

Evaluating One Big Thing 2023

**Technical report on the evaluation and our findings of the
evaluation feasibility study**

Acknowledgements

Thank you to all the participants in this evaluation from across Government, who provided data through the online One Big Thing platform, and participants in Government People Group who completed pre- and post-assessments.

A group of Cabinet Office GSR colleagues completed this evaluation, supported by their line managers and senior leaders.

Evaluation data was collected in the platform provided pro bono by i.AI.

Executive summary

What was One Big Thing 2023?

- One Big Thing (OBT) is a new annual initiative in which all civil servants take shared action around a Civil Service reform priority. OBT 2023 focused on data upskilling and ran from September 2023 to December 2023. Forty-two per cent of all civil servants took part in OBT 2023, based on registrations on the online platform.
- The aims of OBT 2023 were:
 1. To create a practical moment of shared participation to reinforce that we are one Civil Service.
 2. To have a measurable uplift in data awareness, confidence, knowledge and understanding across the Civil Service.
 3. To have a long-term impact on participation in data and other training and initiatives.
 4. To contribute towards achieving better outcomes in the delivery of public services and policy through the use of data.
- OBT 2023 promoted data upskilling through 3 main activities:
 1. New online training materials available to the whole of the Civil Service.
 2. Allocating 7 hours of every civil servant's time between September and December 2023 to self-directed data upskilling activities, with supporting resources.
 3. Asking all line managers to host an activity, conversation or team meeting focused on the use of data in their team's day-to-day work.

Evaluating OBT 2023

- We carried out an evaluation to provide initial results on whether OBT 2023 met its aims. We surveyed all civil servants who took part in OBT 2023, and we also carried out a smaller case study evaluation in a single business unit.
- These evaluations assessed whether OBT 2023 had met its aims to create a shared moment of participation to reinforce that we are one Civil Service, and to have a measurable uplift in data awareness, confidence, knowledge and understanding across the Civil Service.
- We used pre/post tests to measure these aims. We administered a pre-survey and a post-survey to all civil servants who participated in OBT 2023. This survey asked questions about civil servants' shared Civil Service identity, and their data awareness and confidence. We also asked them their views on OBT 2023 and their intentions to act after the training. We administered a data literacy and data behaviours assessment within one business unit,

Government People Group. This complemented the cross-Civil Service evaluation by providing a more objective measure of changes in data literacy and behaviours.

- We also conducted a third study, which tested whether OBT 2023 could be linked to any changing trends in participation in data training. This was done using weekly attendance volumes of data training courses, extracted from the Civil Service Learning Platform.
- The overall evaluation was a feasibility study. We were using these evaluations to test out evaluation methods and learn lessons for the evaluations of future OBT events.
- Our evaluations had several limitations. One of the main limitations is that we did not have a comparison group, which means we cannot test whether OBT 2023 caused any changes we saw in participants' Civil Service identity, data awareness, confidence and knowledge. Another limitation is that the participants in our evaluation are not representative of the Civil Service as a whole. This is because we were not able to use a sampling strategy to achieve a sample from which we could generalise to the wider Civil Service. Participants in our evaluation self-selected into the survey and assessment, which means they may be different to other civil servants, for example in their level of motivation and engagement with OBT 2023, data or evaluation. We will explain these limitations further in the report.

What did we find?

Aim 1: did OBT 2023 create a practical moment of shared participation to reinforce that we are one Civil Service?

- Overall, our evaluation results suggest that OBT 2023 did generate participation across the Civil Service.
- Forty-two per cent of the Civil Service registered for OBT 2023 on the formal, online platform. Eighty-two per cent of those who registered completed the initial online training modules. This means around 1 in 3 (34%) of the whole Civil Service both signed up and completed some training. Around 567,000 data learning hours were recorded on the official platform.
- Twenty-three per cent of those who registered for OBT 2023 recorded completing the target 7 hours of data upskilling (about 10% of civil servants). This suggests that many people who registered for OBT 2023 may not have completed the programme.
- We do not have any data on civil servants who may have participated in local OBT 2023 activities, such as team discussions, but did not register on the online platform, or registered

but did not log the upskilling activities they completed, so the above figures may underestimate overall participation.

- We cannot say whether OBT 2023 was experienced as a “shared moment” based on our data. We asked civil servants about their identity as a civil servant and their sense of connection with other civil servants before and after taking part in OBT but the results were inconclusive and suggested OBT 2023 may not have had an influence on these issues.

Aim 2: did OBT 2023 lead to a measurable uplift in data awareness, confidence, knowledge and understanding across the Civil Service?

- Across the 2 studies, we found some very small positive improvements in participants’ data awareness, confidence and knowledge.
- We found very small increases in participants’ awareness of the relevance and use of data in their data-to-day roles during the period in which they participated in OBT. We also found very small increases in their confidence around data-related ideas (such as ethics) and activities (such as visualising data).
- In our case study, we found very small increases in civil servants’ ability to correctly answer some of the questions we set, which involved applying data to perform tasks (such as calculating something) and about key data-related concepts (such as averages).
- We also found small increases in reported use of data in writing and decision-making but did not find changes in all the data behaviours we asked about.
- Overall, this suggests that OBT 2023 may have resulted in some very small gains in participants’ data awareness, confidence and knowledge, including their ability to apply this knowledge to day-to-day work.
- It is important to emphasise that while the gains we found were statistically significant (for our sample), they were also very small. For example, participants’ average total number of correct answers in the data literacy assessment increased by less than one correct answer (from 6.3 to 6.9). Participants’ assessment of their data awareness and confidence on a 5-point Likert scale from strongly disagree (1) to strongly agree (5) increased on average by only 0.13 points.
- We also explored how relevant OBT 2023 participants found the training. Whether people found OBT 2023 to be relevant or not could help explain why they felt their knowledge, confidence and awareness had improved, or not, after taking part in the training. It could

also give an indicator of how likely it was that their use of data in day-to-day work would change after taking part in OBT.

- On average, we found that people only moderately agreed that OBT was a good use of their time and that the content was relevant to their role. Participants moderately agreed that they were likely to apply learning from OBT 2023 to their roles, but recorded lower scores for their intention to complete specific actions such as booking further training or creating a personal development plan.

Aim 3: Did OBT 2023 have a long-term impact on participation in data and other training and initiatives?

- Our findings for this aim were inconclusive. We did not detect a meaningful change in training courses attended when we analysed weekly volumes of data training courses hosted on the platform Civil Service Learning (CSL) during or immediately after OBT.
- We are unable to definitively say whether OBT was successful or unsuccessful in this aim. We could not capture courses or training undertaken outside of the CSL platform. This means we will not have included all the potential formal and informal training undertaken during the OBT intervention. It is possible that OBT led to an increase in these other training activities that we were unable to observe.
- It was also difficult to determine whether movements in the volume of data training courses attended during and after OBT were attributable to OBT due to the high weekly variance of the CSL data. This resulted in a wide forecast interval, meaning that we had greater uncertainty around the expected level of weekly volumes in the absence of OBT.

What are the lessons learned from our findings?

1. **A training initiative such as OBT may be able to achieve very small increases in participant knowledge, awareness and confidence across a large number of civil servants.**
- Our evaluations found very small improvements in data awareness, confidence and knowledge after taking part in OBT 2023. Even very small improvements may be valuable if they are achieved over a whole organisation. Previous evidence shows that small, but

widespread, changes may offer greater value than an intervention that achieves large effects with smaller groups of colleagues.¹

- Almost 220,000 civil servants took part in OBT 2023. As OBT matures as an annual initiative, and lessons are learned from implementation, it is likely that even higher participation rates could be achieved.
- There is much existing evaluation evidence that can be drawn on when planning future OBT events, focused on different reform priorities, to design an OBT with the best chance of achieving the highest possible impact with the largest possible group of people.

2. The design of future OBT events could do more to support people to apply new learning in their day-to-day roles.

- OBT 2023 took evidence-based steps to support people to apply learning to their day-to-day role by including line manager conversations and local, context-specific activities as part of the programme. This was a sensible place to start because these are relatively low-cost and simple to implement.
- Findings from the cross-Civil Service survey showed that there was still a gap between people's general intentions to use learning from OBT, and their intention to take specific, practical action to do so. Not everyone found OBT relevant. Some small changes in people's reported behaviours were found in our case study, but not across all behaviours.
- In planning future OBT events, further attention could be given to connecting the upskilling content to specific local work and goals, to help people apply new learning in their day-to-day roles.
- For example, this might include more scenario-based content in the training,² or providing evidence-based templates to support line managers help their teams embed the new skills within day-to-day work. These could include structured prompts and cues; action planning;

¹ Lacerenza, C.N., Reyes, D.L., Marlow, S.L., Joseph, D.L., and Salas, E. (2017) 'Leadership training design, delivery, and implementation: a meta-analysis', *The Journal of Applied Psychology*, 102(12), pp. 1686–1718. Available at: <https://doi.org/10.1037/apl0000241>

² Chernikova, O., Heitzmann, N., Stadler, M., Holzberger, D., Seidel, T., and Fischer, F. (2020) 'Simulation-based learning in higher education: a meta-analysis', *Review of Educational Research*, 90(4), pp. 499–541. Available at: <https://doi.org/10.3102/0034654320933544>

self (or team) monitoring; and opportunities to continue to repeat the new skills within work.³

3. It is possible to evaluate OBT again in future, to gain even more extensive and robust evidence to support future delivery of OBT events and other upskilling initiatives.

- OBT 2023 was the first of its kind in the Civil Service. Our evaluation has provided some useful lessons learned for how evaluations of future OBT events could be carried out.
- Overall, our evaluations show that it is feasible to evaluate OBT. The relatively light touch methods we used (pre/post surveys and assessments) could be adapted, improved and used again to understand whether future OBT events achieve their aims.
- Other evaluation methods might also be considered, so the evaluation can be well-tailored to the strategic questions about OBT and Civil Service upskilling we need to answer. Planning and resourcing evaluation from the outset ensures that the widest possible range of suitable evaluation methods are available.

³ Sims, S., Fletcher-Wood, H., O'Mara-Eves, A., Cottingham, S., Stansfield, C., Van Herwegen, J., and Anders, J. (2021) *What are the characteristics of teacher professional development that increase pupil achievement? A systematic review and meta-analysis*. London: Education Endowment Foundation. Available at: <https://educationendowmentfoundation.org.uk/education-evidence/evidence-reviews/teacher-professional-development-characteristics>

Contents

One Big Thing	11
Background to the evaluation of OBT 2023	13
The evaluations and their results	16
Important background to both evaluations	16
Was it ethical to carry out these evaluations?	17
Quality assurance	17
Statistical versus practical significance	18
Study 1 - Cross-Civil Service evaluation of shared Civil Service identity and data awareness, confidence and knowledge	19
Why did we carry out this study?	19
Research questions	19
Evaluation design and methods	20
Data collection	20
Results	23
Research question 1 (RQ1)	23
Research question 2 (RQ2)	24
Research question 3 (RQ3)	25
Research question 4 (RQ4)	27
Research question 5 (RQ5)	28
Discussion	29
Limitations	30
Study 2 - Case study evaluation of whether OBT met its aims within Government People Group, Cabinet Office	32
Why did we carry out this evaluation?	32
Research question	32
Evaluation design and methods	32
Data collection	34
Data literacy and behaviours assessment	34
Assessment roll out	35
Respondents	35
Results	37
Data behaviours	37
Study 2 discussion	41
Data literacy	41
Data behaviours	41
Limitations	42
Study 3 - Interrupted Time Series Analysis of weekly volumes of data training attendance on the Civil Service Learning platform	45
Evaluation design and methods	45
Data collection	46
Results	46

Evaluation discussion and recommendations	48
Lessons learned for future evaluations	50
APPENDICES	56
Appendix 1: Overview of the evaluation methods we considered	56
Appendix 2: Analytical approaches used to address research questions for Study 1	63
Appendix 3: Study 1 OBT survey	66
Appendix 4: Table of Study 1 results	75
Appendix 5: Descriptive analysis exploring the representativeness of the sample for Study 1	80
Appendix 6: Study 2 data literacy and behaviours assessment (One Big Thing knowledge check for GPG)	83
Appendix 7: Study 2 power calculations	91
Appendix 8: Study 2 factor analysis	92
Appendix 9: Study 2 detailed results	93
Appendix 10: Quality assurance for both studies	96
Appendix 11: Analysis overview for study 2	98
Appendix 12: Multiple hypotheses issue	100
Appendix 13: GSR Ethical Checklist for both studies	101

One Big Thing

One Big Thing (OBT) is a new annual initiative in which all civil servants take shared action around a Civil Service reform priority. OBT is sponsored by the Cabinet Secretary and is designed and implemented by the Modernisation and Reform Unit. There is an OBT senior sponsors' network to ensure OBT meets the needs of the whole Civil Service, is feasible to implement locally, and delivers against its aims.

OBT 2023 focused on data upskilling, and ran from September 2023 to December 2023. The better use of data is a priority for government to improve our understanding of complex problems and to target policies and activities to deal with them. OBT 2023 was designed to boost these efforts and help ensure we remain a modern Civil Service able to use data effectively across all our roles.

Forty-two per cent of all civil servants took part in One Big Thing 2023 and 567,000 learning hours were recorded.

OBT 2023 promoted data upskilling through three main activities:

1. New online training materials available to the whole of the civil service.
2. Allocating 7 hours of every civil servant's time between September and December 2023 to self-directed data upskilling activities, with supporting resources.
3. Asking all line managers to host an activity, conversation or team meeting focused on the use of data in teams' day-to-day work.

Online training materials

OBT gave all civil servants access to a new 90-minute data course delivered on the platform Civil Service Learning (CSL). The training was tailored to different levels of skill and experience, offering training aligned to 3 competency levels (awareness, working and practitioner level). Participants completed a pre-course assessment to direct them to the training most suited to their competency level.

Seven hours of self-directed data upskilling

Senior leaders and line managers encouraged every civil servant to spend 7 hours on self-directed data upskilling activities during the period in which OBT 2023 was live. To help civil servants access relevant materials, the online platform gave civil servants access to a catalogue of existing data training and resources, which had been checked for their quality and relevance by data and training

experts. Most departments and professions also made materials and activities available that were tailored to their specific context. Data and digital skills have been a learning and development priority for several years, and previous cross-government communication campaigns have reinforced this priority and signposted materials available for upskilling on the Government Campus. Civil servants could record any data training they had done in 2023 as part of their 7 hours of OBT data learning.

Line manager conversations

After teams had completed the online data training, line managers were encouraged to hold conversations and run activities within their teams to reinforce data upskilling and its relevance to day-to-day work.

Delivering OBT across the whole Civil Service

To ensure the whole Civil Service could participate in OBT in a way that worked for them, a senior sponsors' network was set up. This network met regularly and ensured OBT 2023 met the needs of the whole Civil Service, was feasible to implement locally, and was designed in the right way to deliver against its aims. Cross-Civil Service and local communications channels were used to get the message about OBT 2023 out to all civil servants and promote participation.

The aims of OBT 2023 were:

1. To create a practical moment of shared participation to reinforce that we are one Civil Service.
2. To have a measurable uplift in data awareness, confidence, knowledge and understanding across the Civil Service.
3. To have a long-term impact on participation in data and other training and initiatives.
4. To contribute towards achieving better outcomes in the delivery of public services and policy through the use of data.

In our evaluation of OBT we wanted to test whether the aims of OBT 2023 had been achieved. We also wanted to test out the feasibility of some evaluation methods, to learn lessons for future evaluations of OBT events and similar activities. The following sections set out the evaluation approach and methods, case studies, findings and recommendations.

Background to the evaluation of OBT 2023

Our modern Civil Service systematically evaluates its activities, using lessons learned to improve the quality of delivery and service to the public. Investing in evaluation, creating a system that incentivises testing openly and learning from our mistakes is crucial to fostering innovation. So it was important to use OBT 2023 to test out appropriate evaluation methods which would not only help us understand whether OBT 2023 had met its aims, but would also enable us to build an even higher quality evaluation for OBT 2024 and beyond.

The aims of the evaluation of OBT 2023 were:

1. Develop and feasibility test evaluation methods for assessing whether OBT had met its aims. This will inform the planning of a full evaluation of OBT 2024 and future OBT events.
2. Generate initial results on whether OBT 2023 had met its aims, to inform the planning of future OBT events.

To give us the best possible chance of testing out a range of appropriate evaluation methods, 3 Cabinet Office teams collaborated to deliver 3 evaluations of OBT:

1. **Cross-Civil Service evaluation:** an assessment of the extent to which civil servants who participated in OBT saw a change in measures linked to the aims of OBT during the period OBT was implemented (September to December 2023).

This evaluation used a pre- and post-survey. We asked questions about perceptions of shared Civil Service identity, data awareness, confidence and knowledge. The post-survey included additional questions about how likely participants were to apply new knowledge to their work and take further action based on the training. This evaluation was carried out by the Evaluation Task Force, supported by No. 10's i.AI incubator unit, which designed and built the evaluation platform.

This evaluation is now complete and is reported in this document.

2. **Case study evaluation:** an assessment of how far civil servants within one business unit, Government People Group (GPG), saw an uplift in their data literacy and data behaviours during the period OBT was implemented (September to December 2023). We used a pre-

and post-assessment where participating civil servants completed multiple choice questions that tested their data knowledge and skills, and asked them to report on concrete workplace behaviours related to use of data. This evaluation was carried out by the Government Skills and Curriculum Unit and the Civil Service Data and Insights Team.

This evaluation is complete and is reported in this document.

- 3. Evaluation of whether participation in data training increased as a result of OBT.** We gathered regular data on the number of people who participate in centrally offered data training on the CSL platform. We used this data to carry out an interrupted time series analysis (a statistical analysis of trends over time), to investigate whether OBT achieved its aim to increase participation in data training. This evaluation was carried out by the Government Skills and Curriculum Unit and the Civil Service Data and Insights Team (both based in GPG).

The 3 teams involved collaborated on the design of the evaluation, quality assured one another's work (as well as engaging additional peer review) and jointly developed the results and recommendations presented here.

The evaluations have several limitations, which means care needs to be taken when interpreting the results. These limitations are different for each evaluation, and are explained in each section of the report. An important limitation of both pre/post evaluations is that, due to the design of the intervention, there was no comparison group of civil servants who were not exposed to OBT. Therefore, we do not know whether any changes we saw in people's responses to the survey or assessment reflect an existing trend, were a result of OBT, or were caused by other factors. This means that these evaluations can only provide tentative early evidence of whether OBT showed promise in achieving its aims. They do not provide robust evidence of the causal impact of the programme. This limitation is less of an issue for our third evaluation, as we are able to create an artificial comparison group using our forecasted trend. However, the lack of an observable comparison group still prevents us from fully isolating the effect of OBT from any other coinciding events. It therefore remains a limitation for all 3 evaluations to differing degrees. A full list of the evaluation methods we considered is outlined in Appendix 1.

The main process evaluation question we focused on for OBT 2023 was whether it had met its participation aims. The Modernisation and Reform Unit within Cabinet Office, which runs OBT, also ran a lessons learned exercise with departments to take forward insights for OBT 2024 and beyond. Some survey questions (not analysed as part of our evaluation) also contributed user insights, which formed part of this lessons learned exercise. This was an internal exercise that was not part of the evaluation and is not reported here.

We did not implement an economic evaluation of OBT 2023. This is because measuring the impact of OBT in a reliable way is a necessary first step in being able to quantify or monetise that impact and compare against costs.

The evaluations and their results

We will now report our 3 completed OBT 2023 evaluation studies:

1. The cross-Civil Service evaluation.
2. The case study evaluation carried out within GPG.
3. The interrupted time series analysis carried out using Civil Service Learning training data.

