part tells us the impact of the whole). Rather, a primary assumption, which has driven the process of bringing together these various parts into an enlarged and integrated whole, is that *the whole can be greater than the sum of its parts*.

What does this mean for evaluation? It means that as well as evaluating the impact, implementation and return on investment (ROI) of component parts of the Campus, we need to use methods and measures that can help us capture the difference it has made joining these different parts together. This requires **theory-driven evaluation, based on a theory of change**. The theory of change will be translated by the GSCU analytical team into a strategic programme of research, outlining which elements of the theory of change we will test, when, alongside monitoring and evaluation of specific outputs and outcomes in the time windows in which we would expect to see them. Our interactive version of the theory of change will embed suggested measures, metrics and methods in relation to each specified output and outcome.

Campus evaluation design principle 1:

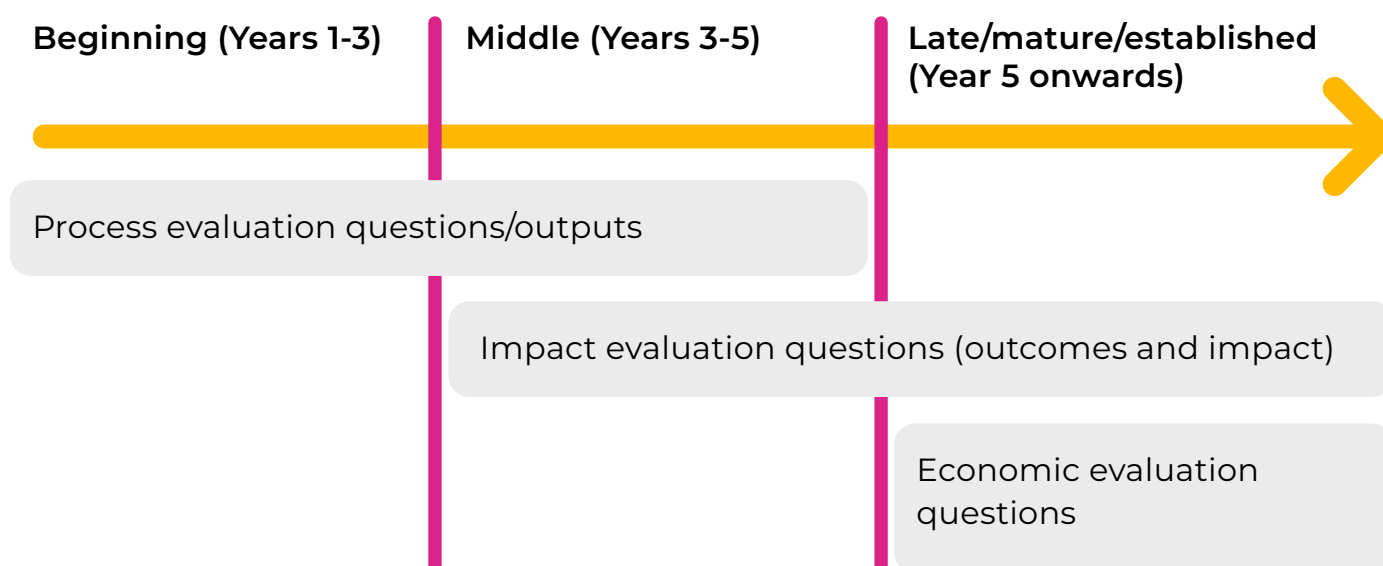
We will use a theory-driven approach to evaluation, based on our Campus theory of change. All evaluation and research work of Campus activities, undertaken by GSCU, must be aligned to the strategic programme of research articulated by the Campus analytical team based on this theory of change. This ensures we generate knowledge cumulatively, and prioritise effectively, thus making best use of our resources. Campus partners in departments, functions and professions are encouraged to make use of this strategic programme of research, the theory of change, and other resources, to support their own evaluation activities; co-ownership of important data generation/sharing and dissemination of findings between GSCU and other Campus colleagues will be an integral part of our partnership working.

