LMEP-funded Digital Excellence Programme Pilot Evaluation

Evaluation Protocol and Statistical Analysis Plan

February 2025

Government Skills

Contents

Evaluation Summary

Project title	LMEP-funded Digital Excellence Programme Pilot Evaluation
Developer (Institution)	Apolitical
Evaluator (Institution)	ICF Consulting Services Ltd
Principal investigator(s)	Dr Matt Barnard
Evaluation protocol author(s)	Dr Matt Barnard, Ali Zaidi, Laura Campbell, Khin Lin, Robert Wishart, Dr Sergio Sallis, Dr Alice Diaz
Statistical analysis plan author(s)	Robert Wishart
Evaluation type	Pilot randomised controlled trial incorporating an implementation and process evaluation
Evaluation setting	Online
Target group	Civil Servants who are not digital and data specialists and who are Senior Civil Servants (SCS1 and above) or grade 6 and grade 7 Civil Servants
Number of participants	Target sample size: treatment group 50 Civil Servants, control group 50 Civil Servants

Version	Date	Reason for revision
1.2		
1.1		
1.0 [*original*]	January 2025	*N/A*

1 Evaluation protocol

This first section sets out the pilot evaluation protocol, beginning with describing the background to the Digital Excellence Programme (DEP, then setting out the structure of the programme (including its theory of change), before going on to describe the pilot randomised control (RCT) methodology and the methodology of the implementation and process evaluation (IPE). The section finishes by setting out the approach to quality assurance and ethical review before noting the key evaluation stakeholders and the evaluation timeline.

1.1 Background

A 2020 independent report for the Digital Economy Council (Digital Economy Council, 2020) explored gaps and opportunities in the Civil Service Digital, Data and Technology (DDaT) function to achieve the then Government's ambition to make UK Government Digital services the best in the world. The report found that DDaT solutions can have a substantial impact on the delivery of public services, through increasing efficiency by reducing officer time and providing a better public service to the UK population. However, it reported seven barriers to achieving the Government ambition.

One of these challenges was the low levels of technical fluency across Civil Service leadership. The report specifically stated:

> *"Underpinning many of the issues surfaced during this review is a general concern around the relatively under-developed level of digital expertise amongst senior Civil Service leadership. This contrasts with the emerging position in the commercial world in which technology is increasingly seen as a critical delivery lever (alongside people and money) and where it is becoming increasingly expected that senior leaders have a clear understanding of how to deploy technology effectively as an organisational lever. At a minimum leaders should be capable of auditing effectively the performance of their digital functions, including having a realistic expectation of how long projects should take, what they should cost, and what questions to ask in order to assess whether delivery is on or off-track."*
>
> *Organising for Digital Delivery* - Digital Economy Council report, 9 September 2020.

A subsequent 2021 report by the National Audit Office on the challenges in implementing digital change (National Audit Office, 2021) identified a "consistent pattern of underperformance" in the delivery of digital solutions. The report also identified a need for upskilling senior leaders so they are better equipped to identify opportunities for business change and support those implementing change.

> *"Whilst digital leaders bring much needed expertise to the public sector, they often struggle to get the understanding and support they need from senior decision-makers, who lack knowledge in this area."*
>
> Gareth Davies, the head of the NAO.

To address these challenges in improving digital public services, *'Transforming for a Digital Future: 2022 to 2025 Roadmap for Digital and Data'* was developed. The roadmap set out progress towards six missions to transform public services. This included a mission specifically on increasing the use of data in public decision-making, alongside wider reforms on digital transformation and increasing digital skills within the Civil Service. The six missions are listed below.

- **Mission One:** Transformed public services that achieve the right outcomes.
- **Mission Two:** GOV.UK One Login.
- **Mission Three:** Better data to power decision making.
- **Mission Four:** Efficient, secure and sustainable technology.

- **Mission Five:** Digital skills at scale.
- **Mission Six:** A system that unlocks digital transformation.

Concurrently, the use of AI was becoming an increasingly important policy priority for the Government. The 2021 National AI Strategy (UK Government, 2021a) set a bold vision to create a "step-change for AI in the UK, recognising that maximising the potential of AI will increase resilience, productivity, growth and innovation across the private and public sectors". In 2024, the Government published a Generative AI Framework for His Majesty's Government (UK Government, 2024). This provided guidance and processes for the effective use of generative AI, acknowledging the huge potential but also the risks associated with this approach, and the need for guidance to the Civil Service on this emerging technology.

To further support the implementation of the roadmap, and as part of its Declaration on Government Reform (UK Government, 2021b), the Government introduced a Digital Excellence Framework for Senior Civil Servants (SCS) (CDDO, 2023). This set out the minimum skills required for SCS to drive digital transformation and to support the effective use of data and AI. To support SCS to achieve the requirements of the Digital Excellence Framework for Senior Civil Servants, the DEP training programme was introduced, among a range of other initiatives.

1.2 The DEP programme

DEP is an online training programme targeting SCS, with the aim of equipping them with skills related to the use of digital solutions and data. This section details the content of the programme, its target learners, delivery approach, a typical participant's journey and the programme's intended outcomes and mechanisms of change. It concludes a summary of the theory of change.

1.2.1 Research informing the evaluation team's understanding

The research team's understanding of the programme and the development of the theory of change was informed by the research activities described below.

- **A review of programme information in order to understand the programme objectives.** This included:
  - the Digital Excellence Framework;
  - promotion materials and presentations of DEP;
  - the DEP participant guide;
  - Senior Civil Servants (SCS) Digital and Data Skills Benchmark Survey;
  - evaluation of the 'One Big Thing' initiative;
  - a Government Campus presentation on 'Digital and Data Skills for All';
  - a draft skills assessment tool produced by Apolitical to assess the impact of the DEP programme.
- **Scoping interviews took place with 15 key programme stakeholders between August and September 2024, to understand the rationale for the programme and its expected or achieved outcomes.** This included:
  - five members of the Cabinet Office (CO) project team;
  - representatives from the Government Digital and Data Team and the Civil Service Central Digital and Data Office;
  - individuals from the programme developers – Apolitical, Ernst & Young (EY), and the London School of Economics (LSE);

    - individuals from different Government departments, including His Majesty's Revenue and Customs (HMRC), the Department for Work and Pensions (DWP) and the Department for Science, Innovation and Technology (DSIT).

- **A review of the DEP course.** After receiving a demo license from Apolitical, ICF reviewed the course to explore the activities and tasks associated with the programme. This further deepened our understanding of what the programme was aiming to teach.

The findings from this research are presented in the rest of this section.

1.2.2 Overview of the programme

DEP comprises four asynchronous training modules.

- Building a Digital Culture in Government (five hours in length).
- Building a Data Culture in Government (five hours in length).
- Building a User-Centred Culture in Government (three hours in length).
- Building AI Confidence in Government (four hours in length).

Each module is supplemented by a one-hour masterclass. These sessions can be attended 'live', but individuals who cannot attend the live sessions can also watch recordings of the masterclasses. There is also a fifth element, which is an innovation masterclass that aims to support senior leaders to lead innovation in their teams. Each module is designed to be standalone and therefore can be undertaken in any order. However, if a participant plans to attend all the masterclasses live then they would need to complete the modules in the order they are listed in above. There are no other activities that are mandated as part of the programme. However, a number of stakeholders reported that in some cases attendees from the same Government department came together to undertake the modules over a similar period and met to discuss the content.

The programme was developed by the LSE against the Digital and Data Excellence for Senior Civil Servants framework (ibid). Online content and teaching materials were then developed by Apolitical. This process was overseen by Ernst & Young. After a development and piloting phase the programme was launched in 2023.

1.2.3 Who the intervention is for

The programme is targeted at SCS across all Government departments. These individuals are responsible for strategy, advising on policy and overseeing the implementation of Government policies and initiatives. They often work directly with Ministers. There are around 7,000 SCS among the overall Civil Service workforce of nearly half a million. The programme is also open to Civil Servants aspiring to be SCS. This primarily includes Civil Servants at Grades 6 and 7. These are experienced Civil Servants who often have responsibility for leading workstreams, and work with executive committees, SCS and Ministers.

DEP is designed to meet the needs of SCS or aspiring SCS regardless of their pre-existing knowledge of digital or data. There are no pre-entry criteria in terms of previous skills, knowledge or qualifications. Likewise, the training is not targeted at specific learning levels. The DEP portal maintained by Apolitical does have a diagnostic tool that allows participants to assess their prior knowledge so they can prioritise which of the four modules they take. Additionally, some departments may include their own eligibility criteria for the programme. For example, in the scoping interviews, one department reported conducting a survey of SCS, where those who rated themselves five on a five-point scale for certain criteria were advised that they do not need to undertake the programme.

1.2.4 How the programme is delivered

Each Government department is expected to identify nominees for the training. Each has a Single Point of Contact (SPOC) for the programme that is responsible for collecting this

information. Nominee information is shared with Apolitical who then send login details that allow attendees to access the course content through a dedicated portal.

Nominees for the training are grouped into cohorts. This is so that they can be assigned to specific masterclasses that are run in a group format. However, participants do not need to attend these masterclasses live to complete the programme. The live masterclasses are typically run over a relatively condensed period (within two months). It is expected that individuals complete each online module prior to the live masterclasses. However, participants have 12 months from enrolling to complete the programme. There is a license fee of £500 for each registration.

The programme has been running since February 2023. Take up has been reasonable, with 1,081 enrolled on the programme by August 2024. However, only 470 of those started the programme and only 240 had completed it at the time of writing. The small proportion of participants starting and completing the programme was attributed by some interviewees as being due to SCS having a busy workload and therefore struggling to make time to undertake the training.