Which OBT aims did we evaluate?

Our 3 evaluations considered whether OBT had met 3 of its 4 aims. These were:

1. To create a practical moment of shared participation to reinforce that we are one Civil Service.
2. To have a measurable uplift in data awareness, confidence, knowledge and understanding across the Civil Service.
3. To have a long-term impact on participation in data and other training and initiatives.

We were not able to evaluate one of OBT's aims. This aim was:

4. To contribute towards achieving better outcomes in the delivery of public services and policy through the use of data.

This is a long-term and complex outcome which could not be captured within the scope and timescale of our evaluation.

Important background to all evaluations

We will explain the methods of each evaluation, the evaluation results and the limitations, to support interpretation of the results. At the end of the report, we discuss the results of all evaluations, and outline our recommendations for OBT 2024. Before this, it is important to highlight some important background information for the evaluations.

Was it ethical to carry out these evaluations?

All studies followed the [code of ethics](#) of the Government Social Research Profession throughout design data collection, analysis, reporting and data storage (see Appendix 13 for Government Social Research ethical checklists for all studies).

Before completing the cross-government survey and (within GPG only) the data literacy assessment, a data privacy impact assessment was created for each project. At the start of each survey we included an explanation of the purpose of the data collection and how participants' data would be processed, stored and used so they could give informed consent (see Appendices 3 and 6). Though participation was encouraged, participants could withdraw at any time and anonymity was preserved in reporting. Participants were only asked to provide non-identifiable demographic data that was relevant to analysis (gender, grade and profession). We also collected participant email addresses to facilitate links between responses and learning records but data was fully anonymised before reporting.

We assessed that survey questions did not risk negative impact on participant wellbeing and did not require the disclosure of sensitive information.

Data will be securely stored for 3 years and then destroyed, consistent with government guidelines.

Quality assurance

We quality assured our analyses so that each team's work was checked for mistakes or omissions and that it had been completed to a high analytical standard. This ensures that what we report here is accurate and that we are confident in the results we present.

We carried out our data analysis in the statistical software environment 'R'. The 2 teams – the Evaluation Task Force and the Civil Service Data and Insights team – checked and cross-checked all code and outputs. A UK Research and Innovation (UKRI) policy fellow working in the Evaluation Task Force who had not been involved in either project performed a final review.

Quality assurance for the third study was undertaken by members of the Border Economic Team who had previous experience of forecasting.

Further information on the quality assurance process can be found in Appendix 10.

Statistical versus practical significance

Statistical significance has a specific meaning in statistics and evaluation. When we find a change to be statistically significant in an evaluation, we mean that we find it likely that the difference is non-random (not occurring by chance alone). Essentially, when we say a result is statistically significant it means an effect **exists**. Practical significance is different from this. This is how **meaningful** the effect we have found is.

We had a good chance of finding a statistically significant effect in our cross-Civil Service evaluation as we had a large sample size. In practical terms, though, we are most interested in whether this change is of a magnitude to be meaningful in practice. We also need to consider whether evidence we have gathered on a sub-group tells us anything about a wider population, even if we do find an effect within that sub-group. For example, we found a statistically significant effect within the GPG Business Unit, but this does not tell us anything about whether a similar effect exists within the wider Civil Service or not.

In our studies reported below, we will be trying to find out if there are statistically significant results, and we will also consider whether these results are practically significant.

Approaches to analysis

The approach to analysis for each study is described in further detail in the relevant sections below. One difference between the first 2 studies is that Study 1 (cross-civil service evaluation) examines mean responses to survey questions, while Study 2 (case study evaluation) examines medians and the distribution of responses. Means were appropriate to use as the basis of the cross-Civil Service evaluation due to the large sample size available. This meant there was less risk of mean responses being influenced by outliers than for the smaller sample size of the case study evaluation. This smaller sample size made the use of medians and distributional analysis more appropriate as the basis for analysis.

Study 1 - Cross-Civil Service evaluation of shared Civil Service identity and data awareness, confidence and knowledge

We now report Study 1. Additional information to support this section can be found in Appendices 3, 4 and 5.

Why did we carry out this study?

In this study we aimed to collect data on all civil servants who participated in OBT 2023. This allowed us to gather the broadest possible picture of whether OBT 2023 met its aims across the whole Civil Service, across a range of departments and grades.

This study was designed to provide evidence against 2 of the 4 aims of OBT 2023. The aims we measured were:

1. To create a practical 'moment' of shared participation to reinforce that we are one Civil Service.
2. To have a measurable uplift in data awareness, confidence, knowledge and understanding across the Civil Service.

Research questions

We developed 5 research questions (RQs) for this evaluation, based on the 2 aims we were measuring:

1. How many people took part in and completed OBT training between September and December 2023?
2. Did participants' sense of shared Civil Service identity change after completing the OBT 2023 training?
3. Did participants' data awareness, confidence or knowledge change after completing OBT 2023?
4. After participating in OBT 2023, did participants believe they could apply the learning to their day-to-day role?
5. After participating in OBT 2023, did participants intend to do anything differently at work as a result of the OBT training?

Evaluation design and methods

This study used a simple pre/post survey design. This compares 2 measures taken at 2 time points, and assesses whether there has been a change (up or down). It does not capture any causal relationship between the intervention (in this evaluation, OBT 2023) and any change found in the measures and cannot explain why any change occurred.

Data collection

Civil servants⁴ who participated in OBT were asked to complete a survey before they started the 90-minute online course. Once they had finished the 90-minute online course and recorded 7 hours of data training they were asked to take another survey. This used the same set of questions, plus some additional questions designed to give a richer picture of their experiences.

Survey questions were designed by the evaluation team to provide the information required to address the RQs. The survey questions are included in Appendix 3. The questions focused on participants' shared identity as civil servants, and their data awareness, knowledge and confidence, before and after participating in OBT. Responses were recorded against a 5-point Likert scale, from strongly disagree to strongly agree. Additional questions in the post-survey asked participants about their experiences of the training and how they expected to apply the content in their work. These additional questions were tailored to the training level they had completed.

It is important to note that while survey items (questions) were designed to assess the main aims of OBT 2023, which we included in our RQs, this survey did not go through a validation process. A validation process checks that a survey is measuring what it is intended to measure, in a reliable and consistent way, based on specialist statistical tests. It was not possible to use a validated survey because the aims of OBT were bespoke. Additionally, it was not possible for us to validate our survey within the implementation timeline for OBT 2023, and given the costs and time involved in validation, it may not have been proportionate to do so, as the aims of OBT 2024 onwards are likely to be different. Since the survey is not validated, this means that we cannot be completely confident that changes seen in the survey data are equivalent to changes in the participants' identity, data awareness, data confidence and data knowledge.

⁴ Approximately 0.1% of people who completed OBT 2023 may have been public servants but not civil servants. We did not exclude this data from our analysis as we did not have sufficiently reliable data to distinguish between these groups.

Surveys were built into the digital platform used to deliver OBT. This ensured that survey and participation data were collected and stored on the same platform. Both surveys were made available from 4 September 2023, the launch date of OBT 2023, and closed after OBT 2023 ended, in the first week of January 2024. The number of registrations and learning hours recorded were monitored throughout the duration of OBT.

In total, 218,583⁵ responses were received for the pre-survey and 32,559 for the post-survey. In other words, about 15% of OBT participants completed the post-survey, which is equivalent to about 6.5% of the Civil Service as a whole.

Participants had to respond to the pre-survey to access the learning platform. This was because it was being used to direct participants to the right level of training. A Data Protection and Impact Assessment explained how the data would be used; if participants did not agree they could opt out of OBT. The post-survey was not mandatory and could be taken at any time. It was signposted to participants using a web link after they had logged 7 hours of training on the platform. This means that there is some variation in the time between the pre- and post-surveys for different participants. We cannot be sure of any effect this may have had, as participants could log any data training carried out during 2023 as part of their 7 hours, including training completed between January and August, before the launch of OBT. This means we do not know whether participants who completed the post-survey later had done more training than those who completed it earlier, or not. We also do not know whether people were already on an upward trend in their data awareness before OBT 2023, or not, which makes it more difficult to know whether it matters that some people took the post-survey later than others.

Approach to analysis

Different analytical approaches were used to address each of the research questions. This is detailed in Appendix 2.

⁵ It was [reported](#) that OBT had been delivered to 212,000 people. This number was an initial estimate, before the final data draw down and analysis took place. The number reported above is the final estimate of participation based on the platform data and includes all participants, including the very small number who may have been public servants not civil servants.

Some survey items were not relevant to our RQs, so we excluded them from our analysis. These items were used by the OBT 2023 team as part of the lessons learned exercise.

We used tests of statistical significance to measure the likelihood that any differences between pre- and post-survey reflect actual differences over time in the population who responded to both surveys, rather than arising by chance. Wilcoxon rank-sum tests⁶ were used because of the ordinal nature of the data. A Bonferroni adjustment was applied to each set of items analysed within RQ2 and RQ3. This correction ensures that we test for significance across multiple hypotheses, rather than individual ones, to reduce the risk of a type 1 error (rejecting the null hypothesis when it is actually true). It is important to emphasise that the methods we used do not allow us to conclude whether any statistically significant results we found in the sub-group who responded to both surveys would apply to the wider Civil Service population, but it is likely that they would not, due to the risks of non-response bias.

We only analysed data provided by participants who completed both the pre-survey and post-survey. This was because we wanted the pre- and post-groups to be comparable. This would not have been the case if we had analysed a much larger group of respondents for the pre-survey than the post-survey. Those who only completed the pre-survey, and not the post-survey, may be quite different to those who did complete the post-survey. For example, they may not have completed the training, and they may have different levels of motivation around data as a topic.

We removed data from those who stated that they undertook more than 1,000 minutes of training (17 hours). This was because this figure suggested that they had either made an error in reporting their participation, or they were very unusual and not typical of most respondents. This produced a final population of 31,437 who were included in the analysis (6% of civil servants). We carried out sensitivity testing to check whether the removal of these participants influenced the overall results. We concluded that it did not, as the differences in mean outcomes between the group we included and those we excluded were never more than 0.01 for any outcome measured.⁷ Therefore, it was appropriate to remove this group's data.

⁶ Wilcoxon rank-sum test is a statistical test used to compare two independent samples. It is a non-parametric alternative to a t-test.

⁷ The interpretation of our numerical scale is explained at the beginning of the next section.

Results

For survey items which used a 5-point Likert scale, from strongly agree (5) to strongly disagree (1), mean scores of 3.1-5 can broadly be interpreted as agreement and mean scores of 1-2.9 can be interpreted as disagreement with the statement. Three is a neutral score (neither agree nor disagree), and responses close to 3 are also interpreted as neutral.

Research question 1 (RQ1)

How many people took part in and completed the training between September and December 2023?

In total, 218,583 people registered for OBT 2023 by completing the pre-survey. This equates to 42% of the Civil Service.⁸ Of those, 178,857 completed the online modules, which is 34% of civil servants, and 82% of those who registered.

In all, 50,955 people recorded completing 7 or more hours of learning (9.8% of civil servants, and 23.3% of those who registered for OBT). This suggests high attrition (almost 75%) between sign-up and completion. However, it is not possible to account for any upskilling activities undertaken but not recorded.

Table 1 shows the number of participants recording different hours of data upskilling activities. The target was for all OBT participants to complete 7 hours of data upskilling activities.

Table 1: Hours of OBT training recorded by participants

Hours recorded	Number recording hours, of post-survey sample	Number recording hours, total
Less than 1 hour ¹	1,635	96,832
1+ hours	4,263	35,429
2+ hours	1,967	13,678
3+ hours	1,153	8,141
4+ hours	990	6,013

⁸ Calculations are based on the most recent estimate of 519,780 civil servants in the UK as of March 2023.

5+ hours	667	4,200
6+ hours	536	3,335
7+ hours	20,226	50,955
Total	31,437	218,583

Source: One Big Thing 2023 pre and post-surveys

Notes:

1. The online training took 1.5 hours to complete. Therefore, in principle, participants should not have recorded less than 1 hour of training. Possible explanations for why some participants recorded less than 1 hour of training are: they mistakenly did not count the 1.5 hours of online training they had completed, they completed this more quickly than expected, or they completed the post-survey before they had recorded any learning hours.

Research question 2 (RQ2)

Did participants’ sense of shared Civil Service identity change after completing the OBT 2023 training?

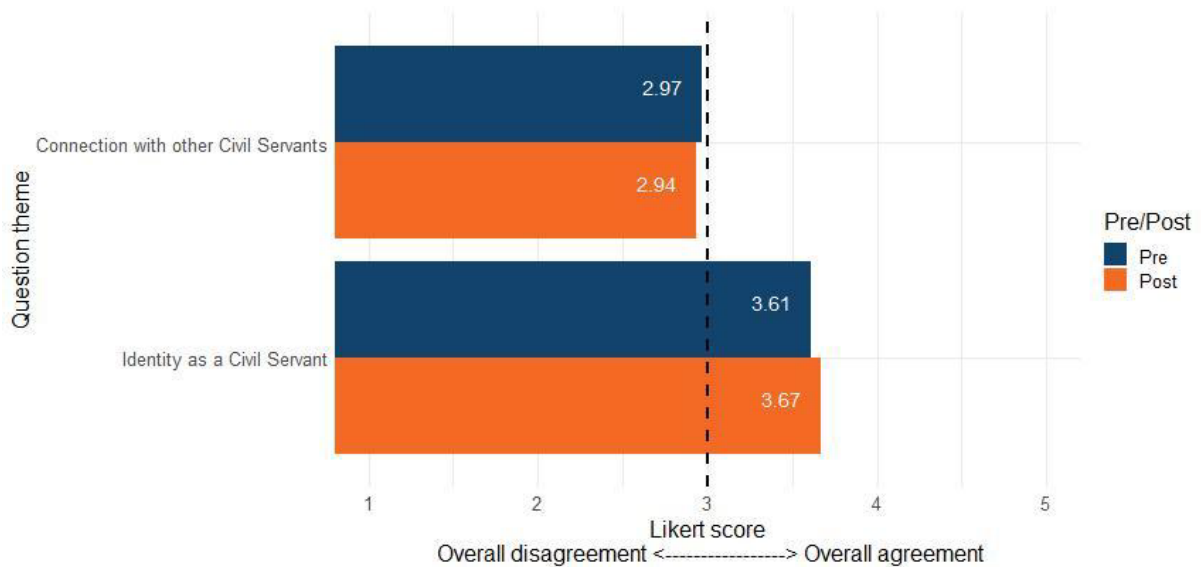
Our results are inconclusive, and on balance suggest that, overall, there was very little change in civil servants’ shared identity after participating in OBT 2023.

There was a very small numerical decrease in participants’ connection to the wider Civil Service after OBT (mean = 2.97 in the pre-survey and 2.94 in the post-survey). A greater number of civil servants agreed that their identity as a civil servant was important to them in the pre-survey, and there was a very small increase in this figure in the post-survey (mean = 3.61 in the pre-survey and 3.67 in the post-survey). As all scores were close to 3, overall participants’ views were relatively neutral on this issue.

These measures are not specific to the OBT content, nor are they closely related to data literacy more generally. There would be a range of factors influencing respondents' sense of connection to the wider Civil Service and their sense of Civil Service identity. Therefore, it is particularly difficult to assess whether OBT had any specific influence on these measures.

Results therefore suggest that, for the participants who completed our surveys, it is inconclusive as to whether OBT 2023 created a practical ‘moment’ of shared participation to reinforce that we are one Civil Service.

Figure 1: Bar chart showing results of pre- and post-tests for RQ2



Source: One Big Thing 2023 pre- and post-survey

Notes:

1. Sample size: 31,437
2. Figure 1 based on responses to survey questions (see Appendix 3):
 - a. Connection with other civil servants: 2a
 - b. Identity as a civil servant: 2b
2. After applying Bonferroni correction, all pre/post comparisons included in this figure were found to be statistically significant at the 95% level or higher. This is expected given the large size of the sample in this study.

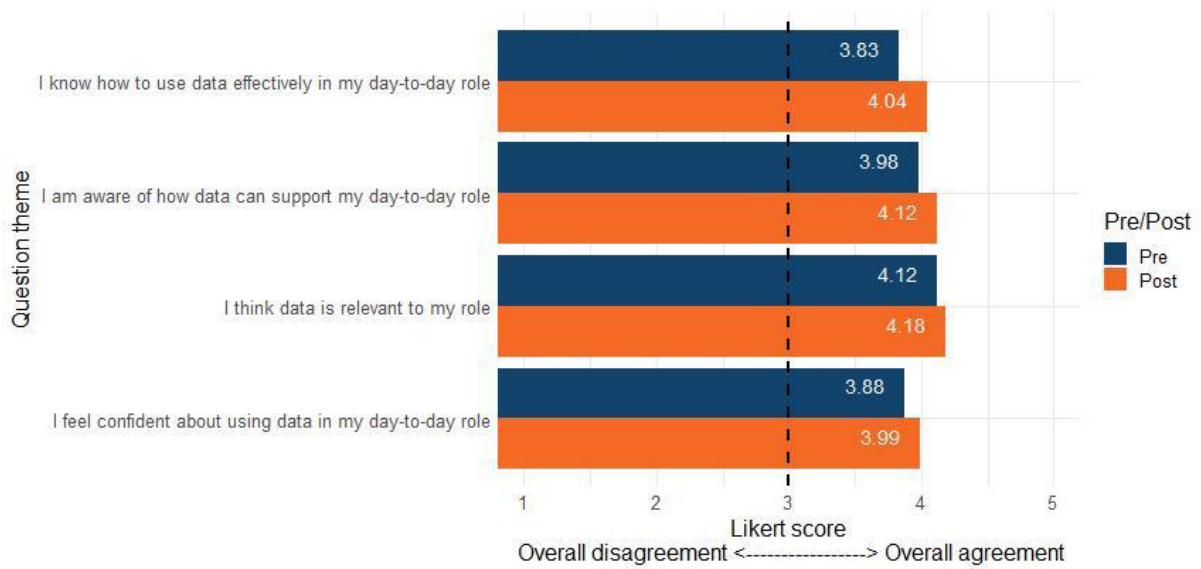
Research question 3 (RQ3)

Did participants’ data awareness, confidence or knowledge change after completing OBT 2023?

Four survey questions tested whether participants reported changes in their data awareness, confidence and knowledge. We asked one survey question which tested participants’ self-reported awareness of how data could support their day-to-day role. As shown in Figure 2, before starting OBT, participants perceived themselves to have a positive awareness of how data could support their day-to-day role (mean = 3.98), and there was a very small increase in this measure after completion of OBT (mean = 4.12).

There were also very small increases in participants’ agreement that data was relevant to their role (from 4.12 to 4.18) and that they were aware of how data could support their day-to-day role (from 3.98 to 4.12). There was a slightly larger increase in participants’ agreement that they knew how to use data effectively day to day (from 3.83 to 4.04). These results should be read in the context that the participants’ baseline scores were already relatively high.