Campus evaluation design principle 2:

Consistent with established practice in the evaluation of complex interventions,¹ we do not expect to evaluate every output, outcome, impact and causal assumption in the theory of change. Evaluation activity will be prioritised on the basis of: a) areas of greatest uncertainty b) areas in which answers are urgently required for work to progress as intended; c) most strategically important areas; and d) areas of greatest spend.

Our sequencing of evaluation foci broadly follows the following plan, notwithstanding where there will be some specific outcomes/impacts and ROI questions which we may be able to begin to address earlier, and some which will logically take longer to realise:

Stage of Campus lifecycle



Colleagues suggested a range of impact, economic and process evaluation questions in our initial strategy development work, presented in detail in our report on findings from that work, and we continue to develop and refine this list. These questions inform our planning of specific evaluations.

2. The Campus is intended to achieve change through multiple, interconnected pathways to change, some of which are human, social and cultural - like leadership, or a learning culture; others are processual, like adopting more consistent and coherent approaches, or gathering and analysing capability data; others are material, such as providing new digital learning environments, or physical collaboration spaces; most of the change pathways mix social/cultural, processual and material levers for change.

¹ [Moore et al. \(2019\)](#)

What does this mean for evaluation? Each change pathway will require bespoke approaches to evaluation which are capable of capturing the intended activity indicators and outcomes in a robust and meaningful way. We also need to assess how they intersect and whether this reinforces the intended changes or leads to unexpected consequences (e.g. trying to commission learning strategically from the centre, removing courses which do not match the central curriculum, could have the unexpected consequence of driving commissioning away from the central learning frameworks, leading to worse value for money, and weaker data on capability). This will require **a combination of quantitative and advanced qualitative methods**, using **robust, specialist measurement tools**, and a **joined up approach to evaluating**, rather than running evaluation projects in siloes.

Campus evaluation design principle 3:

We are proud to be experts in professional learning and will always draw on deep, specialist expertise of this field in our evaluation work. Specifically, we will use this expertise to ensure we are deploying the most effective and appropriate evaluation/research methods for the outcome(s) of interest, not just defaulting to a narrow range of methods, and to ensure we are using robust (and ideally pre-validated and/or peer reviewed) measures/tools to capture that outcome. If we do not have the right expertise in-house we will seek it through our networks of academic expertise and/or external tendering processes.

3. The Government Campus, and the GSCU which leads it, is – in most cases – quite significantly upstream from the changes it is trying to effect. In fact, it often does not own the delivery of much of the L&D offer, and learners and their line managers and teams are distributed right across the Civil Service. In terms of delivery, this requires a co-owned and partnership based approach, with careful consideration of how central levers may/not effect change downstream. In terms of impact assessment, this means that GSCU does not automatically have access to data on business impact or learners' trajectories and performance. In considering informal learning and the development of a learning culture, this is only visible during the course of day-to-day work, and cannot be seen or understood from the 'centre' without active strategies to investigate it.

What does this mean for evaluation? Firstly, the distance between GSCU and the data and participants it would need to access for evaluation, as well as the complex and cumulative nature of many of the evaluation questions that we want to answer, means that we need a genuinely **co-owned and collaborative approach across Government**. Secondly, it requires **careful distinction between where we can reliably ascertain the causal effect of a Campus intervention** (e.g. giving randomly sampled participants a test on statistical analysis skills before & after a relevant training event, to see if the training led to improved skills, versus a control group

who did not undertake the training) **versus where evidence-based demonstration of a plausible effect is more appropriate** (e.g. tracking participants' networks after a programme, where the programme can be shown to make a clear contribution to participants' collaboration with other leaders, but where this is not the only factor influencing this outcome). This distinction is important, because both are valid approaches to evaluation, but must be matched properly to the outcome of interest, to avoid slipping into less robust approaches on the one hand, or wasting resources on inappropriate quasi/experimental methods on the other hand. It will often be relevant to use advanced qualitative analysis methods to understand what the other factors are influencing a downstream outcome (aside from our intervention), so we can build these into future evaluation designs, thus moving closer to demonstrating causal effect for downstream outcomes.