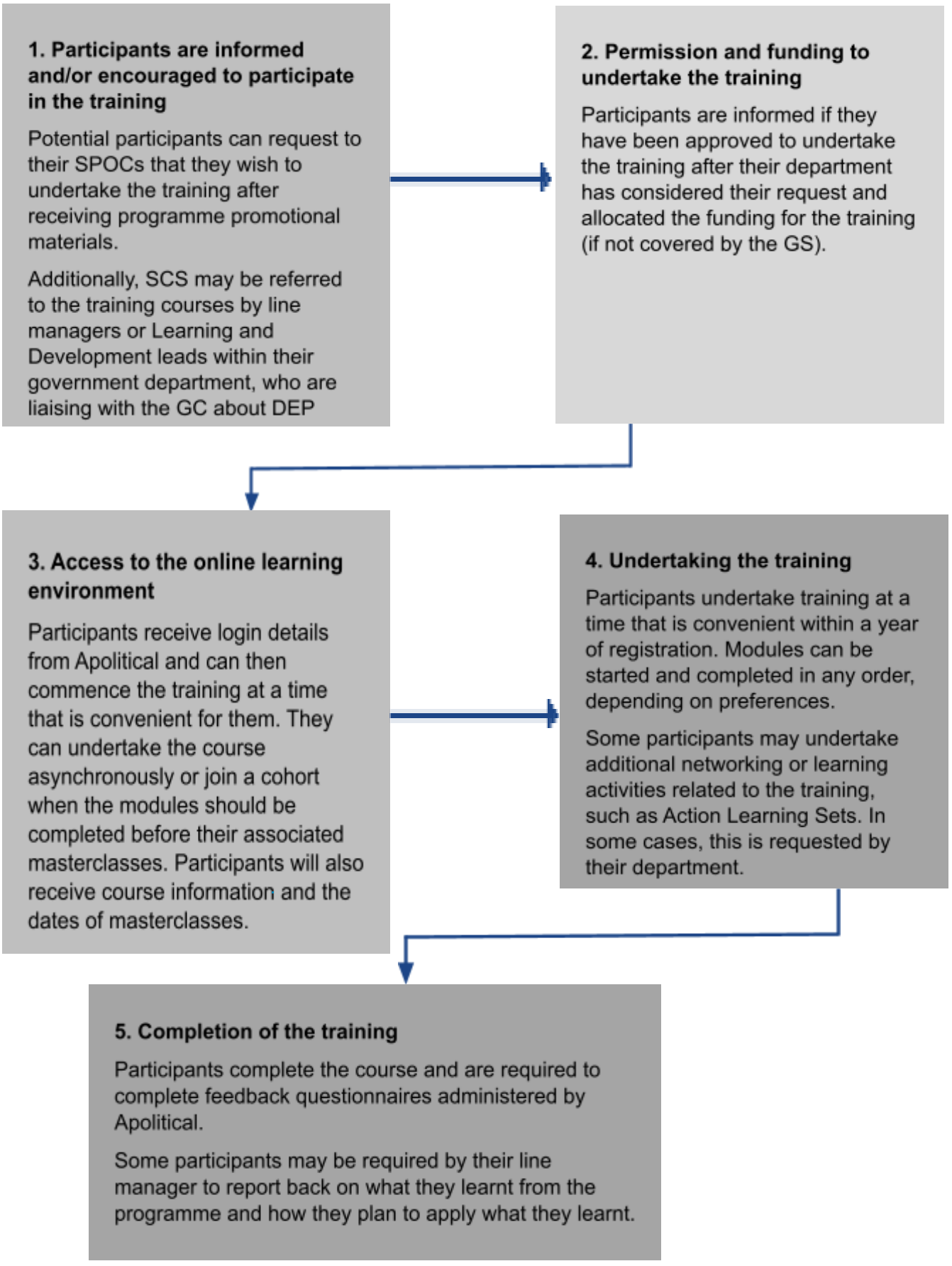
1.2.5 Inputs for the programme

The interviews and desk research identified a wide range of inputs and resources that support the delivery of the programme. These include:

- the development of a curricula by the LSE and Apolitical that is mapped against the Digital and Data Essentials Framework for Senior Civil Servants;
- the maintenance of the online portal and registration of new participants by Apolitical;
- the promotion of the programme through the dissemination of promotional materials by Government Skills (GS);
- the promotion of the programme within Government departments by senior leaders and SPOCs;
- the financial contribution by Government departments or the CO to pay for the licenses for the training;
- the opportunity costs for Government departments of releasing staff to undertake the training;
- the GS to monitor performance and take-up and ensure feedback and lessons learnt are fed into the on-going improvement of the programme;

1.2.6 Typical participant journey

The typical journey of a participant on the DEP course is presented below.



1.2.7 Anticipated DEP outcomes

1.2.7.1 Short-term outcomes

In the scoping interviews, representatives from the developer organisations (Apolitical, EY and LSE) and CDDO both felt the short-term outcomes from the programme are largely defined by the Digital and Data Essentials for SCS, which formed the basis of the training. However, they noted that the DEP programme did not cover all the requirements of the Digital and Data Essentials for SCS.

A review of the course documentation found that four of the five modules of the DEP programme directly relate to four of the five sections of the Digital and Data Essentials for SCS (see Table 1 below). The only section of Digital and Data Essentials for SCS that was not related to a specific module in DEP was technology essentials. DEP also included a module on AI that was not referenced explicitly in Digital and Data Essentials for SCS.

Most of the outcomes listed in the Digital and Data Essentials for SCS guidance are short-term in nature, since they relate to knowledge (starting with 'I understand', 'I know'), attitudes ('I uphold') or skills ('I can') that are acquired immediately from the training. However, some relate to the application of what participants learnt ('I demonstrate', 'I show') that are therefore more likely to be medium-term outcomes (highlighted in **bold** below).

Table 1: Digital and Data Essentials for Senior Civil Servants guidance

Data essentials	Digital essentials	User first	Innovation mindset
I understand the value of good quality data, support ethical data practices and know how to mitigate bias in decision making. **I can demonstrate the use of data in evidence-based decision making, using techniques for analysis and interpretation (such as data visualisation).** I uphold data standards and ensure data protection best practice and regulations (including the Data Protection Act) are followed.	I understand my organisation's digital strategy and objectives, and I know which digital services it provides. I can articulate the benefit of agile ways of working. I support agile methods through my leadership.	I understand that users are at the heart of what Government does. I encourage user research to understand user needs. I understand that Government digital services must be carefully designed to meet user needs. I recommend researching user journeys to provide a joined-up experience. I understand that public sector services must be accessible to all users. I ensure that accessibility regulations and best practice are followed.	**I demonstrate innovation in the design and development of my service, operation and policy. I encourage others to pursue relevant digital, data and technology opportunities.** I know how to access and work with experts to meet the demand for digital services. I can manage the resources required and decide if and how to outsource. **I show the benefits of a multidisciplinary approach by working with a range of teams and experts at the earliest opportunity to develop high quality digital services.**

In the scoping interviews, stakeholders also identified some additional short-term outcomes of the programme. The most commonly reported outcome was related to a change in attitudes, specifically participants **increasingly embracing the use of data and innovation** for supporting public services. This was, in part, felt to be due to participants also having **more confidence in being able to drive forward digital and data solutions.**

For the AI programme, some Civil Servants who had been on DEP or had sent staff to it reported that the programme led to an **increased understanding of key AI concepts**, such as large language models and generative AI. Some participants reported that they also learnt about ethics and when to use AI tools effectively.

1.2.7.2 Medium-term outcomes

The medium-term outcomes relate to the behaviour changes that would largely be expected from participants of the programme after they have applied what they learnt from the programme. The Digital and Data Essentials for SCS identified outcomes related to:

- the increased use of data in evidence-based decision making, including techniques for analysis and interpretation (such as data visualisation);
- employing innovation in the design and development of services, operations and policies, including providing greater encouragement for pursuing relevant digital, data and technology opportunities;
- employing a multidisciplinary approach through collaboration across a range of teams and experts at the earliest opportunity to develop high quality digital services.

In scoping interviews, it was commonly reported that a key output of the programme was for participants to be able to **'ask the right questions'** in order to challenge departmental plans. They noted that SCS are unlikely to be involved in the direct delivery and management of digital programmes, and therefore their key involvement is in being able to challenge projects to ensure they are being developed in a way that maximises the user experience and makes effective use of digital solutions and data.

Some staff in Government departments also reported that they believed the programme would enable SCS to have **more realistic plans and expectations for new programmes**. They note that SCS often set the agenda for digital solutions in key policy areas, and in some cases they can have unrealistic expectations as they do not understand some of the technical or digital limitations inhibiting what they can do.

Nearly all stakeholders also highlighted that a key aim of the programme was to develop SCS so they could be **'agents for change'** in their department. They noted that a key reason the programme targets SCS is because it was felt that they could play an important role in changing the cultures that underpin the effective use of data, digital solutions and innovation within the Civil Service. This suggests the programme is likely to achieve a multiplier effect, as the programme will not only change the behaviours of participants but also other staff in their teams.

1.2.7.3 End outcomes

The scoping interviews identified a wide range of potential long-term outcomes from the programme, which included:

- individual level productivity improvements;
- organisational/departmental level productivity improvements;
- benefits for the public.

In terms of individual productivity improvements, some interviewees noted that the changes in behaviours would likely lead to **fewer issues and delays with projects led by SCS, which would provide a time saving**. This was a result of better understanding the role and limitations of digital solutions, being better able to interpret data, and a greater appreciation of the role of AI and balancing the need of user experience with other Government priorities.

A few interviewees also noted **that more effective leadership of projects could result in a cost saving**, as there would less likely be cost overruns to large projects and the budgets set for activities were reasonable and appropriate. However, they felt it would be difficult to identify the extent to which this would be due to SCS, as it would also be affected by staff involved in the day-to-day delivery of the projects.

Stakeholders identified many organisational benefits that could also arise from SCS having better data and digital skills. This included:

- more effective use of digital solutions in Government departments, which would result in a **reduction in staffing needs** as the online services would reduce the need for staffing to deal with queries or to provide guidance;
- **better staff morale** as staff feel more supported in delivering digital solutions and are less likely to experience problems in their projects;
- i**mproved digital culture** among department teams, which means that individuals are more likely to identify innovative new digital services.

It was also felt that the public would be a beneficiary of the programme. It was agreed that there remained substantial areas for improvement in the Government's digital services, which they expected the programme would help address. These improvements would lead to the benefits described below.

- **Government service users having to spend less time using digital services.** Central to this is a greater use of initiatives that follow the structure of 'Tell Us Once' (Gov.uk, n.d.). This initiative for when a person dies requires the next of kin to enter information and documentation in one digital portal, which is then shared and used by other services such as pensions, HM Land Registry, the Driving Vehicle Licensing Agency and others to update their records.
- **More people being able to use Government digital services to access or provide information.** Some interviewees noted that a lack of clear information and difficult navigation through sites can deter some individuals from using digital tools, particularly those who may lack 'digital confidence'. By focusing these services on the user experience, it was expected they would be more accessible.
- **Fewer user errors in using Government services**. A number of interviewees also noted that the lack of clarity in Government digital solutions can make it more likely that users enter the wrong information. This can have consequences on Social Security payments and access to other services.

1.2.8 Mechanisms of change

This section discusses the mechanisms by which the programme is expected to achieve its outcomes. Describing the mechanisms is important because it provides a framework for understanding if the programme is working as intended and helps identify the barriers and facilitators to the programme's impact. The section first describes underlying approach and assumptions underpinning the design of the programme and providing the context for its mechanism, which are set out in the second section.

1.2.8.1 Underlying approach: Andragogy

Andragogy relates to an approach to learning that is seen as appropriate for adult learners, which is reflected in both the design and delivery of the training and the need for learner engagement with the programme (Knapke et al, 2024). The key assumptions underlying this approach discussed below are drawn from a review of programme documentation alongside literature pertaining to effective adult learning (EEF, 2021; Society of Education, n.d.; DfE, 2011).

- **The course provides engaging learning material**, including:
    - employing an appropriate mix of training methods (e.g. mixing presentations with group work or self-reflection and activities);
    - clarity and quality of the course materials;
    - masterclass speakers' expertise and knowledge of the subject;
    - fidelity to the Digital Essentials for SCS framework;
    - appropriate pace and challenge to the programme;
    - sufficient depth of information to enable SCS to understand the topics well enough to change behaviour;
    - applicability to SCS roles.
- **Participants are sufficiently engaged with the training**. This involves participants actively engaging in the content, undertaking the self-directed learning tasks and undertaking the learning at a time and pace that maximises the value they gain from the programme.
- **The course is relevant to the SCS role and experiences**. Many interviewees note that SCS roles vary greatly, and therefore it would not be possible for all the programme content to be completely relevant to an individual's role. However, the course needs to be structured in a way that participants can see how it applies to their role.
- **Participants are motivated and have the capacity and opportunity to apply what they learnt**. In order to lead to outcomes, it is essential that learning is applied. There can be barriers to doing this, which can stem from participants being unwilling to change their existing practice, or not having the time to do so.