Figure 2: Bar chart showing results of pre- and post-tests for RQ3



Source: One Big Thing 2023 pre- and post-surveys

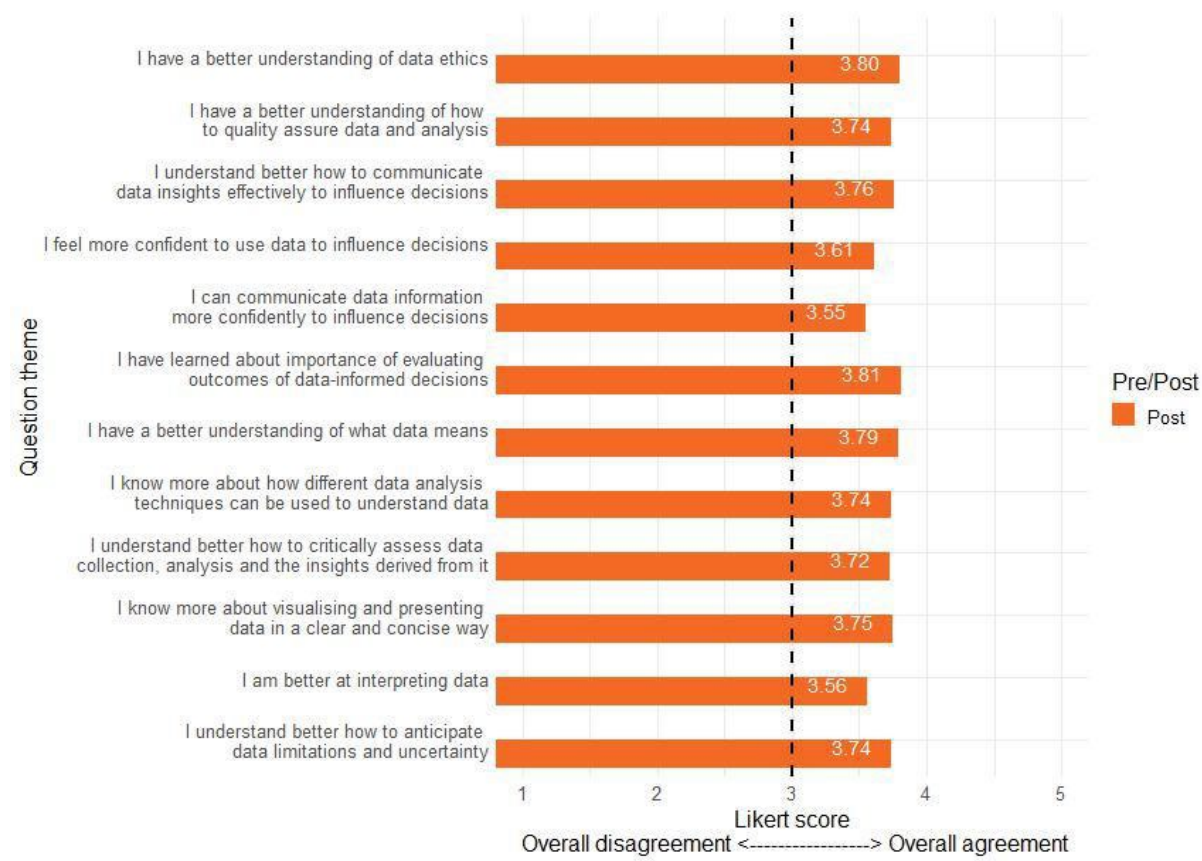
Notes:

1. Sample size: 31,437
2. After applying Bonferroni correction, all pre/post comparisons included in this figure were found to be statistically significant at the 95% level or higher. This is expected given the large size of the sample in this study.

There were 12 other survey questions which provided insight into participants’ self-reported data awareness, confidence and knowledge, as shown in Figure 3. These questions only appeared in the post-survey, so do not tell us anything about changes throughout OBT 2023. The questions asked participants to rate to what extent they agreed that different areas of their data awareness, confidence and knowledge had improved as a result of OBT.

Average participant responses ranged from 3.55 for communicating data more confidently to 3.81 for learning about the importance of evaluating outcomes of data-informed decisions. Overall, participants therefore moderately agreed that their data awareness, confidence and knowledge had improved as a result of OBT.

Figure 3: Bar chart showing results of post-test insights for RQ3



Source: One Big Thing 2023 post-survey

Notes:

1. Sample size: 31,437 (Awareness: 10,525; Working: 15,190; Practitioner: 5,722)

Research question 4 (RQ4)

After participating in OBT 2023, do participants believe they can apply the learning to their day-to-day role?

Three post-survey questions tested participants' beliefs about applying the learning to their day-to-day role following completion of OBT training. As shown in Figure 4, mean scores were consistent across each question, ranging between 3.44 and 3.46. This shows moderate agreement that participants believed they could apply the training to their day-to-day roles.

Figure 4: Bar chart showing post-tests for RQ4



Source: One Big Thing 2023 post-survey

Notes:

1. Sample size: 31,437 (Awareness: 10,525)
2. Figure 4 based on responses to survey questions (see Appendix 1):
 - a. More interested in working with data day-to-day – Awareness a-iii
 - b. Improved understanding of how to use data day-to-day – 7a-ii
 - c. Content was relevant to their role – 7a-iv

Research question 5 (RQ5)

After participating in OBT 2023, do participants intend to do anything differently at work as a result of the OBT training?

Six post-survey questions tested participants' intentions to behave differently following completion of OBT training. As shown in Figure 5, there were a range of scores for this domain. This ranged from 2.83 for intention to find or become a mentor, representing moderate disagreement, to 3.55 for intention to apply learning to their role, representing moderate agreement. Overall, there is a mixed picture here, with slightly higher scores for reported general intention to take action, and slightly lower scores for reported intentions to complete specific actions; for example, a score of 3.40 for intention to participate in further training versus a score of 3.13 for intention to book training.

Figure 5: Bar chart showing post-tests for RQ5



Source: One Big Thing 2023 post-survey

Notes:

1. Sample size: 31,437
2. Figure 5 based on responses to survey questions (see Appendix 3):
 - a. Intend to participate in further training – 7a-iii
 - b. Intend to apply the learning in their role – 7a-v
 - c. Intend to find or become a mentor – Awareness b-iv / Working d-iv / Practitioner f-iv
 - d. Intend to create a development plan – Awareness b-i / Working d-i / Practitioner f-i
 - e. Intend to book training – Awareness b-iii / Working d-iii / Practitioner f-iii
 - f. Intend to add learning to their development plan – Awareness b-ii / Working d-ii / Practitioner f-ii

We also explored whether OBT 2023 participants found the training relevant or not. This could help explain why they felt their knowledge, confidence and awareness had improved, or not, after taking part in the training. It could also give an indicator of how likely it was that their use of data in day-to-day work would change after taking part in OBT, and whether they went on to take further data training after OBT. On average, we found that people only moderately agreed that OBT was a good use of their time and that the content was relevant to their role.⁹

Discussion

Across our five main research questions, we found the following results:

⁹ Survey participants were asked to what extent they believed that OBT was a good use of their time, using a 5-point scale ranging from 1 (strongly disagree) to 5 (strongly agree). The average score was 3.19. Participants were also asked to what extent the content of OBT was relevant to their role, and the average score was 3.44.

1. **RQ1:** in total, 218,583 civil servants signed up for OBT by completing the pre-survey. This accounts for 42% of the Civil Service. Of these, 178,857 staff completed the online modules (34% of civil servants) and 50,955 recorded more than 7 hours of learning (9.8% of civil servants). We received 32,559 responses for the post-survey, of which 20,226 (64.3%) recorded more than 7 hours of learning. These results do not capture any civil servants who may have completed OBT-related activities in their departments, but did not register on the online platform.
2. **RQ2:** results are inconclusive, and suggest overall that there was very little change in civil servants' sense of shared identity after participating in OBT 2023.
3. **RQ3:** we found very small increases in survey participants' self-reported data awareness, confidence and knowledge after participating in OBT.
4. **RQ4:** we found moderate agreement that participants believed they could apply the content of OBT 2023 to their day-to-day roles. Again, this reflects participants' perceptions, and is not a measure of whether they did apply the learning to their work.
5. **RQ5:** results ranged from moderate disagreement through to moderate agreement that participants intended to behave differently as a result of OBT training. This range in scores suggests there may be a gap between participants' general ambitions after OBT and their commitment to take practical steps to implement them. Again, this is a measure of perception and intention, not a measure of whether participants did take any action.

Limitations

Our evaluation did not include any non-OBT comparison group. The methodology, based on pre/post responses from a non-representative sample of civil servants, was not designed to allow us to isolate the impact of OBT relative to other factors or pre-existing trends. These results instead tell us whether average responses changed over time in this group, but without being able to disentangle whether any observed changes were due to OBT or something else.

The sample is not a probability sample, and therefore cannot be generalised to all civil servants.

Our sample included only those participants who completed the pre-survey and the post-survey.

Post-survey responses were substantially lower than the pre-survey, and are likely to represent those who were most engaged with OBT 2023, and thus more motivated to complete the survey. Together, this means the overall sample is likely to overestimate the changes resulting from OBT. This is important to note, given the magnitudes of changes we found were very small, and the levels of agreement with the statements generally fell below 4 (agree) in most cases.

Survey responses captured self-reported change rather than objective changes in data awareness, confidence or knowledge. These types of self-reported measures can often be subject to an optimism bias (participants overestimate their knowledge or skill). As the surveys only covered the period where One Big Thing was live, longer-term outcomes cannot be assessed. The insight that can be gained from measuring short-term self-reported outcomes is limited.

We do not have pre- and post-measures available for all survey domains, meaning evidence is weaker against some of the research questions (RQ4 and RQ5).

The survey was not validated. This means there is a chance that some items do not in fact measure the constructs (OBT aims) they were intended to. This is a risk with any survey instrument that has not been through a validation process to confirm that the survey items measure what is intended.

Analysis has focused on mean changes, not distributional effects. The mean does not tell us about the spread of the data, for example whether there were a lot of participants with neutral responses, or very extreme responses in both directions.

Study 2 - Case study evaluation of whether OBT met its aims within Government People Group, Cabinet Office

Why did we carry out this evaluation?

A limitation of the cross-Civil Service evaluation was that the assessment only measured attitudes and confidence, rather than taking an objective measure of data knowledge, skills (in this report we refer to these 2 ideas together as data literacy) and behaviours. We ran a smaller scale, second evaluation within GPG to pilot a more objective measure of data literacy and behaviours. Data literacy links to OBT's second aim: a measurable uplift in data awareness, knowledge and understanding.

The objective was not to evaluate whether OBT met this aim across the Civil Service, but to run a smaller scale evaluation to see how feasible it was to evaluate gains in data literacy and behaviours (rather than only confidence and attitudes) in a valid way, to inform future evaluations.

Research question

Our research question was:

Is there a measurable uplift in data literacy and behaviours for this business unit sample during the period coinciding with OBT?

Evaluation design and methods

The population for this case study evaluation was GPG. This group was chosen as a case study population for convenience as the GPG Executive Committee wanted to carry out an evaluation of OBT that went beyond the cross-Civil Service evaluation. GPG consists of 966 members of staff mostly in HR, policy, data and digital professions. Everyone was issued with a request to complete either the pre- or post-assessment, regardless of their intention to participate in OBT (for the pre-test) or their participation in OBT (for the post-test).

We used a simple pre/post assessment design, using a data literacy assessment administered to GPG at the start of OBT 2023 (September 2023) and the end of OBT 2023 (December 2023). By using the

same assessment at the beginning and end, we could measure any average changes in scores. We decided to include the whole of GPG as the evaluation population based on power calculations, which showed us we were likely to need around 200 participants to be able to detect a statistically significant effect in our data. This assumed that OBT may have a moderate effect on data literacy.

To improve the rigour of this design, we randomised GPG into 2 groups, where one group was invited to complete the pre-test in September and one group was invited to complete the post-test in December. The reason for this was to remove the risk that people would score more highly on the post-assessment only because they had already completed the assessment once before, and not because their data literacy had genuinely improved. It also offered the best possible chance of getting a high response rate, as it is common for people to opt into the first test in a design like this in higher numbers than the second test (as we saw with the cross-Civil Service design). It is important to note this is not a randomised control trial design. The randomisation in the timing of test administration was used to remove a risk of test-retest bias, but there was no manipulation of OBT roll out and no comparison group.

This evaluation measured whether there was a statistically significant uplift in data literacy and behaviours within the group of GPG colleagues who responded to the request to complete an assessment. We cannot be certain that the measure we take of any changes within this group of people would be the same as a measure for the whole of GPG because of the large risk that people who chose to complete the assessment are systematically different to those who did not (for example, in terms of their motivation, or engagement with OBT), influencing their scores.

The evaluation does not measure whether any changes we detect are attributable to OBT. Colleagues who responded to the request to complete an assessment were exposed to all aspects of OBT 2023, including internal communications about OBT, being able to register for the online course content and potentially participating in individual or collective OBT-related data training. They may have also been exposed to other data content and activities unrelated to OBT 2023, and may already work with data in their roles. Their data literacy and behaviours may already have been improving before the launch of OBT 2023. Our evaluation measures any change in data literacy, which could be due to any of these influences, as well as others we may not be aware of. Our evaluation does not tell us whether OBT itself caused any changes we see in the data. Additionally, GPG delayed its OBT start date to 4 October, one month after the start for the wider Civil Service. There is therefore a

possibility that members of GPG may have been exposed to some aspects of OBT prior to taking part in the pre-assessment.

Data collection

Data literacy and behaviours assessment

We designed a bespoke data literacy and behaviours assessment for this evaluation as we were not able to identify a suitable existing data literacy assessment (having carried out a scoping exercise to determine if any were aligned with the aims of OBT 2023 and our evaluation requirements). We used the OBT training programme as a guide to what we should assess and research on what constitutes data literacy.^{10,11}

We also used factor analysis to check construct validity (that the questions were measuring the aspects of data literacy we expected them to measure) and to increase our confidence that the 11 individual questions we included measured one common construct – data literacy. This approach meant that even though we were not able to fully validate our assessment within the OBT 2023 implementation timeline, we had the best assessment possible within those constraints.

The final assessment included 16 questions. Five behavioural questions focused on data use in day-to-day work. Eleven data literacy questions focused on foundational data ideas, such as which method of averaging is least affected by outliers, and simple mathematical questions which required data manipulation (see Appendix 6). Performance on the 11 data literacy questions were combined into one measure of data literacy.¹²

We collected information on grade (seniority), profession and gender to enable us to check whether the pre- and post-samples were balanced, check for risks of non-response bias, and assess how similar or different our GPG sample was to the rest of the Civil Service. Participants provided their email addresses to allow us to cross-reference their assessment response with their participation

¹⁰ Sternkopf, H. and Mueller, R.M. (2018) *Doing good with data: development of a maturity model for data literacy in non-governmental organizations. Proceedings of the 51st Hawaii International Conference on System Sciences*, Hilton Waikoloa Village, Hawaii, January 3 - 6, 2018.

¹¹ Bratianu, C., Hadad, S., and Bejinaru, R. (2020) 'Paradigm shift in business education: a competence-based approach', *Sustainability*, 12(4), p. 1348.

¹² Factor analysis supported the hypothesis that the 11 data literacy questions measured one underlying construct.

data. We also asked respondents to tell us whether they had participated in the core OBT digital learning. This information helped us to work out whether the group who responded to our assessments were substantially different from the wider GPG directorate. This will help us consider how generalisable our results might be to GPG more widely.

To maximise the statistical power our evaluation method could achieve (the likelihood we could detect a statistically significant effect), it was vital to encourage uptake. This involved repeat emails from both the Civil Service Data and Insights team and the senior leadership of GPG. During the post-assessment collection period, we sent a business unit-wide email to try to maximise uptake, as well as an article in the GPG newsletter. These messages were also reinforced in team communications (for example, team meetings), and cascaded from the senior leadership of GPG, through line managers.

Assessment roll out

We randomised GPG into 2 groups and issued the pre-assessment to one group and the post-assessment to another group. The reasons for this approach were to reduce the risk of test-retest bias, as explained above. The assessment was issued using the Qualtrics survey platform. The pre-assessment was issued to Group One before OBT 2023 was formally launched in GPG, and was open for 23 days. The post-assessment was distributed to Group 2 mid-way through OBT 2023, and was open for 45 days.

The pre-assessment and post-assessment were distributed to 914 members of GPG staff in total, with approximately half of this group of 914 people receiving each assessment. Of the invitations sent, 42 emails were undelivered and so the final population was 872. A total of 392 participants received the invitation to complete the pre-assessment and 479 received the invitation to complete the post-assessment.¹³

¹³ An email distribution list was not identified for the directorate, so emails were sourced through a list of names. Some names did not align with the names used in emails, meaning that not all staff could be contacted. There were also staff listed who had recently left the department. This also accounts for the differences in final sample sizes between the two groups.

Respondents

In total, 288 people completed the assessments. 190 completed the pre-assessment, and 96 completed the post-assessment. The overall completion rate was 33%, but with a much higher completion rate for the pre-assessment (48.1%) than the post-assessment (20.3%).

Approach to analysis

For the behaviour questions we measured the changes between the pre and post-assessment by comparing median scores for the questions about the use of data for analysis and in discussion, and the percentage of respondents answering 'yes' for reported use of data in decisions and in writing.

For the literacy questions we measured the difference between pre- and post-assessments for 3 measures:

1. The overall percentage of correct answers.
2. The raw scores for the set of literacy-based questions as a whole.
3. Differences in score for individual questions.

For each of these measures, we checked whether the results were statistically significant, meaning we checked whether there was a strong probability that they measured real changes in the participants who responded, rather than occurring by chance. For the third set of calculations (differences in score for individual questions) there is a risk we found false positives (so, we found a statistically significant effect where one does not exist) because we ran multiple statistical tests concurrently. For this reason, when discussing the results, we focus mainly on overall scores rather than individual questions. Further discussion of this issue can be found in Appendix 12.

T-test and Chi-squared tests were used to check for differences in the means and distributions for the data literacy questions. We used Wilcoxon ranked-sum and Chi-square tests for the 5 behaviour questions. Further details on our analysis approach can be found in Appendix 11.

Results

Data behaviours

We found a statistically significant increase of 13.6 and 14.1 percentage points for “Yes” responses in participants’ reported use of data in making or suggesting decisions, policy or strategy and their reported routine inclusion of data, facts or numbers in writing/policy recommendations (assessment questions 4 and 5).

Questions 1 to 3 (see Table 2 below) measured the use of data in day-to-day work. We found statistically significant changes in the distribution of responses for questions 1 and 2. This may appear to be a puzzling result, as question 1 saw no change in median response, and question 2 saw only a modest increase, from 2 to 3. This is because the statistical test we used measured the change in the distribution of responses (so, the number of people giving each response), not only the change in the median score. We found no statistically significant difference for question 3. These results are illustrated in Table 2.

While we found a difference between the pre- and post-distributions for questions 1 and 2, which is likely to be non-random, the relatively small change in median responses make the practical significance of these results uncertain.