Campus evaluation design principle 4:

Campus evaluation activity is co-owned by all the colleagues in the GSCU, departments, functions, professions and agencies/ALBs who contribute products and expertise to the Campus offer. Governance of evaluation activity will therefore be through the Campus Design Authority, and all evaluations will begin with careful scoping and consultation work to identify and engage relevant cross-government stakeholders and team members based on programme ownership and delivery, ownership of/access to data sources ('gatekeepers'), relevant expertise, and intended use of the findings.

Campus evaluation design principle 5:

We will be confident using high quality multi- and mixed-methods approaches to evaluation, to ensure we can capture both upstream and downstream outcomes of interest (as specified in the theory of change). We will use the right methods at the right time, to capture outcomes when they can realistically be expected to have emerged based on the programme design. We will use feasibility testing to ensure we only use more resource-intensive evaluation tools which seek causal attribution when our understanding of a programme and how we can apply different methods is sufficiently mature.

4. The Campus is seeking to effect change across the civil service, and (through some elements) the wider public sector. This is not a single, homogenous entity, but is a complex and varied sector or system. Local context, whether that be at the level of the individual learner, the team, the division, the organisation etc., means that the intended changes will play out differently, at different paces, in different contexts. Different barriers that need to shift, and different levers to generate or accelerate change might well be identified in different settings.

What does this mean for evaluation? It will be very important to **identify which contextual factors have been important for shaping how change was/not achieved through both the Campus as a whole, and individual interventions within it.** Over time, this will a) help build up a 'bank' of relevant contextual variables we can control for in quantitative analyses (although it is likely to initially require robust qualitative methods to pinpoint what these might be), and b) help us to adjust implementation approaches for context, boosting impact. This means that a broadly **realist approach to evaluation** will be beneficial - that is not just seeking to identify 'what works', but **'what works, for who, in what context, and why'**. This does not mean we will ask all four questions in every evaluation, but this will be a golden thread that will run through our strategic planning of evaluation activity. It is relevant to note that both robust quantitative and qualitative methods can be used in realist approaches, and that mixed methods designs are often appropriate.

Further information on realist approaches to evaluation can be found in the [Magenta Book annex](#).

Campus evaluation design principle 6:

Our programme of evaluation work will be unified by a realist approach to evaluation - that is, seeking not only to establish 'what works', but also 'for who, in what context, and why?'

3. Approach to Evaluating Campus components and programmes

3.1 Context

1. GSCU has a history of investing in evaluation in key areas. For example, the National Leadership Centre (NLC) established a programme of process, impact and economic evaluation, closely linked to ongoing programme design, and delivered through a partnership with an external research organisation. This substantial project has now been concluded and lessons are being taken forward into ongoing leadership development work. This work was particularly enabled by having a dedicated team and resource for programme evaluation. Where this was not the case, cultures of evaluation and evidence based practice have been harder to build.
2. GSCU has established a small team to deliver this evaluation strategy, led by a Head of Evidence and Impact. The delivery model is lean and efficient, using a mix of external and internal contracting to carry out evaluations, leveraging commercial partnerships that deliver L&D interventions and partnerships with external and internal experts to access evidence and data. GSCU also collaborates with departments and professions to evaluate interventions. The level of specialist resource departments and professions have to support local evidence and evaluation work varies, but is typically not substantial. Whether or not this capacity changes over time will be subject to wider policy on the civil service workforce, and local recruitment decisions. The implication of this resource environment is that the implementation approach to the strategy needs to be realistic and efficient. For this reason we are adopting a 'tiered' approach.

3.2 A tiered approach to evaluating specific components of the GSCU offer

1. Individual Campus projects, programmes and learning offers will be evaluated according to our **tiered approach**. This ensures analytic resource is directed proportionately and purposefully. All colleagues who lead or own the delivery of specific GSCU-owned programmes, courses, schemes etc. are supported in identifying which tier their activity falls into, and plan for evaluation activity accordingly. Each tier is accompanied with **design principles for evaluations within that tier**, which the Campus analytical team support colleagues and external contractors to apply, and which cross-reference methodological preferences and expectations established in the [HMT Magenta Book](#), the Government handbook for evaluation. These design principles are published as an annex to this document.
2. Years 1-3 of the strategy implementation have focused on developing and piloting the approach to tiering interventions, with a view to this being rolled out in its final form in year 3 onwards.