1.2.8.2 Mechanisms

While the general approach of the course provides the assumptions about the necessary conditions for DEP to be effective, the mechanisms described in this section are the ways in which it is believed that the programme's activities (self-directed online learning plus masterclasses) will lead to changes in knowledge, skills and attitudes (the programme's most immediate outcomes). The mechanisms presented below have been identified by the research team based on a review of the material provided about the programme, but they were not directly articulated by the course delivery organisation.

- **Provision of new information:** for some participants the courses may include information they were not previously aware of and therefore the exposure to this material helps them gain new knowledge of the subject matter. This mechanism reflects a knowledge-deficit model of science communication (Grand, 2023), which can be characterised as assuming that learners are akin to 'empty vessels' to which knowledge is added.
- **Information presented in a novel or clearer way:** a second way in which the programme can lead to the immediate outcomes is through presenting information in a way that participants have not seen before and that is more effective in helping them gain an enhanced understanding of it. Unlike within the knowledge deficit model, learners use the course content to help create for

themselves a deeper and more holistic understanding of the topics. They may experience rapid shifts in understanding that are based on a holistic understanding of the topic, reflecting the Gestalt model of learning (Vitello and Salvi, 2023).

- **Experiential learning:** this relates to the application of what participants learnt to their practice (Kolb et al). The programme provides scope for experiential learning both through the exercises and quizzes that are part of the online programme, and through learners applying what they learn between modules. However, the extent to which this is achieved depends on participants being willing to apply what they learnt from the programme and to reflect on their experiences of applying it.
- **Information presented in a more powerful or engaging way:** even if the course material is not new or does not lead to greater cognitive understanding, it may be more vivid or engaging, which can have the effect of being more persuasive and therefore change a participant's orientation or attitudes to the subject. Research indicates that a range of factors are important to changing attitudes, including source credibility, which is clearly present for DEP as the programme has been designed by a well-known and high-status organisation (LSE) (Belch & Belch, 2012). In addition, the course materials reflect the central information processing route of the elaboration-likelihood model of persuasion, which works through stimulating careful consideration of the material and in which attitude change depends on the quality of the arguments presented. This is as opposed to the peripheral route, which does not encourage careful processing and in which attitude change depends on persuasive cues rather than the arguments. (Petty and Cacioppo, 1986).
- **Social influence:** group-based learning is an element that some participants potentially benefit from when they are placed in cohorts. This means their learning can be reinforced and enhanced through discussion with peers (Feldon et al, 2019), both in terms of understanding and potentially attitude change, through the effect of creating new social norms. However, as the learners in the pilot evaluation have not been placed in cohorts, this mechanism is not expected to play a major role for the current participants.

These mechanisms can be categorised using the COM-B model (Michie, van Stralen and West, 2011). The first three mechanisms (provision of new information; information presented in a new way; experiential learning) primarily relate to the **capability** element of the model; the second two (material presented in a more powerful way; social influence) primarily relate to the **motivation** element of the model. Based on the COM-B model, the changes to motivation and capability should directly lead to the medium-term outcomes (behaviour change) which in turn lead to the end outcomes relating to productivity.

1.2.9 Theory of Change

The DEP Theory of Change is summarised in the figure below (Figure 1).

Figure 1: DEP Theory of Change

Need: Transforming for a digital future: 2022 to 2025 roadmap for digital and data' sets out an ambitious Government objective to create a 'Digital First' culture where an increasing number of public services being available online. However, the quality of Government online services and interconnections between different departments in the sharing of data is variable. A 'Report from the Digital Economy Council' and a subsequent 2021 report by the National Audit Office on the 'Challenges in implementing digital change' both identified Senior Civil Servants' digital and data skills as a substantial barrier to achieving this ambition. The Digital Excellence Programme was developed to address this gap by supporting participants to achieve the standards set out in the Digital Excellence Framework for Senor Civil Servants

Inputs	Activities	Mechanisms	Outcomes	Behaviour change	End outcomes
Content development by LSE and Apolitical Apolitical management of learning platform Promotion & management by GS Promotion & coordination by SPOCs Govt financial contribution to cost of programme Staff time for participating in training	**Online self-directed learning modules on:** 1. Building a Digital Culture in Government 2. Building a Data Culture in Government 3. Building a User-Cantered Culture in Government (three hours in length) 4. Building AI Confidence in Government (length not specified) **Masterclasses** Four 'live' one-hour masterclass linked to the online modules. Additional innovation masterclass. **Learning cohorts** Networking and peer-learning undertaken as participants are grouped into learning cohorts with a set start date.	**Underlying andragogy** • Effective and engaging learning • Participants are sufficiently engaged in the programme • Course reflects SCS role and experiences • Participants are motivated and have the capacity and opportunity to apply what they learnt **Mechanisms** • Provision of new information: • Information presented in a novel or clearer way • Experiential learning • Information presented in a more powerful or engaging way • Social influence	**Knowledge** • Understand what is good quality data, ethical data practices and minimising bias • Understand organisation's digital strategy and digital services • Understanding of key AI concepts • Understand digital services must be user-centred and accessible **Skills** • Can uphold data standards and ensure that regulations are followed • Can articulate the benefit of agile ways of working • Manage resources required and decide when to outsource. • Know how to access and work with digital experts • More confidence in being able to drive forward digital and data solutions **Attitudes** • Ensure data best practice is followed • Supports agile methods through leadership • Understand centrality of users • Embracing the use of data and innovation	Use data in evidence-based decision making, including techniques for analysis and interpretation (such as data visualisation). Innovate in the design and development of services operation and policy. Encouraging team members to pursue relevant digital, data and technology opportunities Employ a multidisciplinary approach to develop digital services. Providing better interrogation of digital policies and projects by 'asking the right questions' Setting more realistic plans and expectations for new programmes, Being 'agents for change' in terms of creating a digital first and data culture in Government	**Participant productivity improvements** • Time saving due to fewer issues and delays with projects led by SCSs • Cost saving from fewer project cost overruns and more appropriate budgets **Organisational/ Department level productivity improvements** • More effective use of digital solutions reduce staffing needs for core services • Better staff morale as staff feel more supportive in delivering digital solutions • Improved digital culture **Benefits to public** • Service users having to spend less time using digital services • More people being able to use Government digital services • Fewer user errors in using Government services

1.3 Evaluation design

This section first sets out the evaluation's research objectives, before describing the pilot RCT method and then the method of the IPE.

1.1.1 Research objectives

The pilot evaluation has four objectives, which relate both to the process of conducting evaluation within the context of Civil Service Learning and to gaining an understanding of the nature and effectiveness of the DEP. These are to:

- assess the acceptability of conducting an RCT;
- assess the feasibility of conducting an RCT;
- assess evidence of promise of the Digital Excellence Programme;
- build capacity for future evaluations of this programme and other learning programmes.

1.1.2 Pilot RCT method

1.3.1.1 Trial design

The primary research question addressed by the pilot trial in relation to outcomes is:

> Is there evidence of promise for an online self-paced digital, data and AI skills programme (with additional real-time online masterclasses) on the behaviour and attitudes of Senior Civil Servants and grade 7 and 6 Civil Servants who have not previously undertaken the programme, compared to business as usual?

The evaluation will be run as a pilot two-armed RCT. Randomisation will occur at the level of the individual Civil Servant and will be in the ratio of 1:1. The trial is defined as a pilot as the objectives of the evaluation are to explore the acceptability and feasibility of undertaking an RCT, and so the trial has not been designed with a minimum detectable effect size in mind. A two-arm design was chosen as it is the most straightforward and is therefore appropriate for assessing whether an RCT is feasible and acceptable. A similar logic applies to the choice of the allocation ratio and to randomising at an individual level.

The primary outcome will be attitudes, behaviour and intended behaviour as measured by a bespoke survey designed for the evaluation (see below for more details on the outcome measure). A summary of the pilot design is set out in Table 2 below.

Further work is being undertaken to identify indicators that can be used to assess the impact of the programme on productivity, both as part of this pilot evaluation and in terms of evaluation work that may be carried out in the future.

Table 2: Summary of pilot trial design

Trial design, including number of arms		Two-arm randomised controlled trial
Unit of randomisation		Individual Civil Servants
Stratification variables (if applicable)		None
Primary outcome	variable	Behaviours and attitudes
	measure (instrument, scale, source)	Bespoke survey
Secondary outcome(s)	variable(s)	None
	measure(s) (instrument, scale, source)	N/A
Baseline for primary outcome	variable	Behaviours and attitudes
	measure (instrument, scale, source)	Bespoke survey
Baseline for secondary outcome	variable	None
	measure (instrument, scale, source)	N/A

1.3.1.2 Outcome measure

The evaluation team undertook a rapid review of potential outcome measures that involved investigating competency-based measures, including the Europass digital skills assessment, which is available online, commercial tests (for example, those provided by the company TestGorilla), a diagnostic tool developed by Apolitical, the SCS digital and data skills self-assessment survey, digital and data self-efficacy measures, and approaches to measuring productivity in the public sector. In addition, Government stakeholders were consulted, as was an external academic expert. These investigations indicated that there do not seem to be any existing

tools or measures that are suitable for measuring the outcomes of DEP. Therefore, it was agreed that the evaluation team would develop a bespoke survey that focussed on attitudes, behaviours and behavioural intentions, such as:

- these outcomes can be collected directly from evaluation participants;
- a survey tool to collect these outcomes can be tailored to DEP; and
- these outcomes have a close connection with productivity, the ultimate intended impact of the programme.

In developing the survey, the research team drew on the Digital and Data Essentials for SCS guidance, the SCS digital and data skills self-assessment survey, and the diagnostic tool developed by the programme developer Apolitical. The order of items in the survey was randomised to minimise any primacy or recency effects (i.e., systematic differences in the way people respond to items depending on whether the item is at the beginning, or end or in the middle of the survey). There are three groups of questions covering the following topics: data culture, digital culture and AI confidence. The score ranges by topic are set out below.

- **Data culture**: scored 12-26.
- **Digital culture**: scored 10-50.
- **AI confidence**: scored 7-45.

The Theory of Change does not indicate whether any of these topics are more or less important. For the item response theory (IRT) analysis, each sub-scale will be assessed separately based on the unweighted scores. For the impact analyses, the scales will be transformed into z-scores and then re-scaled with a mean of 50 and SD of 5. The weighted sum of the three sub-scales will be the total score used in the primary analysis.

Before the start of the evaluation, the survey was reviewed by key internal stakeholders and piloted with a small number of Civil Servants who have previously undertaken the DEP. Following baseline data collection, IRT analysis will be conducted to assess whether the instrument can be strengthened by adding, removing or adapting individual items, and if so a revised version used for the follow up data collection. Further information on the IRT analysis approach is detailed in the statistical analysis plan (section 2 below).

1.3.1.3 Sample size

Although the trial was always envisaged as a pilot, originally the intention was to recruit sufficient participants for it to be powered to detect the programme's impact. However, due to practical and logistical issues related to the delivery context, it became clear that the required number of participants could not be recruited. Therefore, it was agreed with the trial sponsor, the Cabinet Office, that the aim was to recruit a minimum sample of 100 Civil Servants, with 50 in the intervention group and 50 in the control group. In practice, a marginally larger sample size was recruited (157 Civil Servants). Of these, 78 were allocated to the treatment group and 79 to the control group.