Table 2: Data behaviours pre/post comparison results

Question	Median response pre	Median response post	Statistically significant difference?
1) How many data reports, dashboards or visualisations have you interacted with in the last week?	3	3	Yes
2) How many times this week have you discussed data trends, metrics or insights with colleagues?	2	3	Yes
3) How many hours this week did you spend analysing, manipulating or interpreting data as part of your regular job functions?	2	2	No

Question	% yes pre	% yes post	Statistically significant difference?
4) Have you made or suggested a decision, policy or strategy based on data analysis in the past week?	45.8	59.4	Yes
5) Do you routinely include data, facts or numbers in your writing/policy recommendations?	68.2	82.3	Yes

Source: One Big Thing 2023 GPG assessment

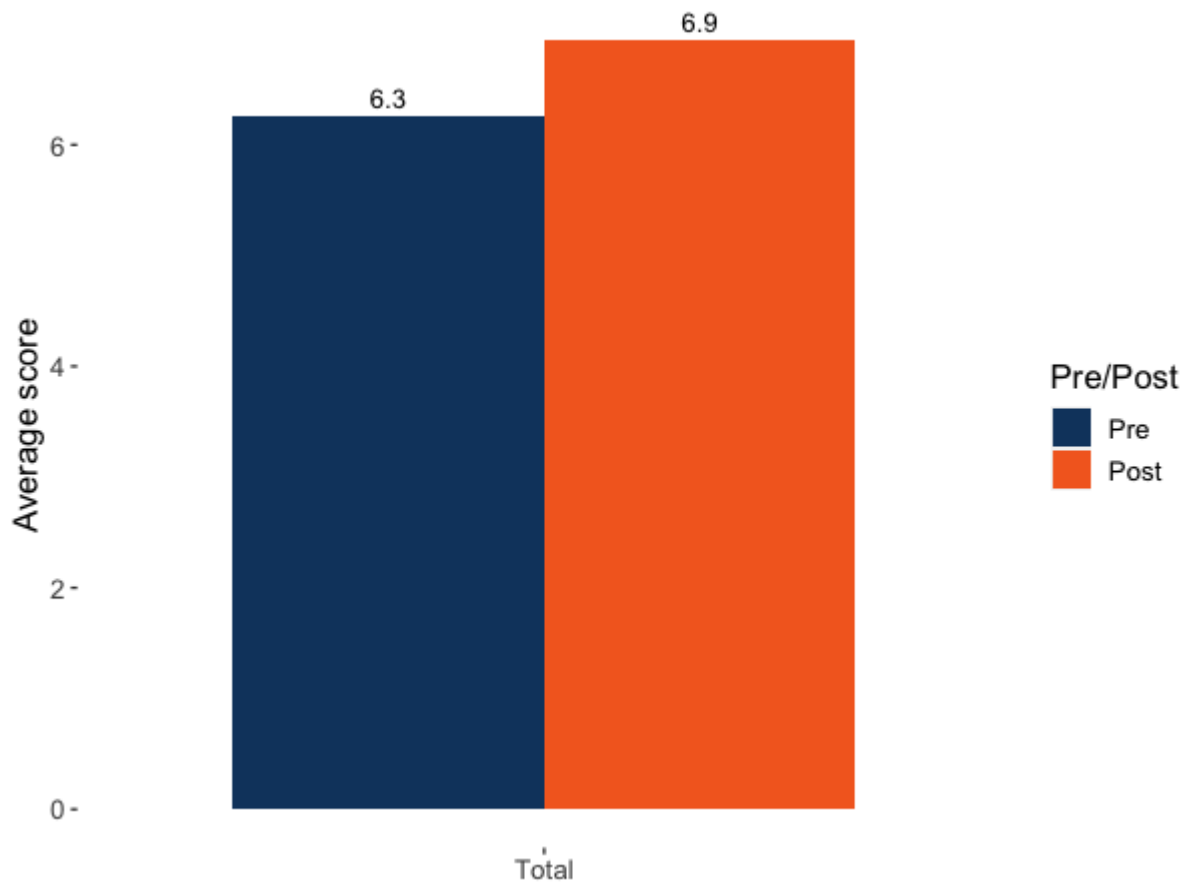
Notes:

1. Sample size: 190 in the pre-assessment, 96 in the post-assessment.
2. Table 2 based on responses to survey questions 1-5 (see Appendix 6 and 11 for more detail).

Data literacy

We found an increase of 6.3 percentage points in the average percentage of correct answers between the pre- and post-assessments. This corresponds to an increase of 0.7 out of a possible score of 11, from 6.26 to 6.95. The increase is statistically significant and suggests that, overall, there was a very small but detectable increase in assessment participants' data literacy across the period in which OBT 2023 took place. See Figure 6, below, for a visual representation. While we are confident that the results demonstrate that there was a very small, non-random change in overall literacy and numeracy scores across our 2 samples, this evaluation cannot tell us whether OBT 2023 had an influence on these scores, nor whether this trend was similar or different to any trends in data literacy that might have already existed prior to the period we studied.

Figure 6: Bar chart showing average scores out of a possible score of 11 in pre- and post-periods



Source: One Big Thing 2023 GPG assessment

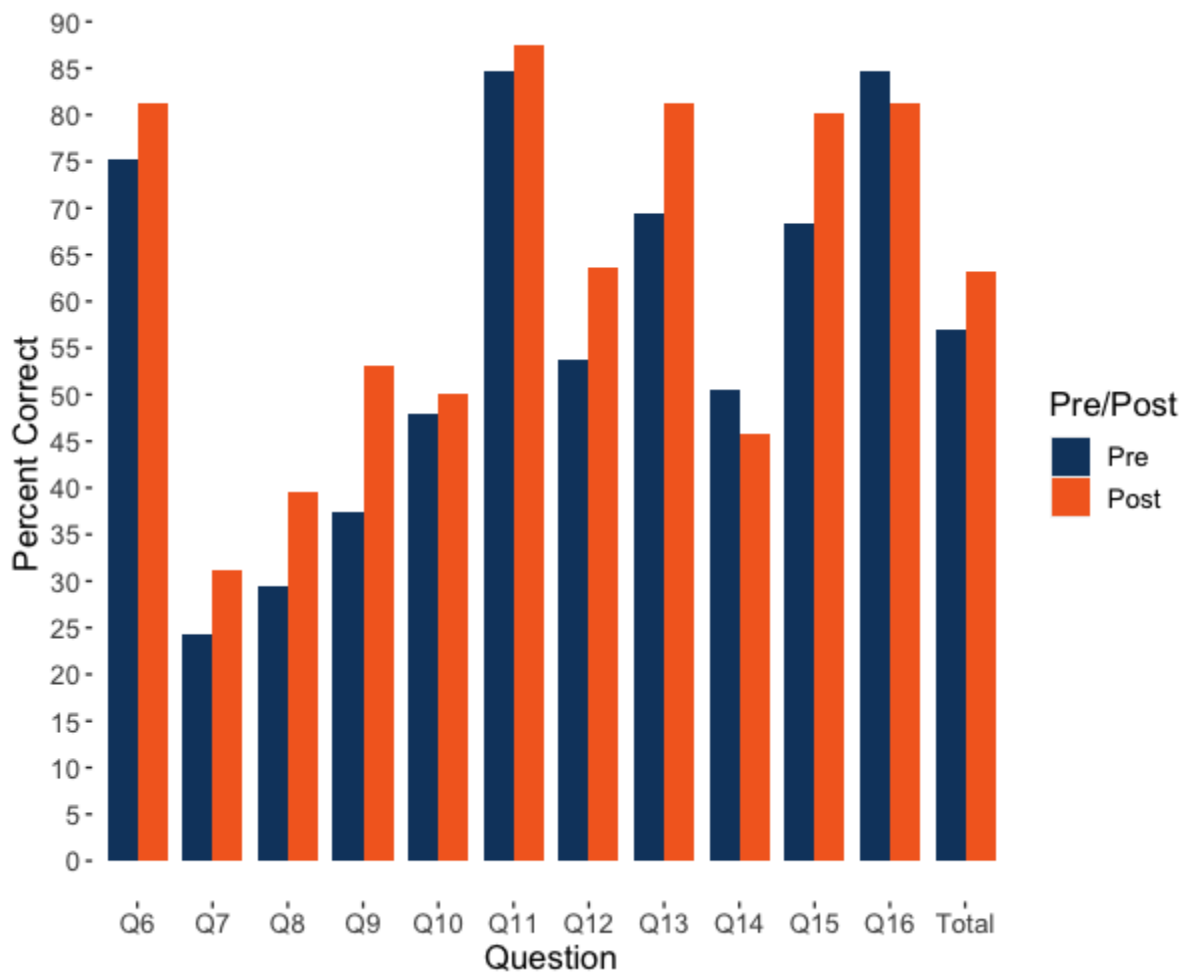
Notes:

1. Sample size: 190 in the pre-assessment, 96 in the post-assessment.
2. Figure 6 based on overall responses to survey questions (see Appendix 6 and 9 for more detail).

Analysis of performance on individual questions showed that though most questions saw an increase of correct answers, we only found statistically significant results for 3 of the data literacy questions. These were the questions on data understanding, data uses and one numerical reasoning question. There were 2 instances where correct answers had decreased (question 14 and 16, numerical reasoning), though this was not statistically significant. Figure 7 illustrates these results.

As mentioned in our analysis approach section, we are not confident in the practical significance of individual questions due to the elevated risk of false positives. While this does not rule out statistical significance for any given question, it does mean that isolating which questions are driving the overall change is difficult. Individual questions are represented as numbers in Figure 7, these numbers are mapped to the full question text in Table 3 below.

Figure 7: Bar chart showing percentage of correct answers by question number in both pre- and post-time periods



Source: One Big Thing 2023 GPG assessment

Notes:

1. Sample size: 190 in the pre-assessment, 96 in the post-assessment.
2. Figure 7 based on responses to survey questions (see Appendix 6 and 9 for more detail).

Table 3: Question numbers mapped to full question text

Question Number	Question
6	Which of the following options can data not do?
7	Which of the following is not an example of a step to consider when checking the quality of data?
8	What does a strong positive correlation imply?
9	Which measure is least affected by outliers?
10	Which plot would be best to visualise the relationship between study hours and exam score?
11	Which graph is best for showing how a variable changes over time?
12	A university researcher...The researcher wants to publish the results to document discrimination. What should they do?
13	Which of the following best describes the data type of a database table containing customer information like...?
14	When did the yellow car park become more popular than the blue car park?
15	Who has spent the most time interviewing?
16	How much overtime will Anil receive for last week before paying tax?

Study 2 discussion

RQ1: is there a measurable uplift in data literacy and behaviours for this business unit sample during the period coinciding with OBT?

We found a very small, measurable uplift in data literacy and data behaviours in our GPG sample following participation in OBT.

Data literacy

Overall, this study found a statistically significant difference for the average percentage of correct responses to data literacy questions between the pre- and post-assessments. This suggests that there was an improvement in data literacy within this sample during the period when OBT 2023 was live. However, these changes were of a very small magnitude; less than one correct answer. Increases in scores were observed for most individual data literacy questions, except for the numeracy section (questions 14 to 16) where results declined for 2 questions, although these changes were not statistically significant.

Data behaviours

There is some limited evidence that there may have been a very small increase in participants' interactions with data and discussions about data during the period coinciding with OBT 2023. Results here are not conclusive as, while statistically significant results were found in our sample, the changes found were very small and therefore of limited practical significance.

Results suggest that there was a small increase in reported use of data in making or suggesting decisions, policy or strategy, and the reported routine inclusion of data, facts or numbers in writing or policy recommendations during the period in which OBT 2023 was live.

The small changes we found in data literacy and behaviours cannot be directly attributed to OBT 2023. They may reflect a pre-existing trend, or be a consequence of other factors. There is a risk of a measurement error because we were not able to fully validate our assessment, and there is likely to be an unobserved variable at play, such as motivation, which may have influenced people's likelihood of filling out the post-assessment, and their assessment score.

Limitations

This study had a number of methodological limitations which are important when interpreting the results.

This evaluation cannot tell us whether OBT 2023 contributed to any upward trend in data awareness, confidence, knowledge and behaviours. It only tells us whether a statistically significant difference is there, not what caused it. It also does not tell us whether that trend is the same or different to any trend before or after the period we collected data about. Additionally, as we have mentioned before, a statistically significant difference only means that the difference was non-random and tells us very little about what the practical implications of this difference is.

The results of this study cannot tell us anything about whether OBT met its aims in the wider Civil Service. This study was conducted in one business unit (GPG). GPG is not representative of the wider Civil Service. The results only tell us something about whether OBT met its aims within GPG. This is demonstrated in Tables 4 and 5, below, where we can see that, on both gender and grade characteristics, the sample we obtained is different from the wider Civil Service.

Table 4: Overview of participants by gender and comparison to the wider Civil Service

Gender	% of assessment respondents	% of wider Civil Service	Percentage point difference
Female	62.2	54.6	7.7
Male	37.8	45.4	-7.7

Source: [One Big Thing 2023 GPG assessment/Civil Service Annual Statistical Bulletin: 2023](#)

Notes: based on a combined sample size of 286 from our assessment.

Table 5: Overview of participants by grade and comparison to the wider Civil Service

Grade	% of assessment respondents	% of wider Civil Service	Percentage point difference
AA-EO	16.4	50.8	-34.3
HEO-SEO	37.8	28.4	9.4
G7-G6	40.9	14.5	26.4
SCS	4.9	1.4	3.5

Source: [One Big Thing 2023 GPG assessment/Civil Service Annual Statistical Bulletin: 2023](#)

Notes: based on a combined sample size of 286 from our assessment.

This study had a relatively small sample size. This means that the evaluation is susceptible to false negatives, where we do not detect a change when there actually is one.

There is a risk of selection and non-response bias in the results, as the assessment was self-selective. OBT completion rates were analysed for the 2 assessment groups at the same time point following the post-assessment. These are outlined below in Table 6.

Table 6: Overview of engagement rates by respondent type

Respondent type	Percentage starting core OBT training, %	Percentage completing core OBT training, %
Pre-assessment respondents	64.7	47.9
Post-assessment respondents	66.7	56.3
Overall GPG mailing list	41.7	29.0

Source: CSL data using the GPG mailing list to identify participants.

Notes:

1. Sample size: 190 in the pre-assessment, 96 in the post-assessment. Overall GPG mailing list totalled 914.

As you can see, there are large differences in participation rates between those who took surveys and those who did not. When we exclude those who started but did not finish the training we can identify an imbalance between participation rates of those who took the pre-assessment and those who took the post-assessment. This suggests that those who responded to our assessment were different from the wider GPG, and those who responded to the post-assessment may, as a group, be different for certain unobservable characteristics. Our concern is that those who responded to the post-assessment may have been, on average, more motivated or interested in data than those who responded to the pre-assessment, hence the differing participation rates.

While we cannot be certain that this issue with unobservable differences is present, the differences in participation rates gives us reason to believe there is a high risk.

Study 3 - Interrupted Time Series Analysis of weekly volumes of data training attendance

Why did we carry out this study?

The objective of this study was to evaluate the third aim of OBT, whether OBT had a meaningful impact on civil servants' participation in data-related training.

Research question

Our research question was:

Did OBT 2023 have a long-term impact on participation in data and other training and initiatives?

Evaluation design and methods

The population for this study was anyone who participated in data or technology-related courses hosted on the online platform Civil Service Learning (CSL) in the years prior, during, and in the months following OBT.

The data source we used covers training available to all civil servants (free and paid-for training) via the specific platform CSL. As there is no single data source covering all training and skills-development initiatives undertaken by civil servants, we were only able to analyse data relating to CSL training. The OBT specific data courses were not included in our analysis.

We were not able to measure the long-term impact of OBT within the scope of this evaluation as we would need to wait longer to obtain more time points. However, we were able to assess the immediate and short term impacts of the intervention using weekly CSL data.

We used an Interrupted Time Series (ITS) design. We used CSL data from January 2021 up to the start date of OBT, September 2023, to produce a forecast. The data measure the volume of courses attended and are aggregated to a weekly frequency; we filtered the data to include only courses related to data and technology themes. Due to the wide window of time pre-OBT we were able to account for seasonality in our forecasts.

The forecast was then used as a counterfactual to the real data that we could observe during and immediately after the OBT intervention. The forecast is produced using an Auto Regressive integrated Moving Average (ARIMA) model¹⁴.

Data collection

Data were drawn from the CSL digital platform, an online platform Civil Servants use to record formal training. This does not include all potential training Civil Servants could participate in, and would not cover any informal training (for example, training delivered internally by team members, shadowing or open-source online courses). We filtered the data to include courses in the following themes:

- Artificial intelligence
- Data & Analytics
- Data and analytics
- Technology & Software
- Technology and software

We used data from 1st January 2021 to 5th March 2024.

Approach to analysis

¹⁴ We opted to use the `auto.arima` function in R to select the parameters of our final forecast model. The function uses Akaike information criterion (AIC), a measure of how well a model fits our data, to compare lots of different possible models - each with different parameters - and pick the one that has the best AIC value.

We use our forecast as a counterfactual to the actual observed data. In order to establish that OBT has had a meaningful impact on weekly CSL volumes we tested whether the actual volume lay outside of the forecast 95% confidence interval of the forecast for a suitably sustained period. The confidence interval represents the area where we are highly confident that the true volume would lie, had there been no intervention.

Results

We did not detect a meaningful change in the weekly volume of data-related Civil Service Learning (CSL) courses during or immediately after OBT. Actual recorded training participation during and after OBT fell within the 95% confidence intervals of our forecasted participation, had OBT not taken place.

There was high variance in the CSL data, which introduced uncertainty in our forecasted trend; this resulted in a wide forecast interval. The width of the interval makes it difficult to determine whether there are any movements in the volume of CSL courses that are attributable to OBT.

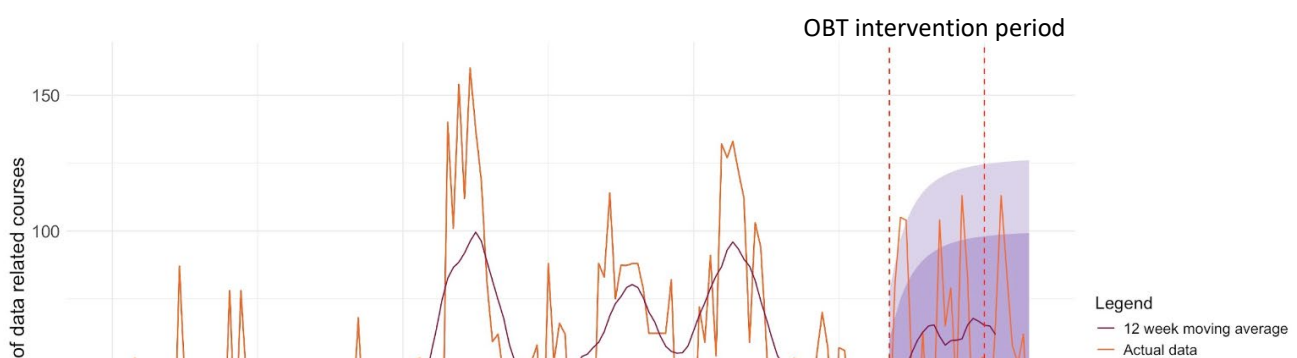
Additionally, many Civil Servants will have participated in courses outside of the CSL platform, for example, informal team-organised training. This will not be captured in the data. It is therefore possible that OBT caused an increase in non-CSL training that we did not observe.

Even if we had found data consistently outside of our forecast interval, the assumption that this was as a result of OBT is still fundamentally untestable. Our forecast provides an estimate of a counterfactual but, without an actual control group, does not allow us to rule out other external factors, excluding OBT, that may have had an impact on the observed data.

For the reasons outlined above; the uncertainty of the estimated forecast and potential training not recorded in the CSL platform, our findings are inconclusive.

Figure 8 shows a visualisation of our ITS analysis. The light purple line and shaded area indicate the point estimate for our forecast and the forecast interval, respectively. The observed data is represented by an orange line and we use a 12 week rolling average to provide a smoother plot, represented by the dark pink line.

Figure 8: Line graph showing actual observed data and our forecast.



Source: CSL digital platform records

Evaluation discussion and recommendations

Our evaluation feasibility study had 2 aims:

1. To generate initial results on whether OBT 2023 met its aims, to inform the planning of future OBT events.
2. To test out methods of evaluation to learn lessons for future OBT events.

Aim 1: Did OBT 2023 create a practical moment of shared participation to reinforce that we are one Civil Service?

Overall, our evaluation results suggest that OBT 2023 did generate participation across the Civil Service. Forty-two per cent of the Civil Service registered for OBT 2023 on the formal, online platform. Eighty-two per cent of those who registered completed the initial online training modules. This means that, of the whole civil service, around 1 in 3 (34%) signed up and completed some training. Around 567,000 data learning hours were recorded on the official platform. Of those who registered for OBT 2023, 23% recorded completing the target 7 hours of data upskilling (about 10% of civil servants).