3. When determining which tier a programme falls into for evaluation purposes we consider a) level of spend; b) level of strategic importance of the programme; c) level of innovation or uncertainty in the evidence base for the programme; d) likely magnitude of the impact generating this evidence will have on future strategic and spend decisions (including considerations of when evidence is likely to be needed to have greatest impact).
4. The **change pathways elaborated in our theory of change act as initial thematic workstreams** linking different evaluation activities together, and ensuring they contribute knowledge to our overarching Campus evaluation. Programmatic research themes are also developed in line with policy and strategic priorities.
5. To support the implementation of this tiered approach, all GSCU colleagues who lead substantial courses, projects or programmes are being supported to deepen their knowledge of, and confidence with, evaluation through participation in appropriate **curriculum-linked learning opportunities**, such as the 'Introduction to Policy Analysis' courses available within the Campus offer, and shadowing and project opportunities with our GSCU analytical team. The GSCU analytical team will also be offered work shadowing and project opportunities with non-analytical teams and colleagues across GSCU, to **build strong inter-professional knowledge and capability, and thus an integrated approach to Campus analysis, strategy/planning and delivery**. This will be particularly important in enabling Tier 2 evaluation.
6. Given the need to map Campus programmes and activities to different tiers, and the subsequent need for detailed scoping of evaluation requirements and possibilities within each programme, we expect the first five year period of this strategy to focus on piloting and implementing new approaches and learning lessons. We expect to move all Campus programmes/components onto this tiered model of evaluation by the end of 2025. We are undertaking this in priority order, beginning with Tier 1 programmes, and all new and redesigned programmes.

Our evaluation strategy, tiered approach and design principles are available to all colleagues across Government to use as a tool for L&D evaluation, where they align with local priorities. You can access our Design Principles for Evaluations here:

**Read Design Principles
for Evaluation**

3.3 Evidence-informed practice

1. Carrying out evaluation has limited purpose if the findings do not contribute to the design of more effective L&D interventions, and, more widely, to organisational learning.
2. Our scoping work and workshop data reveals that while there are many examples of evidence-informed practice both within GSCU and Campus partners, this has not yet become routine nor culturally embedded. This was partly due to poor accessibility of relevant evidence at the time it was needed (compounded by the generally low level of evaluation of L&D interventions in Government settings to date, either internally or by external academic researchers), and partly due to programme design cycles and approaches often not being synchronised with the work (or potential work) of analytical colleagues, and not routinely incorporating attempts to consider robust pre-existing evidence (for example, very limited commissioning of Rapid Evidence Reviews, Systematic Reviews etc.).
3. There is an extensive body of research evidence in sectors including healthcare, policing and education, which have made significant progress in establishing evidence-informed cultures of practice, on the enablers and barriers which are relevant to this goal. This literature base demonstrates that changing practice is not as simple as 'pushing out' research/evaluation findings - cultural changes and new forms of work/new processes are needed to integrate evidence into day-to-day decision making. An early priority of our research and evaluation work will be to draw more systematically on this evidence base, and the experiences of other parts of government, as well as current reality in the L&D space, to identify the steps and interventions that may be needed to ensure that the findings of our evaluations emerging from this 'tiered' approach translate into genuine improvements in practice.
4. There is also significant scope to conduct a greater quantity of internal, robust and systematised research using or collecting organisational data, particularly to understand the nature of the 'problems' we need to solve, and organisational enablers and barriers.
5. One action point will be improving the accessibility of relevant findings, working with other analytical and L&D partners in Government.

4. Enabling Projects

In order to carry out our planned evaluation activity we need to invest in key projects to enable this. During the first five years of this strategy we expect these projects to include, as a minimum:

1. Ensuring we have the networks of internal and expertise needed to ensure up-to-date knowledge of the evidence base, methods and measures.
2. Building capacity, capability and resources within and beyond the Campus team.
3. Optimising our commercial relationships with suppliers of L&D interventions as a lever for our evaluation goals.
4. Working collectively with the L&D and people analytics communities across Government to share and define best practice and ensure evidence is available to support key decisions.