Power calculations were subsequently undertaken to ascertain the Minimum Detectable Effect Size (MDES) that the trial was powered to detect. The calculations were conducted in PowerUp! (Dong and Maynard, 2013) for an individual random assignment design, using the following assumptions:

- a type-one (false positive) error rate of 0.05;
- a type-two (false negative) error rate of 0.20 (synonymous with power of 0.80);
- two-tailed statistical significance testing;

- a 1:1 allocation ratio between intervention and control;
- the variance in the outcome is explained by the baseline measure of the outcome, with an R2 of 0.20;
- attrition between baseline and endline of 10%.

Based on these assumptions, the calculations indicate that the trial would be powered to detect a MDES of 0.47.

The final evaluation report will include further power calculations, informed by the learning from this pilot RCT, outlining the sample size a future trial would need in order to achieve an appropriate MDES. This will take into account the variance in the outcome explained by the same measure at baseline (R2), the sensitivity of the outcome measure, likely rates of attrition and indicative evidence about potential effect sizes (albeit that the analysis conducted in the pilot RCT is underpowered).

1.3.1.4 Participant recruitment

Participants were recruited through a range of approaches:

- Civil Servants who have licenses to participate in DEP (are enrolled), but have not yet started the programme, received an email and information sheet from Departmental SPOCs for DEP, and a message will be integrated into the DEP page of Apolitical's website, inviting Civil Servants to take part in the evaluation;
- Civil Servants who were not yet enrolled into DEP but wanted to take part (or their line manager or department wanted them to take part) were invited to participate in the evaluation after their enrolment, using the same processes mentioned above;
- a selection of Civil Servants in the target population for DEP were offered 'free licences' (i.e., the licences would be paid for centrally by Cabinet Office and the Central Digital and Data Office (CDDO) rather than the Civil Servants' home departments).

1.3.1.5 Randomisation

Once participants completed the baseline survey, their names and email addresses were automatically fed into an Excel file. The trial statistician then used Stata to randomly allocate each individual into intervention or control, with the restriction that the overall allocation ratio was as close to 50:50 as possible. The research team then emailed the participant to let them know which group they were in and emailed CDDO with the contact details of those Civil Servants in the intervention group so that the DEP onboarding process could begin. This process meant the SPOCs, Apolitical and Civil Servants were not aware of the allocation to intervention and control until after all baselining activities were completed.

1.3.1.6 Business as usual

Participants assigned to the control group were asked to not start any of the modules until the evaluation intervention period has finished and were instead assigned to start the learning as part of the subsequent cohort, which will begin after the collection of the follow up data (see evaluation timeline in section 1.7).

1.3.1.7 Consent and data collection

Outcome data will be collected via an online, self-completion survey. Participants who have opted into the evaluation were sent a link to the baseline survey at the start of the evaluation or were directed to it via a link on Apolitical's website. Before the start of the survey, participants were provided information on the evaluation, the randomisation process and its implications for when they will undertake DEP, the data collection process, confidentiality and data protection. They were also provided with contact details for a member of the research team and for the Cabinet Office project lead, in case they wished to find out more about the study. Participants were asked to confirm their consent to taking part in the evaluation before being able to start the survey. Only participants who gave their consent to be part of the evaluation and who completed the baseline survey were randomised and therefore formally entered the trial.

The survey will be administered to participants again at the end of the evaluation intervention period. The survey will comprise the same attitudinal and behaviour questions that were asked in the baseline survey (contingent on any adjustments as a result of the IRT analysis), as well as additional questions about their experience of DEP that will form part of the implementation and process evaluation.

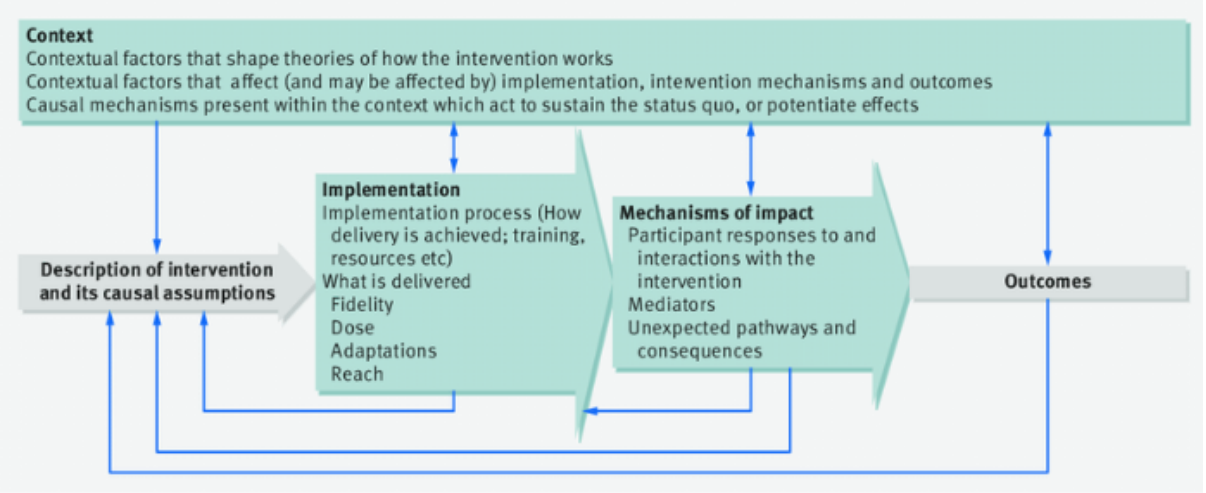
1.3.1.8 Participant retention

Participant retention is important to the validity of the trial. As Civil Servants have busy roles, they can be a challenging group to collect data from. Therefore, the research team is drawing on behavioural science to encourage participants to complete the follow up data collection. When participants are sent reminders to complete the follow up survey, the email will include a graphic indicating what percentage of participants have finished the survey. The graphic will also indicate a target of 90%, which is the level of completion that is seen as important to ensuring the evaluation is a good use of public funding.

1.1.3 Implementation and process evaluation methodology

The IPE will be aiming to assess the acceptability and feasibility of conducting an RCT as well as providing insight into the intervention mechanisms and any unintended positive or negative outcomes. IPE will be informed by Medical Research Council (MRC) guidance for undertaking process evaluations of complex interventions and case study research (Moore et al, 2015; Crowe et al, 2011). The guidance is useful in providing a clear structure for process evaluations that reflects important elements of intervention implementation and the factors that influence outcomes (as illustrated in Figure 2 below).

Figure 2: Key functions of process evaluation and relations among them



Source: Moore et al, 2015

1.3.1.9 IPE research questions

The IPE research questions are set out below grouped under four themes: context, implementation, mechanisms of impact and outcomes.

Context

- How do SCS[1] vary in terms of their need for digital skills learning?
- What is the range and nature of support available for digital tasks?
- What are the range and nature of previous digital learning experiences?
- How does the environment of SCS influence their ability to take part in and benefit from the programme?

Implementation

- What are the range of experiences of SCS of the pilot RCT and what factors influence their experiences?
- What are the range of experiences of SCS on the programme (online content and masterclasses) and what factors influence their experiences?

[1] The pilot trial of DEP included both SCS and G6& G7 Civil Servants, who are potential future SCS. To avoid undue repetition, in this section wherever SCS are being referred to, the G6 & G7 are included.

- What approaches to undertaking the programme and learner behaviour have proved more effective or less effective in supporting participants to achieve positive outcomes?
- What lessons can be learnt from the implementation of the programme and evaluation, including the measurement of outcomes, for future evaluation and outcome measurement of DEP and similar programmes?

Mechanisms of impact

- To what extent are the causal mechanisms influencing programme delivery?
- How do contextual factors affect the extent to which the causal mechanisms are influencing the programme outcomes?
- Are there any other causal mechanisms that affect the achievement of outcomes?

Outcomes

- To what extent have the anticipated outcomes of the programme been achieved?
- Are there any unanticipated positive outcomes for SCS from the programme?
- Are there any unanticipated negative outcomes for SCS from the programme?

1.3.1.10 Overview of methods

To examine the research questions, a mixed methods approach will be taken. This facilitates the triangulation of information from programme data, participant views, delivery partners and the views of other stakeholders such as SPOCs that understand the programme and work alongside SCS, as well as the CDDO that work across departments to improve digital and data skills. The sub-questions to assess the key research questions and evidence sources are presented in the IPE Analytical Framework presented below (Table 3).

Table 3: DEP IPE Analytical Framework

Research questions	Sub-questions/ indicators	Evidence sources
Context		
How do SCS vary in terms of their need for digital skills learning?	• Participants' prior skills in, and attitudes to, digital technologies • Training undertaken by participants to address digital skills needs • Participants' involvement in digital services in the last two years • Perceived skills gaps of SCS on digital technologies	Participant interviews SPOC focus group Delivery partner research CDDO team group discussion Participant survey
What is the range and nature of support available for digital tasks?	• Participant role in supporting digital services in their department • Other training and support provided to SCS	Participant interviews SPOC focus group CDDO team group discussion
What are the range and nature of previous digital learning experiences?	• Understanding of other digital training opportunities within the Civil Service • Perceived limitations of other digital training programme • Barriers that affect SCS access to training	Participant interviews SPOC focus group CDDO team group discussion
How does the environment of SCS influence their ability to take part in and benefit from the programme?	• Current time spent by SCS training • Extent to which digital skills are a priority to SCS, in comparison to other skills areas • Perceived value of training for personal advancement	Participant interviews SPOC focus group
Implementation		
What are the range of experiences of SCS of the pilot RCT and what factors influence their experiences?	• Completion of the trial post-completion survey • Completion of modules during the trial period • Participants' understanding the value of the RCT • Perceptions of the clarity and quality of the recruitment information • Views on the administrative burden of participating in trial	Apolitical administrative data[2] Participant interviews Participant survey Delivery partner research

[2] This is dependent on whether the data can be provided by Apolitical

Research questions	Sub-questions/ indicators	Evidence sources
	• Views on the impact of having to defer entry to the programme • Whether they would be willing to participate in future RCTs	
What are the range of experiences of SCS on the programme and (online content and masterclasses) and what factors influence their experiences?	• Perceptions of the quality of the modules and masterclasses • Perceived value of the asynchronous modules • Accessibility and value of the live masterclasses • Proportion of enrolled participants that have completed modules and attended masterclasses • Length of time between start and completion of the programme	Apolitical administrative data[3] Participant interviews Participant survey Delivery partner research
What approaches to undertaking the programme and learner behaviour have proved more effective or less effective in supporting participants to achieve positive outcomes?	• Comparative analysis exploring prevalence of outcomes by participant characteristics and by pace and level of module/masterclass completion • Views of participants of the benefit of different approaches for undertaking the programme • Description of learner behaviour including proportion of programme completed, time to complete, order of modules undertaken, learner behaviour on the platform (e.g. spending the intended time on tasks or not), attrition from the course, attendance of the masterclasses.	Participant interviews Participant survey Delivery partner research Apolitical administrative data[4]
Mechanisms of impact		
To what extent are the causal mechanisms influencing programme delivery?	• Perceptions of the effectiveness of the programme, in terms of the quality of materials, the mix of learning methods, relevance of content, depth of content and pace of learning. • Level of participant interest and engagement in the programme • Perceived relevance of the programme to SCS role	Participant interviews Participant survey Delivery partner research

[3] As above

[4] As above

Research questions	Sub-questions/ indicators	Evidence sources
	• Scale and quality of networking taking place for SCS participating in cohort model • Degree to which the programme is providing new material, material presented in a new way or more engaging way and degree to which it facilitates using knowledge	
How do contextual factors affect the extent to which the causal mechanisms are influencing the programme outcomes?	• Wider departmental and programme factors that affect the quality and relevance of the programme and ability to network.	Participant interviews SPOC focus group
Are there any other causal mechanisms that affect the achievement of outcomes?	• Views of wider departmental factors (e.g. leadership support, time pressures, opportunity to take time for training) that support or inhibit participation in the programme and achievement of outcomes • Views on support provided by GS and CDDO that support or inhibit the achievement of outcomes • Views programme factors that affect the achievement of outcomes	Participant interviews SPOC focus group
Outcomes		
To what extent have the anticipated outcomes of the programme been achieved?	• Perceived outcomes gained by participants from the programme, in terms of: • Learning (new content and information) • Behaviours (changes to their engagement and contribution to digital projects, their use of data and AI) • Department operations and productivity (tangible changes to the digitalisation of services, the development of new products, and the use of data for strategy and planning)	Participant interviews SPOC focus group Participant survey
Are there any unanticipated positive outcomes for SCS from the programme?	• Perceptions of further outcomes related to personal and organisational development resulting from the programme.	Participant interviews SPOC focus group

Research questions	Sub-questions/ indicators	Evidence sources
Are there any unanticipated negative outcomes for SCS from the programme?	• Views on issues arising from the application of learning from the programme.	Participant interviews SPOC focus group

1.3.1.11 Step-by-Step methodology

This section presents more detail on the research tasks that will be conducted as part of the IPE.