This suggests that many people who registered for OBT 2023 may not have completed the programme. We do not have any data on civil servants who may have participated in local OBT 2023 activities, such as team discussions, but did not register on the online platform, or registered but did not log the upskilling activities they completed, so the above figures may underestimate overall participation.

We cannot say whether OBT 2023 was experienced as a 'shared moment' based on our data. We asked civil servants about their identity as a civil servant and their sense of connection with other civil servants before and after taking part in OBT, but the results were inconclusive and suggested OBT 2023 may not have had an influence on these issues.

Aim 2: Did OBT 2023 lead to a measurable uplift in data awareness, confidence, knowledge and understanding across the Civil Service?

Across the 2 studies assessing this aim, we found some very small positive improvements in participants' data awareness, confidence and knowledge.

Our cross-Civil Service survey assessed people's data awareness and confidence. We found very small increases in participants' awareness of the relevance and use of data in their data-to-day roles during the period in which they participated in OBT. We also found very small increases in their confidence around data-related ideas (such as ethics) and activities (such as visualising data).

In our case study evaluation we found very small increases in civil servants' ability to correctly answer some of the questions we set, which involved applying data to perform tasks (such as calculating something) and about key data-related concepts (such as averages). We also found small increases in reported use of data in writing and decision-making but did not find changes in all the data behaviours we asked about.

Overall, this suggests that OBT 2023 may have resulted in some very small gains in participants' data awareness, confidence and knowledge, including their ability to apply this knowledge to day-to-day work. Our results were statistically significant for the population of respondents. In other words, it is unlikely they occurred by chance, and are more likely to have been caused by something which occurred during the period when OBT 2023 was implemented, or reflect a pre-existing trend. Since OBT was a major Civil Service focus during this time period, it is plausible that there is a link between OBT and the very small changes we found, but it is not something we have been able to test statistically.

We also explored whether OBT 2023 participants found the training relevant. This could help explain why they felt their knowledge, confidence and awareness had improved, or not, after taking part in the training. It could also give an indicator of how likely it was that their use of data in day-to-day work would change after taking part in OBT. On average, we found that people only moderately agreed that OBT was a good use of their time and the content was relevant to their role. Participants moderately agreed that they were likely to apply learning from OBT 2023 to their roles, but recorded lower scores for their intention to complete specific actions such as booking further training or creating a personal development plan.

The evaluations may provide an indication of what sorts of effects it would be realistic to expect from a training intervention such as OBT. The effects that were found were very small. Researchers have pointed out that even small effects can have value, if we are able to achieve them across a whole organisation. This may, in fact, offer greater value than an intervention which achieves large

effects with much smaller groups of colleagues.¹⁵ Therefore, if a robust link can be made between OBT and small positive effects then it can still represent value if high levels of organisational participation can be achieved. The very small effects that we found are also relevant for planning future evaluations, as they suggest that large sample sizes will be needed to detect any statistically significant effects.

Aim 3: Did OBT 2023 have a long-term impact on participation in data and other training and initiatives?

Our findings for this aim were inconclusive. We did not detect a meaningful change in training courses attended when we analysed weekly volumes of data training courses hosted on the platform Civil Service Learning (CSL) during or immediately after OBT.

We cannot say whether OBT was successful or unsuccessful at achieving aim 3 due to two key confounding factors. These are: our inability to capture non-CSL training in our analysis, and the high variability of the CSL data creating higher levels of uncertainty in our forecast. We discuss these in more detail in the relevant section above.

Aim 4: Did OBT 2023 contribute towards achieving better outcomes in the delivery of public services and policy through the use of data?

Our evaluation did not assess this aim as it would require longer term data collection.

Our evaluations were oriented to testing whether OBT 2023 had met its aims. This meant that we only looked for a fixed range of outcomes. It is possible that OBT 2023 may have led to outcomes that were not captured in our evaluation. These could be identified in the lessons learned exercise, to be taken forward into the planning of future OBT events.

Lessons learned for future evaluations

¹⁵ Lacerenza, C.N., Reyes, D.L., Marlow, S.L., Joseph, D.L., and Salas, E. (2017) 'Leadership training design, delivery, and implementation: a meta-analysis', *The Journal of Applied Psychology*, 102(12), pp. 1686–1718. Available at: <https://doi.org/10.1037/apl0000241>

We found that it is possible to evaluate OBT events using a simple pre/post survey design. This type of design can measure changes in confidence, attitudes, self-perceived knowledge and skills, self-reported behaviours and, when objective assessments are used, knowledge and skills.

We found it was feasible to implement a case study evaluation into a supportive and engaged business unit (GPG). Our case study was possible due to the sponsorship of the GPG Executive Committee and Government Chief People Officer, as well as the participation of colleagues. While this did not provide results which generalised to the whole Civil Service, case studies can be an opportunity to explore issues in greater depth, and pilot additional evaluation approaches.

We found it difficult to evaluate the impact of OBT using CSL data for an interrupted time series analysis. We cannot measure other formal and informal activities that are still highly relevant to OBT through this method. This means that we are potentially missing much of the effect of OBT in this area. Additionally, the high variability in the data meant it was difficult to create an accurate forecast to act as a business-as-usual counterfactual.

Lessons learned on improving pre/post survey design

If a pre/post test design were used to evaluate OBT in the future, some enhancements could be made. The enhancements which would make the most difference are set out in Recommendation 3.

It is important to use valid, reliable assessment measures to evaluate OBT, or other training interventions. If bespoke measures are needed, it is advisable to ensure enough time and resources are available for their validation and piloting. This improves the chances of generating clear and valid results and reduces the risk of measurement error. A similar evaluation design to the one used in 2023 could be used again as an efficient route to evaluating OBT, but it would be improved by integrating the identification or design of validated measures into the implementation plan from the outset.

Our evaluations experienced a high rate of attrition between the pre-test and the post-test. Attrition was lower in GPG, where we used a very extensive programme of internal communications to promote participation and only required participants to fill in one assessment. This suggests that consideration of evaluation should be included in the communications strategies of future OBT events. User journeys on the digital platforms should also be considered. For example, if both pre-

and post-tests were embedded in the online learning content (as the pre-survey was for the cross-Civil Service evaluation), rates of completion might be higher. While attrition can be mitigated somewhat in the future, it will not be possible to ensure that everyone completes 2 surveys.

For any future evaluation, careful consideration would need to be given to sampling strategies, to try to generate results that applied to the whole Civil Service and therefore offered better insights into OBT's impact, implementation and value.

Lessons learnt on limitations to pre/post survey approach

The lack of a suitable comparison group who were not exposed to OBT (or were exposed at a later date), an aspect of the design of the OBT intervention, is a major limitation, as it prevents us from identifying any causal relationship between OBT and the very small changes seen. This means it is not possible to conclude whether OBT was responsible for generating these changes or not. Being able to attribute impact is an important prerequisite for robust calculations of value for money. A further limitation of our design is that the survey and assessment were opt-in, and we did not use random sampling, so our results do not generalise to the wider Civil Service.

Changes to the delivery model of OBT could be considered to make it feasible to design experimental evaluations that can test for causal relationships. For example, some civil servants could receive OBT earlier than others, or different versions of OBT could be rolled out to different groups. Care would need to be taken to manage the risks of inadvertent exposure to OBT (such as through cross-government communications, or online access to resources) among the comparison groups, to ensure an experimental evaluation was valid. The work on data interoperability across the Civil Service, which begins to go live during 2024/25, will also make quasi-experimental methods using organisational data to carry out evaluations more feasible than they were in 2023. These types of designs are suitable for answering some research questions, but also have limitations and challenges, which would need to be considered during planning and resourcing.

Our evaluation focused on assessing whether OBT 2023 met its aims. In future, it would be advisable to complement any impact evaluation with a process and economic evaluation, in line with HM Treasury Magenta Book guidance. These evaluations can build on work done in 2023, to ensure they add new insights and are not duplicative.

Lessons learned from the ITS analysis approach

For ITS analysis to be successfully implemented, the data used needs to be complete and of high quality. In order to conduct ITS analysis for future OBT interventions the quality and breadth of data we use will need to improve. Identifying broader indicators of training may be possible-learning spend for example-but will remain a challenge. If these fundamental limitations with the data we use cannot be overcome, an ITS approach may not be suitable for future evaluations of OBT.

Recommendations for OBT 2024

- 1. A training initiative such as OBT may be able to achieve very small increases in participant knowledge, awareness and confidence across a large number of civil servants. OBT design is likely to want to continue to identify the most effective ways of achieving positive effects.**

Our evaluations found very small improvements in data awareness, confidence and knowledge after taking part in OBT 2023. Even very small improvements may be valuable if they are achieved across a whole organisation. Previous evidence shows that small but widespread changes may in fact offer greater value than an intervention that achieves large effects with smaller groups of colleagues.¹⁶

Almost 220,000 civil servants took part in OBT 2023. As OBT matures as an annual initiative, and lessons are learned from implementation, it is likely that even higher participation rates could be achieved. There is much existing evaluation evidence to be drawn on when planning future OBT events, focused on different reform priorities, to plan an OBT design with the best chance of achieving the highest possible impact with the largest possible group of people.

- 2. The design of future OBT events could do more to support people to apply new learning to their day-to-day roles.**

OBT 2023 took evidence-based steps to support people to apply learning to their day-to-day role, by including line manager conversations and local, context-specific activities as part of the programme. This was a sensible place to start because these are relatively low-cost and simple to implement. Findings from the cross-Civil Service survey showed that there was still a gap between people's

¹⁶ Lacerenza, C.N., Reyes, D.L., Marlow, S.L., Joseph, D.L., and Salas, E. (2017) 'Leadership training design, delivery, and implementation: a meta-analysis', *The Journal of Applied Psychology*, 102(12), pp. 1686–1718. Available at: <https://doi.org/10.1037/apl0000241>

general intentions to use learning from OBT, and their intention to take specific, practical action to do so. Not everyone found OBT relevant. Some small changes in people's reported behaviours were found in our case study, but not across all behaviours.

In planning future OBT events, further attention could be given to connecting the upskilling content to specific local work and goals, to help people apply new learning in their day-to-day roles. For example, this might include more scenario-based content in the training,¹⁷ or providing evidence-based templates to support line managers help their teams embed the new skills within day-to-day work. These could include structured prompts and cues; action planning; self (or team) monitoring; and opportunities to continue to repeat the new skills within work.¹⁸

3. It is possible to evaluate OBT again in future, to gain even more extensive and robust evidence to support future delivery of OBT events and other upskilling initiatives.

OBT 2023 was the first of its kind in the Civil Service. Our evaluation has provided some useful lessons learned for how evaluations of future OBT events could be carried out. Overall, our evaluations show that it is feasible to evaluate OBT. The relatively light touch methods we have used to evaluate OBT 2023 (pre/post surveys and assessments) could be adapted, improved and used again to understand whether future OBT events achieve their aims. Other evaluation methods might also be considered, so the evaluation can be well-tailored to the strategic questions about OBT and Civil Service upskilling we need to answer. Planning and resourcing evaluation from the outset ensures that the widest possible range of suitable evaluation methods are available.

Specific evaluation decisions that may need to be taken early would include:

- whether to go beyond a simple pre/post survey evaluation for OBT 2024
- whether to change the delivery model for OBT to enable the establishment of a causal relationship between OBT and any impact identified
- whether to supplement impact evaluation with process and economic evaluation

¹⁷ Chernikova, O., Heitzmann, N., Stadler, M., Holzberger, D., Seidel, T., and Fischer, F. (2020) 'Simulation-based learning in higher education: a meta-analysis', *Review of Educational Research*, 90(4), pp. 499–541. Available at: <https://doi.org/10.3102/0034654320933544>

¹⁸ Sims, S., Fletcher-Wood, H., O'Mara-Eves, A., Cottingham, S., Stansfield, C., Van Herwegen, J., and Anders, J. (2021) *What are the characteristics of teacher professional development that increase pupil achievement? A systematic review and meta-analysis*. London: Education Endowment Foundation. Available at: <https://educationendowmentfoundation.org.uk/education-evidence/evidence-reviews/teacher-professional-development-characteristics>

- whether to issue evaluation advice or guidance to teams who wish to carry out their own case study evaluations of OBT in the future, and how any results could be used by the Cabinet Office to generate broader insights

If a pre/post survey is undertaken for OBT 2024, beneficial enhancements to the design we followed would be:

- identifying or designing validated measures aligned to the evaluation questions
- taking steps to drive up participation to the whole Civil Service
- using a sampling technique, to ensure results were generalisable to the whole Civil Service

In order for an interrupted time series design to be feasible for OBT 2024, we would need the following changes:

- a way to capture volumes of training in addition to those recorded through the CSL platform
- less variation in the historical data to provide a more accurate forecast
- identification of a different outcome variable, for which we are confident that we have robust and complete data

Due to the difficulty in achieving these two aims, it is unlikely that an interrupted time series design would work for evaluating OBT in the future.

APPENDICES

Appendix 1: Overview of the evaluation methods we considered

As this was a feasibility study, this Appendix sets out the evaluation methods we considered and why we did or did not use them.

We assessed that it would be possible to evaluate 2 of OBT’s 4 aims, and to partially evaluate a third objective. The final objective was judged out of scope for the evaluation.

Table 7: Which OBT aims was it feasible to evaluate?

Aim	Assessment
Aim 1: To create a practical ‘moment’ of shared participation to reinforce that we are one Civil Service.	Feasible to evaluate.
Aim 2: To have a measurable uplift in data awareness, confidence, knowledge and understanding across the Civil Service.	Feasible to evaluate.
Aim 3: To have a long-term impact on participation in data and other training and initiatives beyond OBT 2023.	Partially feasible to evaluate. This is because we can access central training data (training owned and procured through the Cabinet Office), but it would be very difficult to access all data on every formal and informal data learning and development initiative accessed across the whole of government.
Aim 4: To contribute towards achieving better outcomes in the delivery of public services and policy through the use of data.	Not feasible to evaluate in 2023. This was because this is a very complex outcome to evaluate, and would require long-term data collection and evaluation, which was out of scope for this project.

We identified several evaluation methods as potentially suitable for evaluating whether OBT 2023 met its aims, and assessed each of these for their feasibility. We focused on methods of impact evaluation as these were most suited to the challenge of assessing whether OBT 2023 met its aims. As OBT was a new initiative in 2023, being delivered at pace, it was necessary to focus our efforts to ensure we could implement the chosen evaluation methods and develop actionable insights in time for the next OBT.

Our assessments of the feasibility of these methods for OBT 2023 were based on the implementation plan for OBT 2023 specifically. This does not mean that some of these methods would not be suitable for evaluations of future OBT events

Table 8: The methods we considered for the cross-Civil Service evaluation

Impact evaluation method we considered for the cross-Civil Service evaluation	Did we use it? Why/not?
<p>Experimental method – randomised control trial (RCT).</p> <p>We considered a phased roll out of OBT ('wait-list control' design), where some civil servants were randomly selected to be exposed to OBT training earlier than others, and the waiting civil servants became a comparison group for those who took part first.</p> <p>We also considered a design where different groups of civil servants received different versions of OBT and we compared them.</p>	<p>Based on the aims and implementation plan for OBT 2023 it was not feasible to use an RCT design.</p> <p>As the goal was for all civil servants to participate in OBT at the same time, we could not construct a comparison group. As this was the first year of OBT, the priority was to develop a single approach to OBT, implement it and learn lessons from the implementation. In this context it was not feasible or desirable to roll out multiple different versions of OBT.</p> <p>As Civil Service workforce data is held at department level, there is no easily available cross-Civil Service sampling frame (central list of all civil servants) available to randomly select people for an RCT. This would have required a method of sampling based on clusters, where</p>

	<p>we chose particular departments, possibly even sampling within them. The implementation plan for OBT 2023 meant this was not feasible.</p> <p>Given the high volume of central communications about OBT it would have been difficult to stop the control group being exposed to OBT content. This would have invalidated any RCT.</p>
<p>Quasi-experimental method based on natural variation in who participated in OBT</p> <p>We looked for a random source of natural variation in who participated in OBT, as those individuals could be used as a comparison group.</p> <p>We identified that it might hypothetically be possible to use civil servants who had been off sick or on another type of leave (such as maternity leave), and therefore had not taken part in OBT for reasons independent of their attitudes towards OBT, as a comparison group.</p>	<p>Based on the implementation plan for OBT 2023, and the way Civil Service workforce data is stored, it was not feasible to use a quasi-experimental evaluation method based on natural variation.</p> <p>The decision to implement OBT over a 3-month period (rather than on a single day or shorter time period, as was initially planned) would have made it very challenging to identify those who had not been exposed to it at all for random reasons. This meant that this evaluation design would not have been valid.</p> <p>As workforce data is held locally within departments, and due to the need for careful data governance when using sensitive personal data, this type of evaluation design was not feasible for Cabinet Office teams to enact across the whole Civil Service within the OBT 2023 implementation plan.</p>
<p>Quasi-experimental method based on trends over time.</p>	<p>We implemented an interrupted time series of participation in centrally offered data training</p>

<p>We considered an interrupted time series analysis, based on a measure relevant to the aims of OBT. This design takes a regular measure, multiple times in the run up to OBT, during it and afterwards, then uses statistical techniques to see if OBT had an influence on the measure, based on the trend.</p>	<p>pre and post-OBT.</p> <p>We considered whether other measures of data attitudes, confidence, knowledge, skills or behaviours were available to carry out an interrupted time series analysis on other OBT aims. We could not identify any pre-existing measures being collected across the whole of government which would have been suitable within the timeframe available for planning and implementing the evaluation. We considered whether pulse surveys could have been used, but these did not align well with the implementation plan for OBT 2023, so were not feasible. There would also be a risk of diminishing response rates for a repeated survey instrument, making it less likely we would be able to generate valid results.</p>
<p>Simple pre/post test.</p> <p>We considered using a survey of participants in OBT before and after their participation, using questions aligned to the aims of OBT.</p>	<p>We implemented a simple pre/post test of all civil servants participating in OBT.</p> <p>We considered whether it would be possible to use a validated measure (meaning, one that had been tested for its validity and reliability). As OBT’s aims were bespoke, it was not possible to implement a validated measure. Instead, the i.AI team designed a survey aligned to the aims of OBT.</p> <p>For the same reasons outlined above, it was not possible to use a probability sampling approach or a comparison group. The survey was provided to all civil servants who participated in</p>

	OBT. They could opt into the survey, or opt out.
--	--

Table 9: The evaluation methods we considered for the case study evaluation

Evaluation methods we considered for the case study evaluation of OBT in GPG	Did we use it? Why/not?
<p>Experimental method – randomised control trial (RCT).</p> <p>We considered a ‘wait-list control’ design, where some GPG civil servants were randomly selected to get OBT earlier than others, and the waiting GPG civil servants became a comparison group for those who received it first.</p> <p>We also considered a ‘stepped wedge’ design where the roll out of OBT consisted of multiple phases, rather than just 2. This was potentially feasible within GPG because, as one business unit, more control could be exercised over the roll out, compared to the cross-Civil Service activities.</p>	<p>It was not feasible to implement an RCT within GPG to evaluate OBT.</p> <p>The time and resources required to design and run an RCT did not align with the implementation plan for OBT 2023.</p> <p>A further concern was the risk of contamination – this means that we could not stop people in the control group being exposed to OBT content during the evaluation, so they were no longer a valid comparison group. This was a particular concern in GPG because it was a single business unit, where colleagues from different teams work together in different office locations and on cross-team projects. This contamination would mean an RCT was not valid.</p> <p>A final concern was that it was not proportionate to invest in an RCT within GPG, because it would have limited external validity. External validity means whether these results would tell us something about the wider Civil Service. The GPG evaluation would not do this because GPG has quite different workforce characteristics to the rest of the Civil Service.</p>

<p>Quasi-experimental method based on trends over time</p> <p>We had already discounted an approach based on natural variation in who undertook OBT for the reasons outlined above. We considered an interrupted time series analysis using a pulse survey which captured a measure relevant to the aims of OBT.</p>	<p>It was not feasible to implement an interrupted time series analysis using a pulse survey.</p> <p>This was because there was not enough time to implement enough repetitions of the pulse survey ahead of OBT 2023 starting, to obtain a before-OBT trend. Anticipated attrition between surveys would introduce an unacceptable level of non-response bias. Furthermore, GPG's size means that non-response between surveys would impact power such that the identification of any effect would be unfeasible. There were no existing measures being regularly captured in GPG that we could use.</p> <p>As with the RCT, above, we had concerns about the proportionality of this approach, due to the lack of external validity, meaning this approach would not have generated results which told us something about the rest of the Civil Service.</p>
<p>Simple pre/post test of data knowledge, skills and behaviours</p> <p>We considered whether we could implement a simple pre/post test that measured something different to the cross-Civil Service evaluation.</p> <p>We were interested in whether we could assess uplift in GPG civil servants' data knowledge, skills and behaviours (as opposed to identity, awareness, attitudes and confidence, which were the focus of the cross-Civil Service evaluation).</p>	<p>We implemented a simple pre/post test of data knowledge, skills and behaviours within GPG, with no comparison group.</p> <p>We considered whether it was possible to use a validated test of data literacy and behaviours, and investigated several options. As we were unable to find a test which matched well with the aims of OBT, we designed our own test and conducted some limited validation of it.</p> <p>To reduce the risk of test-retest bias created by people taking the same test twice, we</p>

	<p>randomised GPG into 2 groups, with one taking the pre-test and one taking the post-test.</p>
<p>A qualitative evaluation using interviews and focus groups carried out after OBT 2023</p> <p>We considered using interviews and focus groups with GPG members after OBT 2023 had finished, and using these to gather richer data on people’s experiences of OBT 2023. This could have been used to test some of the assumptions the design of OBT 2023 was based on, and explore any outcomes not captured in the survey and assessment being used.</p>	<p>We did not carry out a qualitative evaluation of OBT using interviews and focus groups.</p> <p>This was because it was not well aligned with the agreed focus, which was to test whether OBT had met its aims.</p> <p>As GPG’s workforce has different characteristics to the rest of the Civil Service, it was not clear that a qualitative evaluation in GPG would have generated useful results for future OBT events.</p>

It is good practice to also include a process evaluation (an assessment of how effectively an activity has been implemented, and any lessons learned) and an economic evaluation (an assessment of whether the benefits of the evaluation outweigh its costs) in any evaluation project.