4.1 GSCU's enabling role

GSCU is committed to supporting the whole L&D community across Government in our shared mission to be evidence-based and impact-led in our work. Our ownership of the central contracts for procuring training is a key way that we do this, by ensuring these commercial relationships support great evaluation - that is, evaluation which is robust, purposeful and proportionate. We also invest in building shared networks so we all have better visibility on the evidence base that exists, is being developed and what it means. GSCU shares the methods, measures and findings from its evaluations through a range of channels to support colleagues who also want to invest in evaluation, and in turn, seeks out examples of good practice from others that we can apply.

4.2 An educative approach to carrying out evaluation and research

All our evaluation activities - whether carried out internally or externally, by analysts and non-analysts - must have an educative component - that is, they will include features which will actively build the knowledge, skills and/or networks (as defined by the curriculum) that non-specialist colleagues need to be able to participate in, champion and integrate evaluation and evidence-based practice through their work. This expectation will be built into all internal and external contracting for analytical work, and will be part of the role specifications of our Campus analytical teams.

When it is time to evaluate a Campus outcome or activity that you are engaged with, you can expect to be given the opportunity to learn something about evaluation or research, thus developing your own knowledge and skills, and the overall capability of the team(s) you are part of.

There are lots of collaborative and educative evaluation models you might be invited to participate in. For example, you might be involved in co-designing the evaluation approach and tools. You might be able to work as a co-researcher, helping to collect the data and/or analyse it. You might help analysts interpret the findings and consider their implications for practice. You might be invited to a seminar, presentation or 'lunch and learn' to find out more about the evaluation process, findings or something noteworthy within them. In all cases you can expect to know more, be able to do more, and think differently about evaluation than you did at the start of the process.

All we ask of you is to embrace the learning opportunity this presents. When digital and evidence skills are so central to the future of Government this will benefit you on an individual and organisational level.



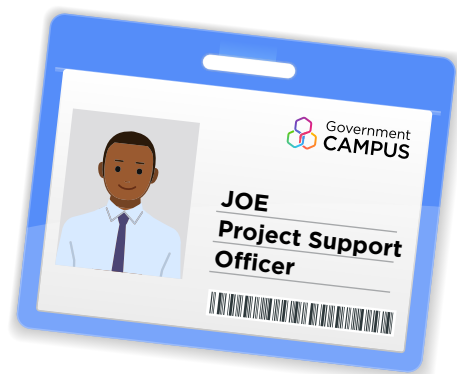
Case study: Learning about research-informed approaches by being part of the project team

In spring and summer 2021 we worked as co-researchers alongside a researcher who was on loan to GSCU from a university. We worked as a research team carrying out the social research which supported the development of this evaluation strategy, and learnt how to carry out and interpret qualitative data using systematic coding. Here we wanted to reflect on what we learnt and gained from being involved in that process, and how it creates a model for non-analytical colleagues to get valuable experience through our planned research and evaluation work.



Esther: My stand out reflection from being a co-researcher is how what looked like an insurmountable quantity of qualitative data can go through a rigorous process that results in a valuable and informative summary and useful findings. I already had some understanding and experience in analysing simple qualitative data, but this was entirely different. It's been enlightening to see how the processes – and rigour in application – have dealt with the complexity, and it gives me great confidence in the report.

Joe: It was certainly interesting to come back to a research project such as this after a year off from university! I've learnt so much about how research-driven work is done in an organisation – as Esther mentioned, it's also been incredibly valuable to be involved in turning an enormous amount of data into concise findings. To have gone from the very beginnings of the formation of the strategy, to watching it become a comprehensive piece, has been fascinating and I'm sure will benefit me as I develop my career.