Analysis of Apolitical administrative data

With Apolitical's consent, the administrative data gathered on those who enrol on the DEP will be analysed, including if possible the following data on learners:

- when they first enrolled on the programme;
- when they began participating in the programme;
- their attendance/completion status for each of the modules/masterclasses;
- date modules/masterclasses attended/completed;
- participant characteristics, including gender, age, department;
- whether they participated in the RCT or not.

The analysis of the administrative data will help us analyse how participants experience the programme and the RCT. This will help understand participation and engagement in the programme.

Follow-up survey of participants

Alongside the outcome data on outcomes (attitudes and behaviour), the follow-up survey of both intervention and control participants will include a range of questions relevant to the IPE. The survey will go to all participants in the evaluation, and will include questions on the following topics:

- attitudes towards the pilot RCT;
- experience of taking part in the pilot RCT;
- attitudes towards the programme;
- experience of the programme (intervention participants only);
- modules taken (intervention participants only);
- experience of each module;
- perceived impact of the programme (intervention participants only).

The survey will comprise mainly closed questions, primarily using Likert scales as answer categories and will enable exploration of the prevalence of certain attitudes, perceptions and behaviours. More in-depth, qualitative information will be collected from the qualitative interviews (see section below).

Qualitative data collection

Qualitative data will be gathered from seven groups of stakeholders. The stakeholder groups, number of encounters, mode of data collection and the focus of the interviews are set out in Table 4 below.

Table 4: Overview of qualitative data collection

Participant group	Number of encounters	Data collection mode	Key topics
RCT – intervention participants	8	Individual	Experience of the programme Perceived immediate impact of the programme Barriers and facilitators to programme impact Experience of the RCT
RCT – control participants	2	Individual	Expectations of the programme

			Experience of the RCT
Previous DEP pilot participants	6	Individual	Experience of the programme Perceived immediate & longer-term impact of the programme Barriers and facilitators to programme impact
SPOCs	1	Group	Training context of department Perceptions of programme in department Experience of RCT
CDDO	1	Group	Perceptions of the programme in CDDO Experience of RC Lessons for future evaluation
Delivery partner	1	Group	Experiences of delivering programme Feedback/ lessons from delivery Experience of RCT Lessons for future evaluation
Masterclass presenters	4	Email	Experience of delivering the programme Extent and nature of contact with participants
Total	24		

Sampling

RCT participants in the **intervention group** will be selected using a purposive sampling approach. Unlike sampling for a survey, which aims to generate a statistically representative sample, purposive sampling aims to capture the range of diversity of experiences and phenomenon (Ritchie et al, 2013).

Based on the data gathered for the baseline survey, the group displays substantial diversity in terms of characteristics, with 35 departments represented (with DWP, CO and HMRC having largest proportions) and 23 role types (with policy, ops/project delivery, HR and digital, data and technology the most common). Apolitical management information also indicates that at the time of writing that there was substantial diversity in terms of progress in completing the courses.

Table 5 sets out the proposed primary sampling criteria, which is comprised of two interlocking criteria, the first being the degree of completion of the modules and the second being grade. Table 6 then sets out the proposed secondary sampling criteria, which will be more loosely implemented, in that the criteria can be achieved by participants in any combination of the six cells created by the primary sampling criteria.

The sampling strategy is aimed at maximising the degree of diversity, but taking into account pragmatic considerations in terms of the total sample size and the timescale within which recruitment and data collection needs to take place.

Table 5: Primary sampling criteria

	Completed 50% - 100% of AI module	Completed 100% of AI module and started at least 1 other module
SCS1/2	Min 1	Min 1
G6	Min 1	Min 1
G7	Min 1	Min 1

Table 6: Secondary sampling criteria

Department	
DWP	Min 2
CO	Min 2
HMRC	Min 2
Role	
Policy	Min 1
Delivery	Min 1
HR	Min 1
Digital, data and technology	Min 1
Sex	
Male	Min 3
Female	Min 3

Similar criteria will be applied to the sampling of the **six previous DEP pilot participants** if relevant data is available, otherwise a convenient sampling approach will be taken.

All **SPOCs** will be invited to take part in the focus group and **all individuals who delivered the most recent masterclasses** will be invited to respond by email. The selection of the appropriate stakeholders from CDDO and delivery partners (Apolitical and Ernst & Young) will be discussed with the relevant teams in those organisations.

Data collection

Interviews will be conducted via Teams and with permission will be recorded and written up. All interviews will be undertaken by the evaluation research team using topic guides to help ensure that there is consistent coverage across interviews. Guides will be developed for each participant group informed by the research questions, but these will be used flexibly as an aide-mémoire rather than topics always being discussed in the same order or only once during the interviews. The guide will not be seen as an exhaustive list of topics and will not prevent unanticipated, but relevant, subjects being discussed.

Central to gaining an in-depth understanding of the issues is the need to probe the answers that participants give. This involves asking participants to amplify and expand answers, explain statements, clarify responses and pick up on inconsistencies (Ritchie et al, 2013, p194). For example, it will be important to explore with learners whether the material delivered was new and which elements were new, whether it helped them develop a new understanding or attitude and which parts of the course had this effect and why. Using the interviews in this way will enable the research team to 'get under the skin' of experiences and therefore gain insight into whether the mechanisms are operating and the factors that influence their effectiveness.

Data synthesis and analysis

The administrative data and follow-up survey will be analysed with descriptive statistics to understand the characteristics of the participants, fidelity to the intervention model and satisfaction with the modules. It will also provide insights into the mechanisms associated with the intervention, such as differences between the intervention and control group in terms of attitudes, as well as the experience of participants in the intervention group of the programme[5].

The qualitative data will be managed using the Framework approach (Ritchie et al, 2013). Within this approach, the data gathered from the interviews will be summarised into a framework developed in Microsoft Excel, subdivided into main themes and sub-themes where columns represent themes, and each row is an individual case. The Framework is developed pragmatically based on the research questions, the topic guide and knowledge of the data. The aim of this is to organise the data without too quickly forcing it into abstract conceptual categories. This means the data is arranged in a systematic way that is grounded in the accounts of the participants while closely tied to the research objectives and allows comparative analysis to take place both between and within cases.

The final stage of analysis involves working through the framework in detail, drawing out the range of experiences and views, identifying similarities and differences, developing and testing hypotheses, and interrogating the data to seek to explain emergent patterns and findings. The aim of the analysis is to develop analytical categories and explanations that are comprehensive in the sense of capturing the full range of views and experiences. Following the Framework tradition, a balance between induction and deduction will be used during the analytical process (Barnard, 2012). Early on, the focus will be inductive, in the sense that it will aim to understand participants from their point of view. As the process moves up the 'analytical ladder', existing concepts and the Theory of Change will be brought in, in order to deductively help to organise and contextualise the findings. For example, this could be comparing the participants' experiences of the programme with the way in which it was intended to be delivered and experienced.

The findings will reflect three broad types of analysis (Ritchie et al, 2013), as described below.

- **Thematic analysis** – this provides the foundation of the findings through categorising the different types of phenomena encountered (for example, different views on how new the course content was or different experiences of outcomes from the course).
- **The identification of typologies** – although typologies do not always exist, where they do exist they can be powerful tools for understanding the nature of

[5] As set out in the statistical analysis plan, cells with fewer than five individuals will be suppressed, with additional suppression (if necessary) to ensure that these values cannot be calculated based on the other reported values.

the phenomena by combining multiple elements identified through the thematic analysis at a case level (for example, types of participants categorised based on a combination of their approach to learning and perceived impact of the course).

- **Explanatory analysis** – this aims to understand the connections between different parts of the process and how they contribute to the outcomes and impact. This is developed through in-depth intra- and inter-case exploration (for example, the mechanisms, barriers and facilitators that explain different course outcomes).

Quantitative and qualitative data gathered as part of the IPE data will be combined and triangulated to test the intervention's logic model and interrogate the causal mechanisms underlying it. Integrating qualitative and quantitative data will be done by synthesising different kinds of data through a question-led approach. For example, the survey data will provide quantitative data on the proportion of participants who felt the RCT and randomisation was acceptable, while the qualitative data will be used to describe the range of experiences of the RCT and randomisation (thematic analysis) and the reasons why it was felt to be acceptable or unacceptable (explanatory analysis).

1.4 Quality assurance

Overall responsibility for the quality of all elements of the trial and all research materials rests with the evaluation's principal investigator, Dr Matt Barnard. Dr Barnard has over 20 years' experience designing and implementing mixed-method research and evaluation studies, including having designed and directed a range of field RCTs funded by organisations such as the UK Health Security Agency, the Education Endowment Foundation and the Youth Endowment Fund.

Dr Barnard will be supported in quality assurance (QA) by three other researchers.

Dr Sergio Salis, ICF's head of impact evaluation, has over 20 years' experience of undertaking complex quantitative analysis, including impact analysis using sophisticated statistical techniques.

Ali Zaidi is a consulting director with over 16 years' experience in conducting studies evaluating education and skills policies and programmes. Ali Zaidi has managed or directed studies for a range of UK and EU public sector clients, including the Department for Education, Education and Training Foundation. He has particular experience in leading complex evaluations using qualitative methods.

Dr Alice Diaz, a senior consultant at ICF, is a psychologist with substantial experience working on quantitative analysis and survey development, including leading large-scale UK-government commissioned random probability surveys.

An additional source of methodological support will be provided by the evaluation advisory board, which will provide advice and guidance at critical junctions of the trial.