Appendix 2: Analytical approaches used to address research questions for Study 1

Table 10: Survey questions and analysis approaches used to address research questions for Study 1

Research question	Survey items	Analytical approach
RQ1: How many people took part in and completed the training between September and December 2023?	Measured based on the completion of the pre-survey (as this was the method of registering for OBT), not through individual survey items.	RQ1 was analysed descriptively using survey data and data from Civil Service Learning.
RQ2: Did participants' sense of shared Civil Service identity change after completing the OBT 2023 training?	Taking part in One Big Thing made me feel connected with other Civil Servants. My identity as a Civil Servant is important to me.	RQ2 and RQ3 were analysed by examining the differences in mean scores in pre- and post-survey responses.
RQ3: Did participants' data awareness, confidence or knowledge change after completing OBT 2023?	Pre/post questions: I feel confident about using data in my day-to-day role. I think data is relevant to my role. I know how to use data effectively in my day-to-day role. I am aware of how data can support my day-to-day role. Post-only questions: I have learned about the importance of evaluating the outcomes of data-informed decisions. I understand better how to	For RQ3, we also carried out descriptive analysis of a set of additional questions only asked in the post-survey.

	<p>communicate data insights effectively to influence decisions.</p> <p>I know more about visualising and presenting data in a clear and concise way.</p> <p>I have a better understanding of data ethics.</p> <p>I have a better understanding of how to quality assure data and analysis</p> <p>I have a better understanding of what data means.</p> <p>I know more about how different data analysis techniques can be used to understand data.</p> <p>I understand better how to critically assess data collection, analysis and the insights derived from it.</p> <p>I am better at interpreting data.</p> <p>I understand better how to anticipate data limitations and uncertainty.</p> <p>I feel more confident to use data to influence decisions.</p> <p>I can communicate data information more confidently to influence decisions.</p>	
--	---	--

<p>RQ4: After participating in OBT 2023, do participants believe they can apply the learning to their day-to-day role?</p>	<p>I have an improved understanding of how to use data in my day-to-day role.</p> <p>I am more interested in working with data in my day-to-day role.</p>	
<p>RQ5: After participating in OBT 2023, do participants intend to do anything differently at work as a result of the OBT training?</p>	<p>I intend to participate in further data training and initiatives.</p> <p>I intend to apply learning from this training in my role.</p> <p>I will find a mentor / become a mentor (practitioner).</p> <p>I will create a development plan.</p> <p>I will book a training course related to data / I will book a related training course.</p> <p>I will add a new area of learning to my development plan.</p>	<p>For RQ4 and RQ5, we carried out descriptive analysis using post-survey data only. This was due to the nature of the questions, which focus on how participants intend to respond to OBT.</p>

Appendix 3: Study 1 OBT survey

PRE-SURVEY

Data Protection

Responses will be stored securely for 3 years. We will not collect any direct personal identifiers, but combined data might make it identifiable. Data will be kept for 3 years to allow comparison of full data sets against future iterations of the survey. After this time, data will only be kept at unidentifiable levels. The raw data will only be accessed by the analysts working on the results.

By proceeding with this survey you are giving consent for the researchers to analyse your responses. Any analysis will be anonymised. Examples may be taken from free text but will not be attributed to an individual or department.

Under the General Data Protection Regulation (GDPR), the lawful basis for processing this information is your consent. This may be removed at any time.

Pre-survey

1. To what extent do you agree or disagree with the following statement?

a. I am aware of the aims of One Big Thing

1 = Strongly disagree

2 = Disagree

3 = Neither agree nor disagree

4 = Agree

5 = Strongly agree

2. To what extent do you agree or disagree with the following statements?

b. I feel connected with the wider Civil Service

c. My identity as a civil servant is important to me.

- 1 = Strongly disagree
- 2 = Disagree
- 3 = Neither agree nor disagree
- 4 = Agree
- 5 = Strongly agree

3. To what extent do you agree or disagree with the following statements?

- a. I feel confident about using data in my day-to-day role
- b. I think data is relevant to my role
- c. I know how to use data effectively in my day-to-day role
- d. I am aware of how data can support my day-to-day role

- 1 = Strongly disagree
- 2 = Disagree
- 3 = Neither agree nor disagree
- 4 = Agree
- 5 = Strongly agree

4a. Are you currently a line manager? If you answer 'Yes' please move onto 4b. (Y/N)

4b. If yes, to what extent do you agree or disagree with the following statements?

- i. I can help my team understand how data is relevant to their day-to-day roles
- ii. I know how to support my team to use data effectively in their day-to-day roles
- iii. I know how to coach team members to make better use of data in their day to day roles

- 1 = Strongly disagree
- 2 = Disagree
- 3 = Neither agree nor disagree
- 4 = Agree
- 5 = Strongly agree

5a. In the last 6 months, have you done any type of training? If yes, move onto 5b.

[Yes / No / I don't know]

5b. If yes, did it have an analytical component (e.g. data, evaluation)?

[Yes / No / I don't know]

POST-SURVEY

Introduction

Thank you for completing the 'One Big Thing' data training. As a final step, please complete the exit survey. This should take approximately 10 minutes. We are keen to understand your experience of One Big Thing to inform future design and delivery of learning and development for Civil Servants.

This survey will provide an opportunity to feedback on your experience, providing the One Big Thing team with valuable information. In the spirit of improving data usage in government, we are gathering this data to help our understanding of whether One Big Thing has been useful.

Exit survey

1. Which level of training did you participate in?
 - a. Awareness / Working / Practitioner / Don't know

2. Please rate how much you agree or disagree with the following statements:

- a. Taking part in One Big Thing made me feel connected with other Civil Servants
- b. My identity as a Civil Servant is important to me

1 = Strongly disagree

2 = Disagree

3 = Neither agree nor disagree

4 = Agree

5 = Strongly agree

What is data?

The word 'data' is used to describe a collection of facts or figures that can be used for decision making. In other words, data is information. For example, we collect a lot of facts and figures in government - prices, weights, addresses, ages, names, temperatures, dates, distances. These are all types of data.

When we organise, analyse and interpret data, it can help us develop a clear picture of a situation which allows us to make more accurate, informed decisions.

3. Following 'One Big Thing', please rate how much you agree or disagree with each of the following statements:

- a. I feel confident about using data in my day-to-day role
- b. I think data is relevant to my role
- c. I know how to use data effectively in my day-to-day role
- d. I am aware of how data can support my day-to-day role

1 = Strongly disagree

2 = Disagree

3 = Neither agree nor disagree

4 = Agree

5 = Strongly agree

4.

a. Are you currently a line manager? If you answer 'Yes' please move onto 3b. If 'No', please move onto question 4. (Y/N)

b. If yes, to what extent do you agree or disagree with the following statements?

- i. I can help my team understand how data is relevant to their day-to-day roles
- ii. I know how to support my team to use data effectively in their day-to-day roles
- iii. I know how to coach team members to make better use of data in their day to day roles

1 = Strongly disagree

- 2 = Disagree
 - 3 = Neither agree nor disagree
 - 4 = Agree
 - 5 = Strongly agree
-

2. Awareness (level 1):

a. Following 'One Big Thing', please rate how much you agree or disagree with each of the following statements (5-point scale):

- i. I have a better understanding of what data means
- ii. I am better at interpreting data
- iii. I am more interested in working with data in my day-to-day role
- iv. I feel more confident to use data to influence decisions
- v. I can communicate data information more confidently to influence decisions

- 1 = Strongly disagree
- 2 = Disagree
- 3 = Neither agree nor disagree
- 4 = Agree
- 5 = Strongly agree

b. Following 'One Big Thing', to what extent do you agree or disagree with the below statements (5-point scale)?

- i. I will create a development plan
- ii. I will add a new area of learning to my development plan
- iii. I will book a training course related to data
- iv. I will find a mentor
- i. Other (please specify)

- 1 = Strongly disagree
- 2 = Disagree
- 3 = Neither agree nor disagree
- 4 = Agree
- 5 = Strongly agree

c. Working (level 2)

Following 'One Big Thing', please rate how much you agree or disagree with each of the following statements (5-point scale):

- i. I know more about how different data analysis techniques can be used to understand data
- ii. I understand better how to critically assess data collection, analysis and the insights derived from it
- iii. I know more about visualising and presenting data in a clear and concise way
- iv. I have learned about the importance of evaluating the outcomes of data-informed decisions

1 = Strongly disagree

2 = Disagree

3 = Neither agree nor disagree

4 = Agree

5 = Strongly agree

d. Following 'One Big Thing', to what extent do you agree or disagree with the below statements (5-point scale)?

- i. I will create a development plan
- ii. I will add a new area of learning to my development plan
- iii. I will book a related training course
- iv. I will find a mentor
- ii. Other (please specify)

1 = Strongly disagree

2 = Disagree

3 = Neither agree nor disagree

4 = Agree

5 = Strongly agree

e. Practitioner (level 3)

Following 'One Big Thing', please rate how much you agree or disagree with each of the following statements (5-point scale):

- i. I understand better how to communicate data insights effectively to influence decisions
- ii. I have a better understanding of how to quality assure data and analysis
- iii. I understand better how to anticipate data limitations and uncertainty
- iv. I have a better understanding of data ethics

1 = Strongly disagree

2 = Disagree

3 = Neither agree nor disagree

4 = Agree

5 = Strongly agree

- f. Following 'One Big Thing', to what extent do you agree or disagree with the below statements (5-point scale)?

I will create a development plan

I will add a new area of learning to my development plan

I will book a related training course

I will find a mentor/become a mentor

Other (please specify)

1 = Strongly disagree

2 = Disagree

3 = Neither agree nor disagree

4 = Agree

5 = Strongly agree

6.

- a. Please rate how much you agree or disagree with the following statements (5-point scale).

- i. The online training helped my learning
- ii. Conversations with my team or line manager helped my learning
- iii. Additional learning resources helped my learning

- 1 = Strongly disagree
- 2 = Disagree
- 3 = Neither agree nor disagree
- 4 = Agree
- 5 = Strongly agree

b. Additional learning

- i. Were there any formats of additional training you found useful?

[Drop down of options included on i.AI app]

7.

- a. Please rate how much you agree or disagree with the following statements (5-point scale):

- i. The OBT training was a good use of my time
- ii. I have an improved understanding of how to use data in my day-to-day role
- iii. I intend to participate in further data training and initiatives
- iv. The content was relevant to my role
- v. I intend to apply learning from this training in my role

- 1 = Strongly disagree
- 2 = Disagree
- 3 = Neither agree nor disagree
- 4 = Agree
- 5 = Strongly agree

8.

- a. Please rate how much you agree or disagree with the following statements (5-point scale).

- i. I am aware of the aims of One Big Thing
- ii. I had sufficient time to participate in One Big Thing during the autumn (September - December)

- 1 = Strongly disagree

- 2 = Disagree
- 3 = Neither agree nor disagree
- 4 = Agree
- 5 = Strongly agree

9. What, if anything, went well about the training? Please use the space below to provide more details.

Free text (150 words max)

10. Was there anything that could have been improved? Please use the space below to provide more details.

Free text (150 words max)

11. Would you be willing to take part in a follow-up discussion?

Yes / No

Appendix 4: Table of Study 1 results

The below table summarises all Study 1 results. The ‘main analysis sample’ portion of this table captures all results presented in the main report for Study 1. The full sample (including people with 17+ hours) portion of the table presents results that include survey responses for people who reported more than 17 hours’ worth of training, and were excluded from the main analysis sample. The table illustrates that the exclusion of these individuals from the main analysis sample makes little difference to the findings.

Table 11: Study 1 results

Research question	Survey question	Main analysis sample					Full sample (including people with 17+ hours)				
		Mean outcome pre-OBT	Mean outcome post-OBT	Difference	p-value	n	Mean outcome pre-OBT	Mean outcome post-OBT	Difference	p-value	n
Did participants’ sense of shared Civil Service identity change after completing the OBT 2023 training?	Taking part in One Big Thing made me feel connected with other civil servants	2.97	2.94	-0.03	0.025	31,437	2.97	2.94	-0.04	0.007	32,559
	My identity as a civil servant is important to me	3.61	3.67	0.06	0.00	31,437	3.61	3.67	0.06	0.00	32,559
Is there any change in participants’ data awareness, confidence and knowledge?	I feel confident about using data in my day-to-day role	3.88	3.99	0.12	0.00	31,437	3.88	4.00	0.12	0.00	32,559
	I think data is relevant to my role	4.12	4.18	0.07	0.00	31,437	4.13	4.19	0.07	0.00	32,559
	I know how to use data effectively in my day-to-day	3.83	4.04	0.21	0.00	31,437	3.84	4.05	0.21	0.00	32,559

		Main analysis sample					Full sample (including people with 17+ hours)				
Research question	Survey question	Mean outcome pre-OBT	Mean outcome post-OBT	Difference	p-value	n	Mean outcome pre-OBT	Mean outcome post-OBT	Difference	p-value	n
	role										
	I am aware of how data can support my day-to-day role	3.98	4.12	0.14	0.00	31,437	3.99	4.13	0.14	0.00	32,559
Is there any change in participants' data awareness, confidence and knowledge? (Insights)	I have learned about the importance of evaluating the outcomes of data-informed decisions	N/A	3.81	N/A	N/A	31,437	N/A	3.81	N/A	N/A	32,559
	I understand better how to communicate data insights effectively to influence decisions	N/A	3.76	N/A	N/A	31,437	N/A	3.76	N/A	N/A	32,559
	I know more about visualising and presenting data in a clear and concise way	N/A	3.75	N/A	N/A	31,437	N/A	3.75	N/A	N/A	32,559
	I have a better understanding of data ethics	N/A	3.80	N/A	N/A	31,437	N/A	3.80	N/A	N/A	32,559
	I have a better understanding of how to	N/A	3.74	N/A	N/A	31,437	N/A	3.74	N/A	N/A	32,559

		Main analysis sample					Full sample (including people with 17+ hours)				
Research question	Survey question	Mean outcome pre-OBT	Mean outcome post-OBT	Difference	p-value	n	Mean outcome pre-OBT	Mean outcome post-OBT	Difference	p-value	n
	quality assure data and analysis										
	I have a better understanding of what data means	N/A	3.79	N/A	N/A	31,437	N/A	3.79	N/A	N/A	32,559
	I know more about how different data analysis techniques can be used to understand data	N/A	3.74	N/A	N/A	31,437	N/A	3.74	N/A	N/A	32,559
	I understand better how to critically assess data collection, analysis and the insights derived from it	N/A	3.72	N/A	N/A	31,437	N/A	3.72	N/A	N/A	32,559
	I am better at interpreting data	N/A	3.56	N/A	N/A	31,437	N/A	3.56	N/A	N/A	32,559
	I understand better how to anticipate data limitations and uncertainty	N/A	3.74	N/A	N/A	31,437	N/A	3.74	N/A	N/A	32,559
	I feel more	N/A	3.61	N/A	N/A	31,437	N/A	3.61	N/A	N/A	32,559

		Main analysis sample					Full sample (including people with 17+ hours)				
Research question	Survey question	Mean outcome pre-OBT	Mean outcome post-OBT	Difference	p-value	n	Mean outcome pre-OBT	Mean outcome post-OBT	Difference	p-value	n
	confident to use data to influence decisions										
	I can communicate data information more confidently to influence decisions	N/A	3.55	N/A	N/A	31,437	N/A	3.56	N/A	N/A	32,559
Do participants think they can apply the learning to their day-to-day role?	I have an improved understanding of how to use data in my day-to-day role	N/A	3.44	N/A	N/A	31,437	N/A	3.44	N/A	N/A	32,559
	I am more interested in working with data in my day-to-day role	N/A	3.46	N/A	N/A	31,437	N/A	3.46	N/A	N/A	32,559
Are participants intending to do anything differently as a result of OBT?	I intend to participate in further data training and initiatives	N/A	3.40	N/A	N/A	31,437	N/A	3.41	N/A	N/A	32,559
	I intend to apply learning from this training in my role	N/A	3.55	N/A	N/A	31,437	N/A	3.56	N/A	N/A	32,559

Research question	Survey question	Main analysis sample					Full sample (including people with 17+ hours)				
		Mean outcome pre-OBT	Mean outcome post-OBT	Difference	p-value	n	Mean outcome pre-OBT	Mean outcome post-OBT	Difference	p-value	n
	I will find a mentor / I will find a mentor/become a mentor (practitioner)	N/A	2.83	N/A	N/A	31,437	N/A	2.83	N/A	N/A	32,559
	I will create a development plan	N/A	3.16	N/A	N/A	31,437	N/A	3.17	N/A	N/A	32,559
	I will book a training course related to data / I will book a related training course	N/A	3.13	N/A	N/A	31,437	N/A	3.14	N/A	N/A	32,559
	I will add a new area of learning to my development plan	N/A	3.24	N/A	N/A	31,437	N/A	3.25	N/A	N/A	32,559

Source: One Big Thing 2023 pre-survey

Notes:

1. 'Main analysis sample' refers to the final sample used for analysis, excluding those who did not respond to the post-survey and those who recorded more than 17 hours of training (n=31,437). 'Full sample' includes those who recorded more than 17 hours (n = 32,559).
2. A 5-point Likert scale was used. 1 = Strongly disagree; 2 = Disagree; 3 = Neither agree nor disagree; 4 = Agree; 5 = Strongly agree. Mean scores of 3.1-5 can broadly be interpreted as agreement and mean scores of 1-2.9 can be interpreted as disagreement with the statement. Three is a neutral score (neither agree nor disagree), and responses close to 3 are also interpreted as neutral.