Esther King, L&D Project lead and Joe Williscroft, Project Support Officer, Campus and Curriculum Strategy and Partnerships, GSCU

5. Resources, oversight and governance

5.1 Governance, Quality Assurance and Reporting

1. Overall Campus analytical activity is governed through existing Campus governance arrangements.
2. We have established an academic advisory network including a small number of internal government analysts/evaluation specialists from other policy/delivery contexts, and a range of external experts and will be broadening and widening these networks of advice over the duration of this strategy.
3. All research and evaluation plans/protocols and findings/reports are peer reviewed (external contractors will be expected to nominate a knowledgeable person of standing to peer review their reports, or to accept a peer reviewer nominated by GSCU).
4. GSCU ensures alignment with all cross-Government evaluation and analytical policy and approaches, maintaining an up to date index of these to refer to in our work.
5. The GSCU analytical team has membership of the relevant analytical professions, and via those professions and their Heads, ensures alignment with the functional standard, all relevant guidance and regulation, and the professional strategies.
6. Integration of research and evaluation is integrated as a regular item on relevant agendas and planning/delivery protocols and processes, at all levels of the Campus programme.
7. We follow **GSR principles for the publication of research and evaluation findings** (and, where relevant, data sets) including prioritising **transparency and rigour**, aiming to make all data and findings (good and bad!) **publicly available in a timely manner**, as far as possible where this does not significantly breach the confidentiality of government business. We also commit to including an **accessible summary for non-specialists, which suggest action points for practice** with all research and evaluation reports.

5.2 Resources

1. Evaluation activities are led by a small specialist GSCU analytical team.
2. This team will take responsibility for GSCU research and evaluation-related relationships with relevant stakeholders and maintaining peer networks with colleagues who supporting evaluation of L&D initiatives across departments, professions and functions, and analysts working in other areas of capability and workforce data.
3. We will adopt business partnering rather than a contracting approach. That is, we will take concrete steps to build true multi-professional working by embedding analysts with programme teams in flexible ways, and ensuring that opportunities are created

for programme colleagues to learn about analysis and evaluation, and vice versa, via relevant activities such as shadowing and project work.

4. In building our team we will prioritise specialist knowledge of evaluation methods, and the research base, for professional learning, capability building, leadership, networks/inter and multi professional working, and organisational change, especially in the public sector and other relevant comparator sectors. This will enable us to work at greater pace and achieve greater relevance and rigour in the work.
5. We will supplement internal expertise with smart, well-targeted external contracting from academic specialists in HEIs, and highly-qualified, specialist research and evaluation focused SMEs, to complement and extend internal capacity. In line with the wider cross-government strategic priority to reduce spending on external consultants, we do not intend to commission from generic consultancies and research organisations who do not add value over and above our internal capability.
6. To supplement our core capability, and benefit from porosity, we will make use of creative routes to bring relevant external research expertise into our team. These are likely to include, but are not limited to: postgraduate, ECR (early career researcher) and non-academic researcher internships, placements and secondments; policy fellows; PhD studentships; and externally-funded research projects, in partnerships with HEIs. We will also make good use of our own advisory panel (see below) and the Government [**Trials Advice Panel**](#), where appropriate.
7. We will also support non-analytical colleagues to undertake relevant graduate and postgraduate qualifications, as far as possible, to build internal capability. These might include, but are not limited to, post-graduate diplomas, accredited professional development units offered by HEIs and professional bodies, various levels of apprenticeships, professional Masters and professional doctorates. Anecdotal data from our workshops and research evidence from other contexts suggests that such qualifications can play a significant role in boosting knowledge, skills and confidence in social research methods relevant to the evaluation of professional learning. This is something we can investigate further more systematically to guide future decision-making in this area.

5.3 Ethical and legal duties

1. When we undertake evaluation which includes collecting data on individuals' knowledge, skills, networks and overall capability we have particular responsibilities to ensure that data is used ethically and legally.
2. Government generates a wide range of '**naturally occurring**' data (that is, data that we haven't collected ourselves for the purposes of research and evaluation, but exists independently of us) that can be used for the purposes of evaluation and research. For example, assessment and accreditation data, skills assessments, 360 degree feedback, personal development plans, quantitative People data, minutes of meetings, survey data etc. Making creative and intelligent use of existing administrative data reduces

the burden on colleagues to supply new data (so has ethical benefits), and often represents efficient use of resources. Therefore, such approaches will form a strong part of our approach, especially at Tier 2, and it is relevant to note that this approach is common in organisational research, and while careful ethical governance is needed, ethics are not generally a barrier to using such data for these purposes.