1.1.4 The QA process

Dr Sallis will provide statistical quality assurance, focusing on verification and validity and applying the principles of RIGOUR (repeatable, independent, grounded in reality, objective, uncertainty-managed, and robust) (Gov.uk, nd). The statistical analysis plan was drafted by the trial statistician reviewed in detail by Dr Salis. Dr Barnard then also reviewed the draft and resolved points of discussion before it was finalised. At the analytical stage, the trial statistician will undertake the analysis, which will be quality assured by Dr Salis. This will involve reviewing the analysis against the statistical

analysis plan to ensure that it was carried out in accordance with the plan and any points of divergence are highlighted and the reason recorded (where the divergence is deemed necessary) or the analysis re-run in accordance with the plan. In addition, Dr Salis will do a line review of the statistical code (in Stata), to ensure that the analysis has been implemented correctly. If any issue arise with the code, Dr Salis will re-run the analysis to check on the results and undertake additional analysis if needed to verify that the results are accurate.

An initial layer of quality assurance for the implementation and process evaluation will be provided by Ali Zaidi, who will lead its design and analysis. Mr Zaidi will review the analytical outputs of team members, including the data management, record of analysis and reporting. Dr Diaz will QA the questionnaire used to measure the trial outcomes and gather data relevant to the IPE. Dr Barnard will then review the questionnaire and IPE analytical outputs, referring to the records of analysis, managed data and interview notes where useful.

Finally, Dr Barnard will review the whole of the final report against the protocol and against the CONSORT reporting guidelines for pilot trials (Eldridge et al., 2016).

1.5 Ethics, data protection and registration

1.1.5 Ethics

The evaluation was reviewed using ICF's research ethics checklist, which was developed based on the Government Social Research (GSR) Ethics Checklist. The review indicated that the evaluation was low risk and did not need to be formally reviewed by ICF's research ethics committee. A copy of the completed ethics checklist is included in the appendix. The decision to use experimental design, and delay treatment from one group, had been made by Government Skills prior to engaging ICF, and had been ethically assessed at that stage. The evaluation ensured that it was undertaken in a way consistent with ICF's and GSR's guidance, including ensuring:

- voluntary participation based on informed consent;
- avoidance of harm to participants;
- that data is kept confidential.

1.1.6 Data protection

The Cabinet Office will be the data controller for the evaluation and ICF will be the data processor. The lawful basis being relied on for data purposes is the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller (as defined by the General Data Protection Regulation Article 6). A data privacy notice will be made available to participants during the recruitment and consent processes, which will inform participants of their rights and provide further information on the study (detail of the information being collected, how ICF will store, process and protect personal data, and who the data subject should contact if they have any concerns). Data will be stored securely on ICF servers within a UK-based server and only be accessible to the study team. The evaluation team will destroy its copy of the data sets two years after completion of the final evaluation report.

1.1.7 Registration

The evaluation has been pre-registered on the Government Evaluation Registry.

1.6 Stakeholders and interests

Cabinet Office steering group

The following are the lead contacts within the Cabinet Office steering group who will be responsible for overseeing the evaluation:

- Siobhan Dickens, Head of Evidence and Impact, Government Skills: Senior Responsible Analyst;
- Emma Gibbs, Evidence and Impact Lead, Government Skills: Evaluation Project Lead;
- Tanzia Ahmed, Evidence and Impact Research Officer, Government Skills: Project Assistant;
- Neil Sherringham, Capability Lead, Government Skills: Project Advisor (DEP delivery);
- Philip Wilson, Head of Digital, Data, and Innovation, Government Skills: Project Advisor.

ICF

The following are the evaluation staff from ICF with primary responsibility for implementing the trial:

- Dr Matt Barnard, consulting director and director of the ICF's Centre for Behaviour Change: principal investigator for the trial with overall responsibility for its design and delivery;
- Ali Zaidi, consulting director consultant: lead for undertaking the scoping interviews, developing the Theory of Change and leading the implementation and process evaluation;
- Laura Campbell, senior consultant: project researcher, responsible for leading the development of the data collection portal and survey;
- Khin Lin, senior consultant: project researcher, supporting all aspects of the evaluation;
- Robert Wishart, ICF associate: trial statistician and lead for impact analysis.
- Dr Sergio Salis, head of impact evaluation: quantitative analysis quality assurance lead;
- Dr Alice Diaz, senior consultant: questionnaire quality assurance lead.

Members of the project advisory board are listed in Table 7 below.

Table 7: Advisory Board

Role	Name	Affiliation
Chair	Martin Petto	Strategy, Policy and Partnerships Deputy Director, Government Skills
Civil Service Members	Philipp Dreyer	Senior Evaluation Adviser, Evaluation Task Force
	Eff Blank	Evidence and Impact Lead, Government Skills
	Philip Wilson	Head of Digital, Data and Innovation , Government Skills
	Suzanne Moore	Deputy HR Director, HMRC
	Jess Arnold	Deputy HR Director, DSIT
	Dr Jack Blumenau	Associate Professor in the Department of Political Science, UCL & ESRC Policy Fellow in the Evaluation Task Force
	Barbra Webber	Head of Digital Learning and Frameworks, Government Digital and Data Capability team
Wider public sector stakeholders	Beth Thompson	Strategic Lead for Research and Evaluation, NHS England
External academic advisors	Dr Amanda Taylor-Beswick	Professor of Digital and Social Sciences, Director - Centre for Digital Transformation, University of Cumbria
	Professor Tom Crick	Professor of Digital Education and Policy, Education and Childhood Studies, Swansea University and Chief Scientific Advisor, DCMS
	Professor Christian Schuster	Professor in Public Management in the UCL Department of Political Science and Academic Co-Director of the UCL Policy Lab.
Secretariat	Elinor Cosgrave	Executive Support Office, Government Skills

Funding

The evaluation is funded by HM Treasury's Labour Market Evaluation and Pilots fund.

1.7 Timeline

Table 8: Timeline for the evaluation

Dates	Activity
01/08/24 – 30/09/24	Evaluation scoping and planning
01/10/24 – 01/11/24	Evaluation design, outcome survey design & testing, trial recruitment
04/11/24 – 14/11/24	Baseline outcomes survey
15/11/24	Randomisation and programme onboarding
18/11/24 – 31/01/25	Delivery of intervention
03/02/25 – 07/02/25	Follow-up outcomes survey
03/02/25 – 14/11/25	Control group undertakes programme
10/02/25 – 31/03/25	Analysis and reporting

2 Statistical analysis plan

This second section sets out the evaluation's statistical analysis plan, including describing the primary analysis, item response theory analysis (IRT) and the additional analysis that will be carried out alongside the primary analysis. The section also describes the approach to addressing any imbalance at baseline or missing data. It concludes by setting out how the effect size will be calculated.

2.1 Primary analysis

The primary analysis will estimate an intention-to-treat (ITT) effect size using a single-level OLS regression. Using a regression model is preferred to a t-test as it has greater statistical power (and therefore reduced uncertainty in effect estimates). This is because the baseline measure of the outcome, which is included as an independent variable (regressor), is likely to explain some of the variance in the outcome at endline. Additionally, including the baseline of the outcome measure should control for any existing differences between these two groups on the observed difference in outcomes (though it will not control for any imbalances on unobserved characteristics). This is because while randomisation should, *on average,* achieve balance on observed and unobserved characteristics, it is possible that there is an imbalance by chance. The regression equation is outlined below:

$$Endline_i = \beta_0 + \beta_1 Allocation_i + \beta_2 Baseline_i + e_i,$$

Where:

- $Endline_i$ is the total weighted score from the behaviour and attitudes outcome measure collected at endline for individual *I*;
- β_0 is the coefficient for the regression intercept;
- β_1 is the coefficient for the impact estimate that is, the average effect of the programme on those individuals who were randomly allocated to the intervention $(Allocation_i = 1)$;
- β_2 is the coefficient for $Baseline_i$, which is the total score from the behaviour and attitudes measure at baseline for individual *i*. The coefficient represents the association between pre- and post-intervention scores for respondents in the control group;
- e_i is the error term.

The analysis will be conducted in Stata 18.5, using the following syntax:

reg endline i.allocation baseline

As part of the sensitivity analysis, the effects will be re-estimated using an equation that includes a range of baseline variables (such as gender, department etc) as co-variates.

2.2 Interpreting the primary analysis

As noted above in the discussion of sample size calculations (section 1.3.1.3), it is unlikely that the pilot trial will detect a statistically significant difference between the outcomes of the intervention and control groups. Recognising this, the research team will follow the guidance set out by the Youth Endowment Fund that "effect size estimates should be interpreted cautiously to avoid over-interpretation and undue enthusiasm or pessimism" (YEF, n.d., p8). In practice, as the guidance notes, this

means describing any positive difference in average outcomes between intervention and control as 'evidence of promise' rather than evidence of efficacy or effectiveness. In terms of reporting quantitative data, following the extension to the CONSORT framework to randomised pilot trials (Eldridge et al., 2016), estimates of effects will be reported with 95% confidence intervals, but without P values.

2.3 IRT analysis

The IRT analysis will be conducted at one or two time points:

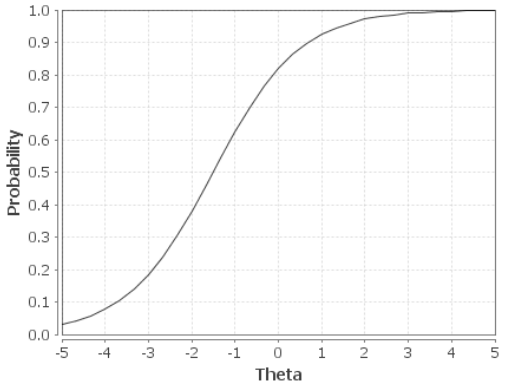
- **December 2024**: After randomisation to explore the properties of the outcome measure at baseline
- **February 2025** (only if the baseline outcome measure is revised): Using the endline data to explore the properties of the outcome measure at endline

The baseline IRT will be used to explore if the outcome measure can be revised or strengthened before the endline data collection begins. Although it is preferable not to revise the outcome measure between baseline and follow up, one aim of the pilot study is to explore the reliability of the outcome measure, and the baseline data provides a good opportunity to do this and for any learning to be applied at the follow up point. This will also contribute to the aim of the pilot to build capacity for future evaluations of the DEP or other learning programmes. If the outcome measure is revised, additional IRT will be conducted at endline to explore the properties of the outcome measure.

The IRT analysis will be conducted separately by topic: data culture, digital culture and AI confidence. Analysis will be conducted in Stata and jMetrik, to explore the issues set out below.

- **Internal consistency statistics:** This indicates how consistent responses to the test are. Guttman's λ_2 will be estimated where:
  - a value of less than 0.70 indicates poor reliability;
  - a value of greater than or equal to 0.70, but less than 0.80, indicates an acceptable level of reliability;
  - a value of greater than or equal to 0.80, but less than 0.90 indicates a good reliability level;
  - a value of greater than or equal to 0.90 indicates excellent reliability.
- **Item discrimination**: This indicates the extent to which an item can discriminate between Civil Servants with better or worse attitudes to data culture, digital culture and AI confidence. This is estimated as the probability that respondents answer correctly (or achieve a greater number of marks) based on their overall ability (measured by their total score across items). Item discrimination scores range from -1.0 to 1.0. We consider a value of greater than or equal to 0.30 to be acceptable, and greater than or equal to 0.50 to be excellent.
- **Item characteristic curves**: This provides information on the probability of a learner answering an individual item correctly based on their overall score (for the relevant sub-scale). The probability of answering an individual item should increase as a learner's overall scores increase. An example item characteristic curve is presented in Figure 1 below.