Appendix 5: Descriptive analysis exploring the representativeness of the sample for Study 1

Study 1 uses a sample of civil servants that responded to both the pre- and post-survey. This sample is not representative of all civil servants, since people could opt out of taking part in either survey. In this Appendix we carry out some descriptive analysis to illustrate the extent to which the Study 1 analysis sample is similar or different to the wider population of civil servants. This helps us to assess whether the final analysis sample is composed of people who have similar average characteristics to the wider Civil Service, or were very different.

Table 12 compares the differences in pre-survey responses between respondents in the main analysis sample (n=31,437) and the full pre-survey sample (n=218,583). Here we observe differences in mean responses of between 0.01 and 0.07. The larger differences relate to confidence and knowledge, with smaller differences related to data awareness and shared identity.

Table 12: Descriptive analysis exploring representativeness of Study 1 sample

Research question	Survey question	Main analysis sample		All pre-survey respondents		Difference
		Mean outcome pre-OBT	n	Mean outcome pre-OBT	n	
Did participants' sense of shared Civil Service identity change after completing the OBT 2023 training?	Taking part in One Big Thing made me feel connected with other Civil Servants	2.97	31,437	2.92	218,583	0.05
	My identity as a Civil Servant is important to me	3.61	31,437	3.60	218,583	0.01
Is there any change in participants' data awareness, confidence and	I feel confident about using data in my day-to-day role	3.88	31,437	3.81	218,583	0.07
	I think data is relevant to my role	4.12	31,437	4.08	218,583	0.04

knowledge?	Know how to use data effectively	3.83	31,437	3.75	218,583	0.07
	Awareness of how data can support role	3.98	31,437	3.93	218,583	0.05

Source: One Big Thing 2023 pre-survey

Notes

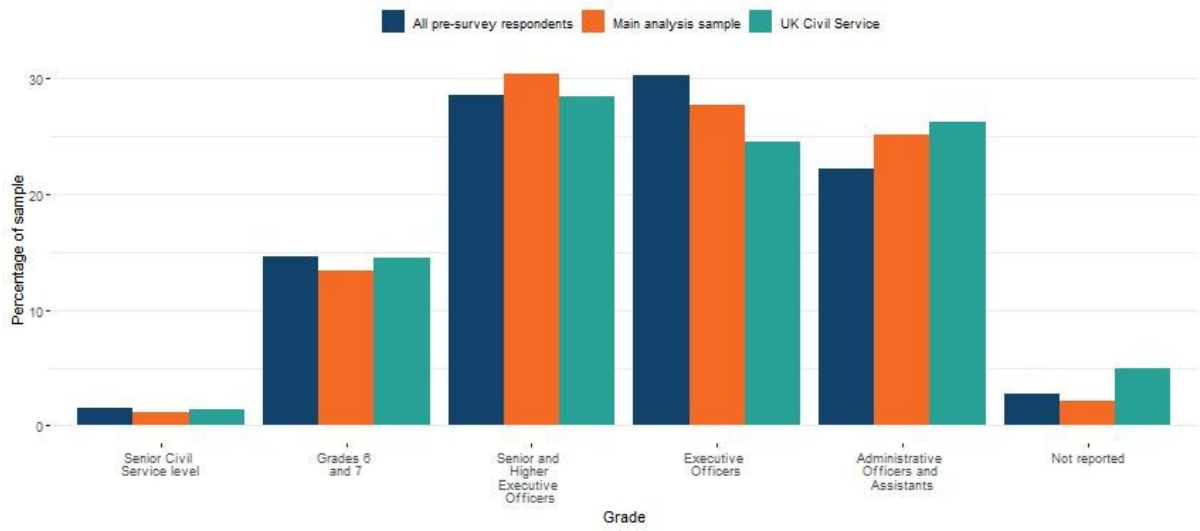
1. 'Main analysis sample' refers to the final sample used for analysis, excluding those who did not respond to the post-survey and those who recorded more than 17 hours of training (n=31,437). 'All pre-survey respondents' refers to all those who responded to the pre-survey (n=218,583).
2. A 5-point Likert scale was used. 1 = Strongly disagree; 2 = Disagree; 3 = Neither agree nor disagree; 4 = Agree; 5 = Strongly agree. Mean scores of 3.1-5 can broadly be interpreted as agreement and mean scores of 1-2.9 can be interpreted as disagreement with the statement. Three is a neutral score (neither agree nor disagree), and responses close to 3 are also interpreted as neutral.

In the chart below we explore how the grade structure of the OBT analysis sample compares with the wider Civil Service. The chart shows the following groups:

- the main OBT analysis sample (n=31,437)
- all respondents who completed the pre-survey (n=218,583)
- the wider Civil Service (n=519,780)

For each sample group, the chart shows the proportion of that group belonging to each grade structure. This illustrates that our main analysis sample is largely similar to the wider UK Civil Service population in terms of grade. The analysis sample includes a higher proportion of Senior and Higher Executive Officers than the wider Civil Service, but the differences are not large. Note that there may be other differences between our sample and the overall population that we have not been able to analyse with the available data.

Figure 8: Comparison of the grade structure in the Study 1 main analysis sample and wider Civil Service



Appendix 6: Study 2 data literacy and behaviours assessment (One Big Thing knowledge check for GPG)

Introduction.

We are delighted to welcome you to the One Big Thing (OBT) evaluation knowledge check!

The purpose of this survey is to help determine the capabilities of Government People Group staff **after** the roll out of OBT. You have been randomly selected to take part in the knowledge check after the course content was distributed, and your results will be compared with the pre-course knowledge completed by your peers.

If you haven't taken part in One Big Thing, you can access the content here:

onebigthing.civilservice.gov.uk. Please continue with the survey if you have completed part of the training - even if you haven't completed the full 7 hours.

This knowledge check should only take **10 minutes** to complete.

As the aim is to measure your current knowledge, please complete the questionnaire quickly and by responding with the answer that first comes to mind. **Please do not discuss the questions with anyone else or search for answers online.**

Your responses are completely anonymous and will not be shared with anyone outside the research team. *We will ask for your email address which will be used to match responses to actual course participation data provided by the digital learning team.* However, all reporting will be based on aggregated data.

We are extremely grateful for your participation, which will provide insights to aid continuous development and improvement in learning for all Civil Servants.

[Privacy Notice](#)

Data will be kept in strictest confidence, anonymised and confidential. It is also advised that you do not enter any identifying information in the available free text boxes. Data will not be used to

identify individual persons and will be kept in accordance with Cabinet Office and GDPR data security requirements. Please find a link to Privacy Notice [internal link provided].

If you have any questions about the survey, please contact the Data services, Analysis, and Research team (DART) in CS Data and Insights [email address provided].

Thank you, once again, for your participation, which will provide insights to aid continuous development and improvement in learning for all Civil Servants.

One Big Thing Knowledge Check

Q1. How many data reports, dashboards or visualisations have you interacted with in the last week?

Examples of interaction could include viewing, taking information or insight from, sharing with others, or creating.

Q2. How many times this week have you discussed data trends, metrics or insights with colleagues?

Please provide an estimate as a numerical response.

Q3. How many hours this week did you spend analysing, manipulating or interpreting data as part of your regular job functions?

Please provide an estimate as a numerical response, you can use decimals, e.g. 0.5 for half an hour.

Q4. Have you made or suggested a decision, policy or strategy based on data analysis in the past week?

Yes

No

Q5. Do you routinely include data, facts or numbers in your writing/policy recommendations?

Yes

No

Q6. Which of the following options can data **not** do?

- Help inform decision-making
- Provide insights
- Allow for greater understanding of a topic
- Provide definitive answers to questions
- Don't know

Q7. Which of the following is **not** an example of a step to consider when checking the quality of data?

- Contacting the data owner(s) and requesting documentation related to the dataset
- Keeping a log of what you are doing
- Considering how the data will be used to provide insights, and tailoring analysis towards that
- Spot checking the data to see if there is anything that looks unusual
- Don't know

Q8. What does a strong positive correlation imply?

- Change in one variable causes change in the other
- Variables increase together
- No relationship between variables
- Negative relationship between variables
- Don't know

Q9. Which measure is least affected by outliers?

- Mean
- Range
- Standard deviation
- Median
- Don't know

Q10. Which plot would be best to visualise the relationship between study hours and exam score?

- Bar chart
- Pie chart
- Scatterplot
- Time series
- Don't know

Q11. Which graph is best for showing how a variable changes over time?

- Bar chart
- Pie chart
- Line graph
- Scatterplot
- Don't know

Q12. A university researcher conducts a study using students' academic records without their knowledge or prior consent. The study finds certain demographic groups are more likely to fail courses than others. The researcher wants to publish the results to document discrimination. What should they do?

- Publish the full analysis to influence reforms, as academic failure harms students
- Publish only aggregated statistics without identifiers to balance benefits and privacy risks
- Do not publish as consent is required to analyse individual academic records
- Publish select examples with student names redacted to personalise the issue
- Don't know

Q13. Which of the following best describes the data type of a database table containing customer information like name, address, phone number, email, etc?

- Unstructured data
- Structured data
- Time-series data
- Audio data
- Don't know

Q14. The table below shows the average number of cars using 4 car parks between 2014 and 2019. Each are recorded in thousands.

Car Park	2014	2015	2016	2017	2018	2019
Blue	50	40	50	35	30	50
Red	40	40	30	30	25	45
Yellow	45	35	40	50	50	55
Green	20	30	25	25	20	30

When did the yellow car park become more popular than the blue car park?

Between 2014 and 2015

Between 2015 and 2016

Between 2016 and 2017

Between 2017 and 2018

Between 2018 and 2019

Don't know

Q15. The first table below provides information on the interviews scheduled and completed. The second table presents the gender by age breakdown of the people interviewed.

Name	Interviews Scheduled	Interviews Completed	Average interview length	Work Completed
Rachel	23	19	31	0.83
Karim	33	23	22	0.70
John	28	16	26	0.57
Geraint	27	14	32	0.52
Ekaterina	33	20	29	0.61

Age	Male	Female
18-24	13	3
25-30	12	8
31-36	29	7
37-40	0	13
40-50	0	7

Who has spent the most time interviewing?

Rachel

Karim

John

Geraint

Ekaterina

Don't know

Q16.

Anil has worked overtime last week and is trying to work out how much in overtime he will be paid.

He receives an hourly wage of £12.61.

Any overtime paid between Monday and Friday receives the hourly rate.

The rate for Saturday is one and a half times the hourly rate.

The rate for Sundays is double the hourly rate.

The table below shows the number of hours that Anil worked overtime last week

Day	Hours attracting overtime
Monday	2
Tuesday	1
Wednesday	0
Thursday	3
Friday	2
Saturday	7
Sunday	7

How much overtime will Anil receive for last week before paying tax?

£277.42

£365.69

£409.83

£453.96

£510.71

Don't know

One Big Thing Knowledge Check

Q17. What is your email address?

This will be used to link your responses to your CSL learning record, your responses will remain anonymous and will not be shared with anyone outside the immediate research team. You can opt out of this if you prefer.

Q18. How would you describe your gender?

Male

Female

Other

Q19. What is your grade?

AA-EO

HEO-SEO

G7-G6

SCS

Q20. What is your profession?

Analysis Commerical and Procurement Corporate Finance Communications Counter Fraud Digital, Data and Technology Government Economic Service Government Social Research Government Statistical Service Finance Human Resources Intelligence Analysis Internal Audit International Trade and Negotiation Knowledge & Information Management Legal Medical Occupational Psychology Operational Delivery Operational Research Planning Policy Project Management and Delivery Property Science and Engineering Security Tax Veterinary No profession Don't know

Q21. Have you taken part in the One Big Thing e-learning?

No

Yes

One Big Thing Knowledge Check

Q23. One Big Thing involves 7 hours of data training, how many hours of training have you completed?

Q24. Did you have a team conversation about data?

Yes

No

Q25. Do you have any comments/ feedback about this knowledge check?

[Survey Completion](#)

0%

Current Progress 85% 1

0

0

%

One Big Thing Knowledge Check

We thank you for your time spent taking this survey.

Your response has been recorded.

Appendix 7: Study 2 power calculations

Before deciding on a methodology, a power analysis was conducted to determine the number of participants needed based on effect size. This suggested that the anticipated sample size of around 200 participants would detect effect sizes of 0.4 and above. This was predicted based on a 2-tailed test at a .05 significance level. Please see below for added context.

Table 13: Study 2 power calculations

Expected effect size	Minimum total sample size - two-tailed test
0.1	3,124
0.2	788
0.3	352
0.4	200
0.5	128

Appendix 8: Study 2 factor analysis

Factor analysis was used to assess construct validity for the questions relating to data literacy. Correlations across questions were tested, and Cronbach's alpha score was calculated as 0.73. Cronbach's alpha, if other items are deleted, is either equal to or below 0.73. Results of a polychoric factor analysis were commensurate with a one-factor solution, explaining 33% of the variance in scores. This singular construct is referred to as 'data literacy' throughout the main body of the report

Appendix 9: Study 2 detailed results

Impact evaluation findings

We are comparing between the pre and post groups:

1. Average % of correct responses for questions 6-16.
2. Median responses for questions 1-3.
3. % yes for questions 4 and 5.
4. Overall average % correct for questions 6-16 aggregated.

Table 14: Study 2 detailed results

Question	Median pre	Median post	Significant?	Test	Test stat	P-value
1 How many data reports, dashboards or visualisations have you interacted with in the last week?	3	3	Y	Wilcoxon Rank-sum		0.47
2 How many times this week have you discussed data trends, metrics or insights with colleagues?	2	3	Y	Wilcoxon Rank-sum		0.49
3 How many hours this week did you spend analysing, manipulating or interpreting data as part of your regular job functions?	2	2	N	Wilcoxon Rank-sum		0.48

Question	Median % yes pre	Median % yes post	Significant?	Test	Test stat	P-value
4 Have you made or suggested a decision, policy or strategy based on data analysis in the past week?	45.8	59.4	Yes	Chi-square	4.18	0.041
5 Do you routinely include data, facts or numbers in your writing/policy recommendations?	68.2	82.3	Yes	Chi-square	5.55	0.018

Question	% correct	% correct post	Significant?	Test	Test stat	P-value
----------	-----------	----------------	--------------	------	-----------	---------

pre						
6 Which of the following options can data not do?	75.3	81.3	N	Chi-square	0.98	0.32
7 Which of the following is not an example of a step to consider when checking the quality of data?	24.2	31.3	N	Chi-square	1.28	0.26
8 What does a strong positive correlation imply?	29.5	39.6	N	Chi-square	2.51	0.11
9 Which measure is least affected by outliers?	37.4	53.1	Y	Chi-square	5.84	0.016
10 Which plot would be best to visualize the relationship between study hours and exam score?	47.9	50.0	N	Chi-square	0.04	0.83
11 Which graph is best for showing how a variable changes over time?	84.7	87.5	N	Chi-square	0.20	0.65
12 A university researcher...The researcher wants to publish the results to document discrimination. What should they do?	53.7	63.5	N	Chi-square	2.14	0.14
13 Which of the following best describes the data type of a database table containing customer information like...?	69.5	81.3	Y	Chi-square	3.95	0.047

14 When did the yellow car park become more popular than the blue car park?	50.5	45.8	N	Chi-square	0.39	0.53
15 Who has spent the most time interviewing?	68.4	80.2	Y	Chi-square	3.86	0.049
16 How much overtime will Anil receive for last week before paying tax?	84.7	81.3	N	Chi-square	0.34	0.56
Overall	56.9	63.2	Y	T-test	2.64	0.0089

Appendix 10: Quality assurance for all studies

Analysis of data was conducted in R, a statistical software environment. All code and outputs were checked internally by each of the 2 teams (the Evaluation Task Force and the GPG Civil Service Data and Insights team). They were then cross checked by the other team involved in this project. Finally, a UKRI policy fellow working in the Cabinet Office performed a final review of the R script.

Table 15: Quality assurance log

Team	Internal check	Notes	Cross-team check	Notes	External UKRI check	Notes
DART	18/01/2024	Ensure assumptions are checked for t-tests.	07/02/2024	Data and key outputs checked.	13/02/2024	Suggested change to bootstrap by using a permutation inference procedure. This ensures that the correct null hypothesis is created for each sample.
ETF	05/02/2024	No errors found	08/02/2024	Some changes made to charts. Error noted and corrected on line 142.	13/02/2024	No errors found
DART - ITS	13/06/2024	Minor error when running diagnostics.	13/06/2024	Logic behind forecasting method	n/a	n

				discussed and signed off. This was QA'd by the borders economics team, not the ETF.		
--	--	--	--	---	--	--

Appendix 11: Analysis overview for study 2

OBT baseline assessment analysis:

- baseline analysis was conducted on the pre assessment results
- the data were filtered to remove those who had not fully finished the assessment
- for the first 3 questions the responses were converted to numeric and a median table constructed
- the fourth and fifth question were converted to a binary 1/0 response, 1 for “Yes”, 0 for “No”. This was then used to calculate the overall % yes for these questions
- questions 6 to 16 were also converted to binary responses. This was a 1 for the correct answer and 0 for other answers
- % correct by questions and overall were created. These were subsequently split by gender, profession, and grade. Visual representations of this were also created and some significance tests run to compare these subgroups
- due to the small sample size and additional issues with our evaluation, it was decided that subgroup analysis should not take place. This was to avoid generating conclusions on sensitive topics (such as gender differences) that we could not be sufficiently confident about
- comparisons were made between the gender and grade breakdowns of our assessment respondents compared to the wider CS in order to demonstrate that our sample was not necessarily representative of the wider CS

OBT pre/post assessment analysis:

- similar data cleaning was done to the post-assessment data. Incomplete assessment responses were removed and questions were converted to binary variables where appropriate
- one additional step taken was to remove spurious emails that had been introduced when the post-assessment was sent to all on the GPG mailing list in an attempt to increase awareness and therefore sample size
- a pre/post indicator column was added to both datasets before they were joined
- visual representations of overall scores were created split by the pre/post column
- Chi-square tests were conducted across the pre/post split for each question. A T-test was run for overall test scores and the CI interval for this test plotted
- similar comparisons with the overall CS were constructed and represented visually

- questions 1 to 3 had differences between pre and post tested using Wilcoxon Ranked-Sum tests to compare distributions. This is because we are comparing median responses, not mean. Initially, statistical significance was found for some questions
- visual inspection of distributions indicated that there were outliers present, which the Wilcoxon test may have been vulnerable to
- to mitigate this, we re-ran the Wilcoxon tests on bootstrapped distributions using permutation inference to shuffle pre/post labelling and create null distributions. This confirmed statistical significance
- as an additional robustness measure, the responses for questions 1 to 3 were converted to bins (i.e. 0-5). These were then checked using Chi-square tests, with no significant differences found
- differences for questions 4 and 5 were tested using Chi-square tests

OBT participation rates analysis:

- this was done using CSL data to establish whether our respondent groups had different participation rates between pre and post and compared to the entire mailing list
- we constructed a binary variable of 0 for “IN PROGRESS” and 1 for “COMPLETED”
- we then filter pre/post respondent and overall GPG mailing lists so that only those who have completed the training are included. We repeat this step to include those who started but did not finish also

Appendix 12: Multiple hypotheses issue

There are some issues associated with running multiple hypothesis tests for questions 6 to 16 concurrently. Through this we are increasing the likelihood of type 1 errors - detecting a change where there is none. We did find a strong statistical significant ($p\text{-value} < 0.01$) increase in overall mean scores for the aggregate measure of questions 6-16. Additionally, there were measured increases in 9/11 questions (albeit with wide confidence intervals). This adds to our confidence that the key message, which is that there was a very small but noticeable increase in average test scores from 6.26 to 6.95 out of a possible score of 11, remains true and *statistically* significant. It is worth noting that this issue with multiple testing *does* mean that we are less confident about which exact questions are driving the overall change in scores.