3. In general, **data assembled or collected for the purposes of evaluation is collected for the purpose of assessing the intervention, not assessing an individual, so should be handled fully confidentially**, so that data points are not attributable to named individuals, and should be used only for the purposes of answering the evaluation questions. There may be occasional exceptions to this (usually for ethical reasons,² or where the participant has given informed consent for additional use), which need to be carefully considered on a case-by-case basis. It is important in the context of evaluating internal L&D interventions, that colleagues are reassured that evaluation data is fully distinct from performance management data, and would never be used as such, unless this had been specifically agreed with the participant in advance.
4. **Assessment can often be formative as well as summative**, that is, as well as reaching a judgment, it can be used developmentally. For example, in a recent leadership development programme participants consented to being assessed and the assessment data was shared with them as part of the programme, for developmental conversations. The participants have highlighted this as a beneficial element of the programme. **We should be confident in embedding assessment into our programmes where it is necessary for better evaluation, but should also make sure this plays an educative role for participants** as far as possible.
5. It is relevant to note that for Tier 1 activities, the ethical obligations of ensuring impact and value-for-money of substantial areas of public expenditure renders the educative element of assessment relatively less significant - in this case evaluation is needed because of our duty to ensure we are spending public funds wisely, and are delivering on the most strategically important parts of our offer, thus contributing to overall effective government. At Tier 1, it is fully consistent with the civil service code to ask colleagues to contribute to these priorities through participating in assessment and/or data-sharing for the purposes of evaluation. All other normal ethical considerations should still apply.
6. A distinction can be made between administrative data which is collected for non-evaluative purposes, which we then use for the purposes of evaluation, and data which we collect specifically for evaluation. In both cases, data must be handled ethically and in line with our legal obligations, but the specific ethical and legal issues that will need to be considered will differ. Particular research designs also have particular ethical implications to navigate – for example, an experimental

² For example, in educational research, if in the course of an evaluation data emerged which suggested a child was at risk of serious harm, this would always be passed on, due to legal and ethical safeguarding obligations.

design requires one group to be the 'control' and not receive the 'treatment' (the L&D intervention), and it is thus important to ensure this does not unfairly disadvantage them, for example via a design where the waiting list becomes the control group, receiving the 'intervention' on its next run.

7. It is relevant to note that many parts of the public sector such as the healthcare/ the NHS, education and policing sectors already undertake a wide range of L&D evaluation work, including trials and quasi-experimental methods, successfully and ethically, and have established principles for data use and sharing which fall well within legal obligations. **While it is right to emphasise and focus on our ethical and legal duties, they are not a barrier** to conducting the sorts of evaluation we want to do. Our strong, existing public sector research and evaluation partnerships will enable us to learn from their solutions.
8. Responsibility for identifying ethical and legal issues and proposing approaches for their appropriate management will be the responsibility of the individual designing or (for external evaluation) commissioning a specific evaluation activity. This must be reviewed and signed off by a '**responsible person of standing**' from an approved list, and who must be a badged member of an appropriate Government analytical profession (most likely to be GSR) or, where appropriate, Occupational Psychology. This mirrors the ethical approval processes common in most HEIs.
9. If the 'responsible person of standing' deems that there are ethical complexities in the proposal that require further consideration, the proposal must be reviewed and discussed by a minimum of two 'responsible persons of standing', seeking external advice via one of our academic advisory boards, and/or from other internal stakeholders and experts, as required.
10. Desk-based research of existing literature in the public domain does not normally require such approval, but research using institutional documents or data, not in the public domain, generally would, as permission and/or ethical approval is likely to be required to re-use this data for different purposes.
11. There are specific protocols governing publication of research undertaken in Government settings, particularly when this involves internal or Government-produced data sets (quantitative or qualitative), and these protocols must be followed at all times, including in the citation of reports that have not been publicly published. This protects the reputation of Government research and prevents less robust findings circulating as 'fact', which is in everyone's best interests. It is relevant to note that these guidelines also emphasise our obligation to publish our findings.

GSCU regularly reviews its evidence and evaluation commitments and responsibilities in relation to current Government, civil service and organisational policies. We maintain an internal register of relevant policies to support our coordination with those.