Figure 1 Example item characteristic curve



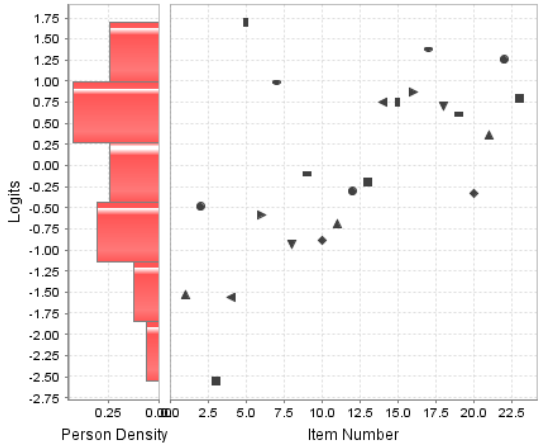
The x-axis displays the overall score of the individual, where zero represents the mean, and 1, 2, 3 etc. indicates the standard deviation away from the mean.

The y-axis indicates the probability of answering this item correctly, based on the individual's overall score.

In this example, the item discriminates well for most individuals.

- **Item-person map:** This indicates, graphically, whether there are enough items that are targeted at a range of outcome levels. This compares the participants 'ability', measured by their total score, to the item difficulty. Item difficulty is the probability of answering an item correctly as a function of the participants 'ability', measured by their total score.

Figure 2 Example item-person map



In this example, the left-hand graph indicates the density of the outcome distribution.

The right-hand graph indicates the probability of items being answered correctly based on the total score.

An instrument with a good range of targeting should ensure that there is a good distribution of items across the outcome distribution; as is the case in this example.

- **Distribution of the outcome:** The distribution of the total score is analysed graphically to indicate whether there are potential issues with floor or ceiling effects.

The results of the IRT analysis will be reported in a stand-alone technical report. This report will synthesise the results and consider whether changes to the survey questions are required. It may suggest the need for revising or removing items or adding further items to the questionnaire. Changes to the outcome measure will be reported in the final evaluation report. If changes are made, the same IRT analysis and technical reporting will be provided at endline, providing further recommendations for revisions (if necessary) for any future trial.

2.4 Additional analysis

In addition to the primary analysis, there will be three secondary analyses. These will explore the impact separately on the items related to the three modules. The analyses will follow the same approach as the primary analysis, substituting the baseline and endline variables accordingly. The process evaluation will also explore how compliance could be measured for a future trial and whether this could be used to undertake a dose-response analysis (i.e., to understand whether different levels of exposure to the intervention influence its impact) or complier average causal effect (CACE) analysis, which calculates the effect on the intervention on those who actually received it, rather than just being allocated to the intervention group.

2.5 Imbalance at baseline

Randomisation should, *on average*, achieve balance in observed and unobserved characteristics between the intervention and control group. However, it is possible that by chance an imbalanced sample is achieved, meaning that impact estimates would suffer from selection bias. While imbalance on unobserved characteristics cannot be assessed, it is possible to assess balance across baseline and time invariant characteristics.

The following data will be reported descriptively for the 'as randomised' and 'as analysed' samples:

- **outcome at baseline**: the behaviour and attitudes score at baseline;
- **demographics**: such as age, sex and gender identity, physical health conditions, ethnicity;
- **role and team**: primary area of work, civil service grade, whether they have previously held a role with a significant digital, data or technology element.

Statistically significant differences will be tested for as follows:

- **continuous variables**: differences will be assessed using t-tests;
- **categorical variables**: differences will be assessed with Chi-Square tests.

A p-value of less than 0.05 will be considered indicative of a statistically significant difference between the treatment and control group. If one or more variables is found to be statistically different, additional sensitivity analysis will be conducted and the effect size will be re-estimated using an inverse-probability weighting approach. In the first stage the propensity score will be estimated using the model below:

$$Allocation_i = \beta_0 + \beta_1 X_i + e_i,$$

Where X_i represents the vector of baseline and time-invariant characteristics. The inverse of the predicted values from this regression are used to weight observations based on their likelihood of being in the intervention group based on observed characteristics. This will be estimated using the *teffects ipw* command in Stata.

Given the small sample size, it is possible that there will be small cell counts, which could lead to statistical disclosure. Therefore, the following statistical disclosure controls will be applied: cells with fewer than five individuals will be suppressed, with additional suppression (if necessary) to ensure that these values cannot be calculated based on the other reported values.

2.6 Missing data

Missing data could occur for two reasons:

- item non-response;
- survey non-response.

Item non-response will be dealt with through a simple imputation process, using the mean of the items responded to within a sub-scale to impute the value of the item.

Survey non-response is harder to address. Multiple imputation is unlikely to be appropriate given the sample size and lack of auxiliary variables. Therefore, the following steps will be taken to identify any missing data patterns:

- conduct descriptive analysis, to explore the proportion of cases that are excluded from the complete case analysis for the treatment and control groups, respectively, against the baseline outcome, demographic, role and team variables outlined in the 'Imbalance at baseline' section;
- a 'drop-out' model will be estimated to explore patterns of missing data. This will use a logistic regression where the dependent variable is a binary indicator reflecting whether the observation was excluded from the primary analysis due to missing outcome or covariate data. The independent variables will be all the baseline outcome, demographic, role and team variables.

2.7 Effect size calculation

The effect size will be estimated as a *Hedge's g* effect size using the following formula:

$$g = J \times \left(\frac{\overline{Y}^{T}_{adj} - \overline{Y}^{C}_{adj}}{SD_{pooled}} \right),$$

Where:

- $\overline{Y}^{T}_{adj} - \overline{Y}^{C}_{adj}$ is the adjusted mean difference in outcomes, given by the coefficient β_1 in the primary analysis model.
- SD is the pooled standard deviation, given by:

$$SD_{pooled} = \sqrt{\frac{(n_T - 1)SD_T^2 + (n_C - 1)SD_C^2}{n_1 + n_2 - 2}}$$

- J is the correction factor, given by:

$$J = 1 - \left(\frac{3}{4(n_T + n_C - 2) - 1}\right)$$

Confidence intervals will be estimated as:

$$g - zv_g \leq g \leq g_{WT} + zv_g,$$

Where z is the critical value for statistical significance testing, multiplied by the variance (the square of the pooled standard deviation).

ANNEXES

Annex 1 ICF Research Ethics Checklist

Employees should familiarise themselves with the research ethics policy in its entirety. In addition, the below research ethics checklist must be completed during the proposal development stage. This checklist has been developed based on the Government Social Research Ethics Checklist. Upon completion, you will be signposted to specific points that should be considered as well as relevant sections of the research ethics policy, wider ICF policies and available templates and resources as appropriate to support you. These outputs should be referred back to if the project is won, and throughout the project lifespan.

Considering the ethical implications of our work is required for each and every research or evaluation project. However, where a project relates to the questions in red, Project Directors and Project Managers should take extra care and caution to ensure that we are meeting the highest ethical standards possible. This may involve for instance including research ethics as an agenda item at each project meeting with team members or the client, putting together a dedicated ethics steering group and submission to the ICF research ethics committee for approval.

1. Does the project involve collecting data from participants?
2. Is there a risk that certain groups will be excluded from the research?
3. **Are potential participants aged 16 or under?**
4. **Could the potential participants be considered vulnerable adults?**
5. **Might some of the research questions cover stressful or culturally sensitive subjects?**
6. Will incentives be offered to participants?
7. Will the project involve offsite data collection?

Project: Digital Excellence Programme Pilot Evaluation

Q1) Does the project involve collecting data from participants? If yes, consider the following…	
Proposed methodology: • Is the research design appropriate to the participants? • Is the level of respondent burden appropriate for the groups of people involved in the research? • How will the research consider the diverse perspectives of people according to their gender, disability, ethnicity, religion, sexual orientation, socio-economic status and age? Yes, the research design and level of burden is appropriate to the groups. Data on participant characteristics is being collected and will be reported descriptively.	
Informed consent: • What processes are in place to ensure that participants are informed and understand the project, the purpose, the client, topics and that their participation is voluntary?	■ **Research ethics policy:** 2.4 Informed Consent

• What can you do to ensure that participant agreement is made before the interview is conducted? The purpose of the evaluation and what was required of the participants will be included in the recruitment emails. Information and consent forms are included at the start of the survey and participants cannot proceed without indicating their consent. Consent to contact for follow up interviews is included in the survey and consent will be requested again of those people who are invited to take part in qualitative interviews.	■ Online surveys - a guide to good practice *(page 21-23)* ■ Practical guidance on integrating data protection into E&A services: Interviews *(page 7-8)* ■ Model consent forms and information sheets
Data protection: • What procedures are in place to ensure adherence to the Data Protection Act and other government data security requirements? • Reporting should not allow the identification of any individual. What checks are in place to ensure that no one can be identified? (for both quantitative and qualitative work) Data protection protocols will be in place, with data stored in secure locations only accessible to the research team. Quantitative data will be reported at an aggregate level, with no subsample analysis where groups are too small (i.e., below 5 people). Qualitative data will be reported in terms of themes and quotes or perspectives will be written in such a way as to avoid identification. This will be checked by the project director prior to the report being shared with the commissioner, the Cabinet Office. A DPIA will be drawn up and agreed with the Cabinet Office before data collection takes place.	■ ICF's data protection policy and procedures
Safety and wellbeing • What considerations have been taken to ensure participants' safety and wellbeing? The research team does not believe that there are any substantial safety or wellbeing concerns for participants or researchers in undertaking the study.	■ **Research ethics policy:** 2.3 Participant safety and wellbeing

Q2) Is there a risk that certain groups will be excluded from the research?	
• Might the research, sampling design or data collection method exclude some groups of people? • What steps can be taken to encourage and widen participation? (e.g. travel costs, childcare, varying times and locations of interviews, accessibility of venues, advance letters in different languages etc) • Do you need interviewer assistance such as offering help with the completion or a translator? • Do you need to consult with others so that barriers to participation for certain groups are reduced? • Have the interviewer/researchers demonstrated awareness of equality issues and an ability to work inclusively? • What is our role/responsibility to different stakeholders and research participants around dissemination? • Are there any accessibility or equality issues about how findings are made available or presented? The research team does not believe there are any major barriers to any potential participants taking part in the study. The research team is committed to making any reasonable adjustments for participants with disabilities that are required.	ICF Guidance: Inclusion of participants

Q3) Are potential participants aged 16 or under? If yes, consider the following.	
Informed consent: • Consent from a parent or legal guardian is required for children aged under 16 to participate in research, what processes are in place to ensure this is done? • How can you ensure that the children are also adequately informed about the work? N/A	**Research ethics policy:** 2.4.3 Gaining informed consent (vulnerable populations)
Chaperones: • It is sometimes recommended that an adult accompanies children and young people during an interview. What processes are in place to ensure this is in place when required? • Who is best to accompany the child(ren)? N/A	
Safety and wellbeing • What procedures are in place to ensure interviewers are properly trained and vetted (e.g. DBS check)?	**Research ethics policy:** 2.3

• What procedures are in place for disclosure of abuse? • What processes are in place if there is a concern for the safety and wellbeing of the participant or that of others? N/A	Participant safety and wellbeing ICF Europe & Asia Policies: Safeguarding Policy