Appendix 13: GSR Ethical Checklist for all studies

GSR Ethical Checklist for One Big Thing evaluation

Study 1

<u>GSR Principle 1: Research should have a clear and defined public benefit</u>		
<u>Principle components</u>	<u>Considerations and mitigations</u>	<u>Sensitivity rating</u>
<p>a) Identifying a user need</p> <ul style="list-style-type: none"> - <i>Does the research aim to meet a clearly defined, legitimate and unmet user need?</i> - <i>Have you engaged with relevant stakeholders in order to fully establish the user need?</i> - <i>Is other research already taking place with the same groups, which could be amalgamated to prevent</i> 	<p>This study was used as a case that will inform a more in-depth roll out of evaluation to the wider Civil Service.</p> <ul style="list-style-type: none"> ● Findings of the study will be used to inform further developments of data literacy training for the Civil Service. ● This will feed into wider government initiatives around upskilling Civil Servants in One Big Thing (OBT) training. 	<p><u>Green</u></p>

<p><i>over-researching small populations?</i></p>	<ul style="list-style-type: none"> ● Relevant stakeholders were engaged with and user needs were established to be legitimate. 	
<p>b) Public benefit</p> <ul style="list-style-type: none"> - <i>How will the findings from this research benefit the public?</i> - <i>Are there any risks that public benefits will not be realised?</i> - <i>Could the research disproportionately benefit or disadvantage a particular group?</i> - <i>Is it necessary to conduct this research in order to realise the public benefits?</i> 	<p>The objectives of OBT highlight the numerous potential benefits of the programme.</p> <p>However, an intervention of this size requires significant planning, time and public money. Additionally, the training itself has an opportunity cost for civil servants who could be undertaking other work activities during this time; the equivalent salary cost of all civil servants taking part in 7 hours of training is estimated at over £70 million.</p> <p>It is therefore vital to evaluate whether the training fulfils its aims, especially since this is</p>	<p><u>Green</u></p>

- Does the public benefit outweigh any identified risks?

to become a repeated annual event.
Therefore, this evaluation case study was deemed to be sufficiently beneficial for public interest without any associated risks.

c) Transparency and Dissemination

- *Have you got a clear dissemination strategy in place? i.e. where, when and how you will disseminate findings?*
- *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?*
- *How will you ensure that research findings are brought to the attention of relevant stakeholders?*
- *Will the research process be fully transparent?*

- There is already an engaged stakeholder group for this research. Our responsibility to them is to present a case study evaluation with full transparency about strengths and limitations of the approach, as well as recommendations for future evaluations.
- High-level findings will be disseminated to stakeholders initially through summary documents and presentations.
- Following this, an in-depth report will be delivered to stakeholders and subsequently will be put into the public domain.
- Research participants will remain anonymous in all result-sharing and findings will only be shared in general terms (for example, avoiding reporting on small, identifiable sample sizes).

Green

--	--	--

GSR Principle 2: Research should be based on sound research methods and protect against bias in the interpretation of findings		
Principle components	Considerations and mitigations	Sensitivity rating

a) Proposed methodology

- Is the research design appropriate to the groups being interviewed?

- Is this 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?

- Is the proposed methodology the best and most cost-effective way of answering the research questions?

-Have you considered all the possible potential biases in the data, methods and analysis techniques that will be used in the project?

- Participants were allowed time during their working day to complete the surveys. As the surveys were relatively short, this was deemed to be an appropriate burden.
- As participation was open to the entire Civil Service, this naturally included diverse participants.
- Participants were only asked to provide non-identifiable demographic data that was relevant to analysis (gender and grade for both studies).
- The proposed methodology was developed in line with a small available budget and tight time constraints. Significant consideration was put into determining the best and most cost-effective methodology to answer the research questions.
- A full report is provided alongside results highlighting all potential biases

Green

- Are you using new, emerging, or controversial methodologies or techniques? If so, what steps have been taken to ensure the integrity of the methods and results?

in data, methods, analysis and techniques used.

- We are not using new, emerging or controversial methods.

<p>b) External ethical scrutiny</p> <ul style="list-style-type: none"> - Has your project been subject to independent ethical review? - Does the project fall within the remit of the UK Policy Framework for Health and Social Care Research? (See section 3.13-3.15 in the main guidance for further information and links to decision making tools) - Will contracted partners be required to go through internal ethics committees? 	<ul style="list-style-type: none"> ● A DPIA was completed for this project and accepted. ● This research would not fall within the remit for UK Policy Framework for Health and Social Care Research. ● We did not use contracted partners for this research. 	<p style="text-align: center;">Green</p>
---	--	---

GSR Principle 3: Research should adhere to data protection regulations and the secure handling of personal data		
Principle components	Considerations and mitigations	Sensitivity rating

<p><u>a) Data Protection</u></p> <ul style="list-style-type: none"> - <i>What procedures are in place to ensure adherence to the GDPR, Data Protection Act (2018) and other government data security requirements?</i> - <i>What is your legal basis for processing of personal data?</i> - <i>How will you inform and assure participants that you will treat their data in accordance with the relevant data protection legislation (e.g. privacy notice)?</i> - <i>Do you need to complete a Data Protection Impact Assessment?</i> 	<p>A privacy notice and DPIA were completed for this study ensuring that it meets GDPR and other relevant guidelines.</p>	<p><u>Green</u></p>
---	---	----------------------------

<p><u>b) Research findings</u></p> <p>- <i>How can you ensure that the data collected during the research is not going to be used for any other than its originally defined purpose?</i></p> <p>- <i>What checks are in place to ensure that no one can be identified in reporting? (for both quantitative and qualitative work)</i></p>	<ul style="list-style-type: none"> • Data has only been presented alongside relevant caveats around its limitations. This is an agreement made between researchers and relevant stakeholders. • No personally identifying data was included in reporting. • Data was presented in line with Office of National Statistics guidelines which state that data will not be presented for sample sizes of less than 10. 	<p><u>Green</u></p>
---	---	----------------------------

Study 2

<u>GSR Principle 1: Research should have a clear and defined public benefit</u>		
<u>Principle components</u>	<u>Considerations and mitigations</u>	<u>Sensitivity rating</u>

<p>a) Identifying a user need</p> <ul style="list-style-type: none"> - <i>Does the research aim to meet a clearly defined, legitimate and unmet user need?</i> - <i>Have you engaged with relevant stakeholders in order to fully establish the user need?</i> - <i>Is other research already taking place with the same groups, which could be amalgamated to prevent over-researching small populations?</i> 	<p>This was used as a case study that will inform a more in-depth roll out of evaluation for OBT 2024.</p> <ul style="list-style-type: none"> ● Findings of the study will be used to inform further developments of OBT evaluation for the Civil Service. ● This will feed into wider government initiatives around upskilling civil servants in One Big Thing (OBT) training. ● Relevant stakeholders were engaged with and user needs were established to be legitimate. ● A study undertaken by the Evaluation Task Force is running at the same time which will be used to form a wider case study evaluation. 	<p><u>Green</u></p>
--	---	----------------------------

b) Public benefit

- *How will the findings from this research benefit the public?*
- *Are there any risks that public benefits will not be realised?*
- *Could the research disproportionately benefit or disadvantage a particular group?*
- *Is it necessary to conduct this research in order to realise the public benefits?*
- *Does the public benefit outweigh any identified risks?*

The objectives of OBT highlight the numerous potential benefits of the programme.

However, an intervention of this size requires significant planning, time and public money. Additionally, the training itself has an opportunity cost for civil servants who could be undertaking other work activities during this time; the equivalent salary cost of all civil servants taking part in 7 hours of training is estimated at over £70 million.

It is therefore vital to evaluate whether the training fulfils its aims, especially since this is to become a repeated annual event.

Therefore, this evaluation case study was deemed to be sufficiently beneficial for public interest without any associated risks.

Green

c) Transparency and Dissemination

- *Have you got a clear dissemination strategy in place? i.e. where, when and how you will disseminate findings?*
- *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?*
- *How will you ensure that research findings are brought to the attention of relevant stakeholders?*
- *Will the research process be fully transparent?*

- There is already an engaged stakeholder group for this research. Our responsibility to them is to present a case study evaluation with full transparency about the strengths and limitations of the approach, as well as recommendations for future evaluations.
- High-level findings will be disseminated to stakeholders initially through summary documents and presentations.
- Following this, an in-depth report will be delivered to stakeholders and subsequently will be put into the public domain.
- Research participants will remain anonymous in all result-sharing and findings will only be shared in general terms (for example, avoiding reporting on small, identifiable sample sizes).

Green

--	--	--

GSR Principle 2: Research should be based on sound research methods and protect against bias in the interpretation of findings		
Principle components	Considerations and mitigations	Sensitivity rating

a) Proposed methodology

- Is the research design appropriate to the groups being interviewed?

- Is this 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?

- Is the proposed methodology the best and most cost-effective way of answering the research questions?

-Have you considered all the possible potential biases in the data, methods and analysis techniques that will be used in the project?

- Participants were allowed time during their working day to complete the assessment. As the assessment was relatively short, this was deemed to be an appropriate burden.
- Surveys were open to all GPG civil servants who took part in OBT. The pre-survey was attached to OBT training and the post-survey was open to all those who had completed the training. This naturally included a cross-section of demographics within the business unit.
- Participants were only asked to provide non-identifiable demographic data that was relevant to analysis (gender and grade for studies 1 and 2, which collected primary data).
- The proposed methodology was developed in line with a small available budget and tight time constraints. Significant consideration was put into

Green

- Are you using new, emerging, or controversial methodologies or techniques? If so, what steps have been taken to ensure the integrity of the methods and results?

determining the best and most cost-effective methodology to answer the research questions.

- A full report is provided alongside results highlighting all potential biases in data, methods, analysis and techniques used.
- We are not using new, emerging or controversial methods.

<p>b) External ethical scrutiny</p> <ul style="list-style-type: none"> - <i>Has your project been subject to independent ethical review?</i> - <i>Does the project fall will in the remit of the UK Policy Framework for Health and Social Care Research? (See section 3.13-3.15 in the main guidance for further information and links to decision making tools)</i> - <i>Will contracted partners be required to go through internal ethics committees?</i> 	<ul style="list-style-type: none"> ● A DPIA and privacy notice were completed for this project. ● This research would not fall within the remit for UK Policy Framework for Health and Social Care Research. ● We did not use contracted partners for this research. 	<p style="text-align: center;">Green</p>
---	---	---

GSR Principle 3: Research should adhere to data protection regulations and the secure handling of personal data		
Principle components	Considerations and mitigations	Sensitivity rating

<p><u>a) Data Protection</u></p> <ul style="list-style-type: none"> - <i>What procedures are in place to ensure adherence to the GDPR, Data Protection Act (2018) and other government data security requirements?</i> - <i>What is your legal basis for processing of personal data?</i> - <i>How will you inform and assure participants that you will treat their data in accordance with the relevant data protection legislation (e.g. privacy notice)?</i> - <i>Do you need to complete a Data Protection Impact Assessment?</i> 	<p>A privacy notice has been completed for this study ensuring that it meets GDPR and other relevant guidelines.</p>	<p><u>Green</u></p>
---	--	----------------------------

<p><u>b) Research findings</u></p> <p>- <i>How can you ensure that the data collected during the research is not going to be used for any other than its originally defined purpose?</i></p> <p>- <i>What checks are in place to ensure that no one can be identified in reporting? (for both quantitative and qualitative work)</i></p>	<ul style="list-style-type: none"> • Data has only been presented alongside relevant caveats around its limitations. This is an agreement made between researchers and relevant stakeholders. • No personally identifying data was included in reporting. • Data was presented in line with Office of National Statistics guidelines which state that data will not be presented for sample sizes of less than 10. 	<p><u>Green</u></p>
---	---	----------------------------

GSR Principle 4: Participation in research should be based on specific and informed consent		
Principle components	Considerations and mitigations	Sensitivity rating

<p><u>a) Consent to take part in primary research</u></p> <p><i>- 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? Will you ensure that participants have given fully informed consent before taking part in the research?</i></p> <p><i>- If you intend to follow up participants with further research, has this been made clear and consent given?</i></p>	<p>Before completing the survey, participants were provided with a data privacy impact assessment (an explanation of the purpose of the data collection and how their data would be processed, stored and used) so they could give informed consent. Participants could withdraw at any time and participation was anonymous.</p>	<p><u>Green</u></p>
---	---	----------------------------

<p><u>b) Consent via gatekeepers or proxy</u></p> <ul style="list-style-type: none"> - <i>Is this required? If so, what processes need to be in place?</i> - <i>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?</i> 	<p>This is not required.</p>	<p><u>Green</u></p>
<p><u>c) Children and young people (aged 16 and under)</u></p> <ul style="list-style-type: none"> - <i>What processes are in place to ensure consent from a parent or legal guardian has been sought for children under the age of 16 and how has this been done?</i> - <i>How can you ensure that the children are also adequately informed about the research?</i> - <i>What processes are in place to ensure, where required, an adult accompanies children and</i> 	<p>This is not required.</p>	<p><u>Green</u></p>

<p><i>young people during an interview? Who is best to accompany the child(ren)?</i></p>		
<p><u>d) Vulnerable adults</u> <i>- Are you interviewing participants who may lack the mental capacity to provide informed consent for themselves? If so, the successful contractor may be required to obtain clearance from an NHS Research Ethics Committee.</i></p>	<p>This is not required.</p>	<p><u>Green</u></p>

<p>- How can you ensure that participants are adequately informed about the work?</p>		
<p>e) Access protocols</p> <p>- Are there any particular access protocols for certain groups, does this apply to your respondent group?</p> <p>Access protocols could apply to: Courts, Police, Prisons, Schools</p>	<p>This is not required.</p>	<p><u>Green</u></p>
<p>f) Secondary Research</p> <p>- Does the consent cover all potential future uses of the data?</p> <p>- If your legal basis for processing data is not consent, have you still considered whether individuals have been (or should be) given</p>	<p>This is not required.</p>	<p><u>Green</u></p>

<p><i>the choice of their data being included in this research?</i></p>		
<p>g) Incentives?</p> <p><i>- Is the use of incentives necessary? What evidence do you have that the use of incentives will significantly improve the research?</i></p> <p><i>- Is your use of incentives in keeping with the GSR ethical principles? (See section 2.33-2.35 in the main guidance for further information)</i></p>	<p>This is not required.</p>	<p><u>Green</u></p>

GSR Principle 5: Research should enable participation of the groups it seeks to represent		
Principle components	Considerations and mitigations	Sensitivity rating
<p><u>a) Identifying and reducing the barriers to participation</u></p> <p>- <i>What steps have you taken to identify potential barriers to participation?</i></p> <p>- <i>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)</i></p> <p>- <i>Do you need interviewer assistance such as offering help with completion, or a translator?</i></p>	<ul style="list-style-type: none"> • The surveys were simple to access and could be done so from anywhere with internet connection. • As participant samples were taken from a wider sample who conduct their day-to-day work in English, the fact that the assessment was presented in English was not deemed to be an issue. 	<p><u>Green</u></p>

<p><u>b) Ensuring that hard to reach groups are included</u></p> <p><i>- Is the research and sample design appropriate?</i></p> <p><i>- Might the data collection method exclude some groups of people?</i></p> <p><i>- Do you need to consult with others (e.g. support groups, charities and other relevant stakeholders) so that barriers to participation for certain groups are fully identified and reduced?</i></p>	<ul style="list-style-type: none"> • The sample design was appropriate for the aims of the study. Researchers ensured that reporting of data did not state that it was relevant to any groups who weren't represented in the sample. • Consultation with external bodies was not required for this research. 	<p><u>Green</u></p>
---	--	----------------------------

GSR Principle 6: Research should be conducted in a manner that minimises personal and social harm		
Principle components	Considerations and mitigations	Sensitivity rating

<p><u>a) Research participants</u></p> <ul style="list-style-type: none"> - <i>Do any of the research questions cover stressful or culturally sensitive subjects? If so, how will stress and sensitivities be minimised?</i> - <i>How can interview length be kept to the minimum?</i> - <i>Do you need to ensure that there is post-interview support?</i> - <i>How will you offer support to those that are approached but decide not to participate in the research?</i> 	<p>Survey questions were assessed and did not risk a negative impact on participant wellbeing.</p> <p>They did not require the disclosure of sensitive information.</p>	<p><u>Green</u></p>
--	---	----------------------------

b) Interviewers/ researchers

- *What procedures are in place to ensure interviewers are properly trained (for example in methods, relevant legislation such as the Equality Act)?*
- *Do all interviewers /researchers have appropriate security clearance (e.g. criminal record checks or disclosure Scotland if interviewing/ working with children)?*
- *What procedures are in place for handling disclosures of abuse, self-harm or suicidal ideation?*
- *What procedures are in place to ensure the safety of the interviewer/ researcher?*
- *Has consideration been given to exposure of researchers and analysts to sensitive topics?*

This study did not include interviews.

Green

(e.g. potential for vicarious trauma)

<p><u>c) Wider Social Groups</u></p> <p><i>- How will you mitigate any potential for harm to those who have not taken part in the research? For example, research focussing on specific groups has the potential to impact the wider social group.</i></p> <p><i>- Have you considered or sought the public's views on the research?</i></p>	<p>As this research was a case study of how to employ future OBT evaluations, rather than a study aiming to attribute overall success or failure of the training to the wider Civil Service, it was not anticipated that there would be potential harm for those who had not taken part in the research.</p>	<p><u>Green</u></p>
---	--	----------------------------

Relevant legislation

<p><u>Will your research comply with all relevant legislation?</u></p> <p><u>For example:</u></p> <ul style="list-style-type: none"> · Anti-Terrorism, Crime and Security Act (2001) · Crime and Disorder Act (1998) · Data Protection Act (2018) · Freedom of Information Act (2000) · General Data Protection Regulation (2016) · Health and Social Care Act (2012) · Human Rights Act (1998) · Mental Capacity Act (2005) · Equality Act (2010) - Public Sector Equality Duty <p><u>Do you need to ensure compliance with any additional legislation, policy, code of practice or guidance?</u></p>	<p>Yes</p>
---	------------

<p>Summary</p>	<p>Overall sensitivity rating</p>
-----------------------	--

What are the key sensitivities?

We collected personal data from participants including grade, department and profession. In addition, participants could volunteer information through free-text comments, so it was possible that personal data was shared here.

How are you addressing them?

Data was anonymised before sharing outside of the analysis team. For example, this meant removing participants from the data where it was possible to identify them. In practice, this meant removing profession data where participants had less than 10 responses. Participants could withdraw consent at any time.

How often will you re-visit this research ethics assessment?

This ethics assessment will be re-visited whenever findings are shared to new groups and in future OBT evaluations.

Green