Appendix 1: Process for developing the evaluation strategy and theory of change

1. The first phase of scoping involved a Social Researcher using stakeholder discussions and documents to understand the nature of the Government Campus as an intervention for professional learning and organisational change, and to scope the affordances and challenges for evaluation of different features of the Campus and the systemic context in which it is situated.
2. This initial scoping phase also included consideration of the wider professional and policy context for social research and evaluation in Cabinet Office and across Government, and the development of a deeper appreciation of the cross-civil service data ecosystem, and how this enabled and created challenges for Campus evaluation.
3. Between June and September 2021 we carried out two structured social research projects to inform our evaluation approach in a more systematic way. These were focused on i) evaluation aspirations and approaches, and ii) developing a theory of change for the Government Campus.
4. We invited Campus stakeholders across the civil service (departments, functions, professions, agencies/ALBs and central Cabinet Office analytical and CSHR teams), sampled on the basis of range and relevance (purposive sampling) to participate in a series of 90 minute digital workshops focused on Campus evaluation and a theory of change for the Government Campus. These workshops were recorded, transcripts generated, and physical outputs in the form of digital collaboration boards and the videoconference 'chat', were confidentially stored. Overall, 147 people registered for the workshops. Some of these attended all three available workshops, others only attended one or two. A small number registered but did not attend any workshops.
5. The theory of change workshops focused on reviewing and critiquing a draft theory of change which had been generated from framework-driven qualitative coding of documents and notes from stakeholder discussions, sampled on the basis of range and relevance (purposive sampling). The evaluation strategy workshops used a semi-structured approach based around a series of prompt questions.
6. In advance of the theory of change workshops, a systematic framework-driven synthesis was conducted of theoretically sampled documents and stakeholder interviews to populate an outline model of a theory of change. Inductive coding was used within framework categories to generate specific content e.g. types of output and outcome. This was used as a stimulus for discussion in the workshops. Following the workshops, 409 suggestions made by participants were identified from the data, transcribed, assessed and either acted upon or rejected, with a rationale being recorded.

7. Qualitative thematic coding was used to analyse all data generated from the evaluation-focused workshops, initially using framework-driven (deductive) coding, to ascribe material to research questions of interest, with open/inductive coding then being used to generate codes and themes within each broader category.
8. Two coders undertook the analysis of the theory of change data and three coders analysed the evaluation strategy data, with QA checks being conducted throughout by a coder who had not undertaken the initial coding. We used low-inference, content-based coding throughout, to ensure high reliability.
9. The findings of the analysis of the theory of change data are contained within our theory of change for the Government Campus. The findings of the analysis of our evaluation strategy data are contained in a separate linked report, which informed this strategy.
10. Our draft strategy was tested with key stakeholders before it was finalised.
11. Limitations: Due to the 'opt in' nature of the workshop invitations, it is likely that colleagues who are interested in, or recognise the importance of, evaluation as a part of L&D and capability activity were more highly represented than those who were less interested or saw evaluation as less relevant. Therefore, the findings are a systematised account of the current practices, aspirations and ideas of those stakeholders most likely to be engaged with Campus evaluation at the earlier stages; these findings are **not generalisable** across the L&D and capability building field across Government as a whole (and were not intended to be, c.f. our sampling strategy outlined above).
12. Limitations: This was a piece of rapid qualitative research. After an initial round of practice coding by two/three coders to establish common approaches and calibrate decision-making, data was coded by one coder, and 'QA' checked by a second coder, who primarily checked that all raw data eligible for coding had been coded, and checked the homogeneity of codes and themes within our codebooks (i.e. all data coded to that code was highly similar). We did not undertake rounds of coding in parallel, and calculate inter-rater reliability, as would be common in a more extended piece of qualitative research.
13. Limitations: As we lacked the resources for full transcription of the workshop transcripts, we relied on the automated transcript generated by the videoconferencing app. Where content was unintelligible it was manually corrected, but in some instances the best-fit meaning had to be inferred due to transcription which was not fully accurate. If the meaning of the original content could not be established by any means, the text was excluded from coding.

Further details on our research design and data collection and analysis processes can be obtained from our GSCU Campus analytical team.