Q4) Could the potential participants be considered to be vulnerable adults? If yes, consider the following. *A vulnerable adult is someone aged 18 or above who may need community care services for reasons like mental health issues, disability, age or illness. They may not be able to take care of themselves or protect themselves from harm or exploitation*[6].	
Informed consent • Are there any groups that might have difficulty giving informed consent themselves? • How can you ensure that participants are adequately informed about the work? • Is consent via gatekeepers required? If so, what processes need to be in place? • What steps can be taken to ensure representativeness, i.e., to ensure that participants are not "hand-picked" by gatekeepers or that there is a minority view promoted? N/A	**Research ethics policy:** 2.4.3 Gaining informed consent (vulnerable populations) Model consent forms and information sheets
Safety and wellbeing • What procedures are in place to ensure interviewers are properly trained and vetted (e.g. DBS check)? • Have the interviewer/researchers demonstrated awareness of equality issues and an ability to work inclusively? • What procedures are in place for disclosure of abuse? • What processes are in place if there is a concern for the safety and wellbeing of the participant or that of others? N/A	**Research ethics policy:** 2.3 Participant safety and wellbeing ICF Europe & Asia Policies: Safeguarding Policy

Q5) Might some of the research questions cover stressful or culturally sensitive subjects?	
Participant safety and wellbeing: • How will stress and sensitivities be minimised? • How can interview length be kept to the minimum?	■ **Research ethics policy:** 2.3 Participant

[6] https://www.mencap.org.uk/advice-and-support/safeguarding/safeguarding-adults

• Do you need to ensure that there is post-interview support? • What procedures are in place for disclosure of abuse? N/A	safety and wellbeing ■ ICF Europe & Asia Policies: Safeguarding Policy ■ **Research ethics policy:** 2.4.4 Gaining informed consent (sensitive topics) ■ Model consent forms and information sheets
Researcher wellbeing: • What procedures are in place to ensure the wellbeing of the researcher?	■ **Research ethics policy:** 3.1 Researcher safety and wellbeing

Q6) Will incentives be offered to participants?	
• Are incentives appropriate for the research topic and population of interest? • Could this be viewed as a form of coercion or negatively impact the power dynamic between the researcher/client and participant? • Are participants engaging in the research as part of their professional or personal time? Incentives are not being offered to participants, though some participants are being given 'free licences' to take part in the intervention. In this context 'free' means their home departments do not have to pay, but the cost is still borne by the Government.	■ Research ethics policy: 2.5 Incentives ■ Online surveys - a guide to good practice *(page 13)*

Q7) Will the project involve offsite data collection?	
• What procedures are in place to ensure the safety of the researcher? N/A	■ **Research ethics policy:** 3.1 Researcher safety and wellbeing

	■ Dynamic risk assessment for lone working ■ Lone working policy ■ Lone working procedures ■ Lone working risk assessment

Annex 2 Randomisation syntax

```
log using "$logs\01.1 Civil Service Skills Randomisation FINAL.log", replace

********************************************************************************
**# 01.1 CS Skills Randomisation - Final
*** Wishart Research Consulting Limited on behalf of ICF Consulting Services Ltd
*** CS Skills - 14/11/2024
********************************************************************************

/*
Change log:
        - 14/11/2024: Robert Wishart - Created do-file
*/

**********
**# 1. Load data
**********

import excel "$data_original\Baseline - List for randomisation.xlsx", clear firstrow
de // 4 vars, 157 obs
rename *, lower

**********
**# 2. Randomisation
**********

/*
Randomisation process:
        - Simple randomisation (no blocking)
        - The observations will be ordered by a random number
        - The random number will be generated using a "seed" to ensure replicability
        - The "seed" will be a random number, generated by random.org (between 1 and 1,000,000) - a screenshot of which will be saved in the syntax folder
        - The first half _n<(_N/2) will be randomised to the intervention, the rest to control.
*/

*****
* 2A. Generate random number for ordering
```

```
*****

set seed 449420 // seed 1 in folder
gen rand_order = runiform()

*****
* 2B. Sort by random number
*****

sort rand_order

*****
* 2C. Allocation
*****

lab def randomisation 0"Control" 1"Intervention"
lab val randomisation randomisation
replace randomisation = 1 if _n<(_N/2) // intervention
replace randomisation = 0 if randomisation==. // control
fre randomisation

**********
**# 3. Prepare data for export and save
**********

drop rand_order

compress
lab data "CS Skills Randomisation file - 20241114"

export excel "$data_intermediate\CS Skills Randomisation.xlsx", replace firstrow(varlabels)
save "$data_intermediate\CS Skills Randomisation.dta", replace

********************************************************************************
**# End of do-file
********************************************************************************

log close
```

Annex 3 References

Barnard, M. 2012. Critical Qualitative Theory: Opening up the Black Box of Criminal Justice Interventions, in Erica Bowen, Sarah Brown (ed.) *Perspectives on Evaluating Criminal Justice and Corrections* (Advances in Program Evaluation, Volume 13), Emerald Group Publishing Limited, pp.129-141 (DOI 10.1108/S1474-7863(2012)0000013010)

Belch, G.E., & Belch, M. A. 2012. *Advertising and promotion: An integrated marketing communications perspective* (9th ed.). New York: McGraw-Hill quoted in Hogg, M. A., & Vaughan, G. M. .2014. *Social Psychology* (7th ed.). London: Pearson

Central Digital and Data Office. 2023. *Digital and Data Essentials for Senior Civil Servants.* [online] GOV.UK. Available at: https://www.gov.uk/government/publications/digital-data-and-technology-essentials-for-senior-civil-servants/digital-data-and-technology-essentials-for-senior-civil-servants [Accessed 28 Oct. 2024].

Dalkin, S.M., Greenhalgh, J., Jones, D., Cunningham, B. and Lhussier, M. 2015. *What's in a mechanism? Development of a key concept in realist evaluation. Implementation Science*, 10(1). doi:https://doi.org/10.1186/s13012-015-0237-x.

Department for Education. 2011. *Teachers' standards guidance for school leaders, school staff and governing bodies*. [online] *GOV.UK*. Available at: https://assets.publishing.service.gov.uk/media/61b73d6c8fa8f50384489c9a/Teachers__Standards_Dec_2021.pdf.

Digital Economy Council, 2020. *Organising for Digital Delivery.* Available at: https://www.gov.uk/government/publications/organising-for-digital-delivery/organising-f+or-digital-delivery

Dong, N. & Maynard, R. 2013. *PowerUp! A tool for calculating minimum detectable effect sizes for experimental and quasi-experimental studies*. Journal of Research on Educational Effectiveness, 6:1, 24-67, DOI: 10.1080/19345747.2012.673143

Education Endowment Foundation. 2021. *Effective Professional Development*. [online] EEF. Available at: https://educationendowmentfoundation.org.uk/education-evidence/guidance-reports/effective-professional-development.

Eldridge, S.M., Chan, C.L., Campbell, M.J., Bond, C.M., Hopewell, S., Thabane, L. and Lancaster, G.A. 2016. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *BMJ*, 355, p.i5239. doi:https://doi.org/10.1136/bmj.i5239.

EU, [data unknown]. *Test Your Digital Skills*. [online] Available at: Digital Skills Assessment Tool (europa.eu) [Accessed 2nd October 2024]

Feldon, D.F., Litson, K., Jeong, S., Blaney, J.M., Kang, J., Miller, C., Griffin, K. and Roksa, J. 2019. Postdocs' lab engagement predicts trajectories of PhD students' skill development. *Proceedings of the National Academy of Sciences*, [online] 116(42), pp.20910–20916. doi:https://doi.org/10.1073/pnas.1912488116.

Grant, W. 2023. The Knowledge Deficit Model and Science Communication. *Oxford Research Encyclopedia of Communication.* Retrieved 8 Dec. 2024, from https://oxfordre.com/communication/view/10.1093/acrefore/9780190228613.001.0001/acrefore-9780190228613-e-1396.

GOV.UK. n.d. *The Aqua Book: Guidance on Producing Quality Analysis for Government*. [online] Available at:

https://www.gov.uk/government/publications/the-aqua-book-guidance-on-producing-quality-analysis-for-government. [Accessed 22nd November 2024]

GOV.UK. n.d. *What To Do After Someone Dies*. [online] Available at: https://www.gov.uk/after-a-death/organisations-you-need-to-contact-and-tell-us-once. [Accessed 2nd October 2024]

Knapke, J.M., Hildreth, L., Molano, J.R., Schuckman, S.M., Blackard, J.T., Johnstone, M., Kopras, E.J., Lamkin, M.K., Lee, R.C., Kues, J.R. and Mendell, A. 2024. *Andragogy in Practice: Applying a Theoretical Framework to Team Science Training in Biomedical Research*. British Journal of Biomedical Science, [online] 81. https://doi.org/10.3389/bjbs.2024.12651.

Kolb, D. A., Boyatzis, R. E., & Mainemelis, C. 2001. Experiential Learning Theory: *Previous Research and New Directions.* In R. J. Sternberg & L.-f. Zhang (Eds.), *Perspectives on Thinking, Learning, and Cognitive Styles* (pp. 227–247). Lawrence Erlbaum Associates Publishers.

Michie, S., van Stralen, M.M. and West, R. 2011. The Behaviour Change Wheel: a New Method for Characterising and Designing Behaviour Change Interventions. *Implementation Science*, 6(42). doi:https://doi.org/10.1186/1748-5908-6-42.

Moore G F, Audrey S, Barker M, Bond L, Bonell C, Hardeman W et al. 2015. *Process Evaluation of Complex* I*nterventions: Medical Research Council guidance*. BMJ; 350 :h1258 doi:10.1136/bmj.h1258

National Audit Office, 2021. *The Challenges in Implementing Digital Change.* Available at: https://www.nao.org.uk/insights/the-challenges-in-implementing-digital-change/

Petty, R.E. and Cacioppo, J.T. 1986. The Elaboration Likelihood Model of Persuasion. *Advances in Experimental Social Psychology*, 19, pp.123–205. doi:https://doi.org/10.1016/s0065-2601(08)60214-2. Quoted in Hogg, M. A., & Vaughan, G. M. .2014. Social Psychology (7th ed.). London: Pearson

Ritchie, J, Lewis, J, McNoughton-Nichols, C. 2013. *Qualitative Research Practice (2nd edition)*. London: Sage

Society of Education. n.d. *The Professional Standards | Society for Education and Training*. [online] Available at: https://set.et-foundation.co.uk/your-career/the-professional-standards/the-20-professional-standards-to-enhance-your-practice.

UK Government. 2021a. *National AI Strategy*. [online] GOV.UK. Available at: https://www.gov.uk/government/publications/national-ai-strategy.

UK Government. 2021b. *Declaration on Government Reform*. [online] Available at: https://www.gov.uk/government/publications/declaration-on-government-reform.

UK Government. 2024. *Generative AI Framework for HMG (HTML)*. [online] Available at: https://www.gov.uk/government/publications/generative-ai-framework-for-hmg/generative-ai-framework-for-hmg-html.

Vitello, M. and Salvi, C. 2023. Gestalt's Perspective on Insight: A Recap Based on Recent Behavioral and Neuroscientific Evidence. *Journal of Intelligence*, [online] 11(12), p.224. doi:https://doi.org/10.3390/jintelligence11120224.

Youth Endowment Fund. N.d. *Pilot Study Guidance*. Available at: https://youthendowmentfund.org.uk/wp-content/uploads/2022/03/3.-YEF-Evaluations-Guidance-Pilot-studies-March-2022.pdf [Accessed 17 Dec. 2024].