



The Civil Service Future Leaders and Senior Leaders Accelerated Development Schemes - Evaluation Update

This document provides a mid-point update on our evaluation of the Future Leaders and Senior Leaders schemes, and summarises our findings from an analysis of participants' in the 2023/4 cohorts motivations. **It provides a snapshot of our progress as of August 2024, and is not a completed evaluation report.**

This report provides an update on our progress in evaluating two accelerated leadership development schemes run by the Leadership College for Government (LCG): the Future Leaders Scheme (FLS), and Senior Leaders Scheme (SLS). It outlines our interim impact evaluation results at a mid point of the evaluation, and research on motivation undertaken on the two programmes.

Our evaluations of FLS and SLS are interested in whether the programmes result in changes for the participants in the leadership attributes and outcomes the programmes are designed to develop. We are also investigating the different motivations people have for applying for each programme, and whether certain motivations mean that people are more likely to be accepted onto either programme.¹ We are using specialist quantitative evaluation methods to carry out the research. Our evaluations will cover two cohorts of each programme. So far we have collected and analysed data on one cohort for each programme. Therefore, this is an interim update on the evaluation and does not provide any final findings on either programme.

The FLS and SLS evaluations were developed in 2022, at a time when the decision had already been made to reform both programmes. This impact evaluation therefore forms a baseline against which to compare the impact of the later, reformed programmes, and to learn lessons to inform reform. The first year of our impact evaluation therefore focuses on the impact of the 2023/24 SLS and FLS programmes (those that started delivery in early 2023 and finished in early 2024) only, and not the current programme or the future, fully-reformed programme.

¹ In future we hope to examine whether participants' motivations are predictive of their programme outcomes. This is dependent on achieving a large enough sample size, and it was therefore not possible to carry out in year one of the evaluation.





This update report provides information on the programmes, and an outline of the approach, results, and next steps for each set of analysis.

The FLS and SLS

The Future Leaders Scheme (FLS), and Senior Leaders Scheme (SLS) are two of the Accelerated Development Schemes (ADS) delivered by the Leadership College for Government (LCG). The ADS also includes the Directors Leadership Scheme (DLP), and two programmes which provide additional opportunities for those who gain a place on the FLS and identify as coming from an ethnic minority background (META) or as having a disability (DELTA).

The ADS aims to develop high-potential individuals to build a robust and diverse pipeline to the most senior and critical Civil Service roles, with each of the three main programmes targeting different grades, having different number of participants, and being developed at different times:

- FLS is aimed at Grades 6 and Grade 7 and equivalent (the two grades immediately below the Senior Civil Service), has approximately 400 participants per cohort, and began running in 2019.
- SLS is aimed at Deputy Directors (the entry grade into the Senior Civil Service), has approximately 100 participants per cohort, and began running in 2012.
- DLP is aimed at Directors (the middle grade in the Senior Civil Service), has approximately 30-35 participants per cohort, and began running in 2023, replacing the previous High Potential Development Scheme (HPDS).

Each programme runs one 12 month cohort per year. Each programme primarily focuses on developing system leadership, networking, and collaboration. The programmes generally include a mixture of residentials and online sessions, although the exact content and pedagogical approaches used differ depending on the programme.

The impact evaluations

The Leadership College for Government decided in late 2022 to investigate reforming ADS programmes, primarily the FLS and SLS, to ensure these programmes were maximising positive outcomes in **systems leadership, networking, and collaboration**, for those taking part. As part of this reform process, it was decided to undertake an impact evaluation of FLS and SLS to assess the extent that each programme delivered these intended outcomes. As a decision to reform these programmes had already been made, the value of this impact evaluation is in forming a baseline against which to compare





the impact of the later, reformed programmes and to pinpoint any areas in which the current, pre-reform programmes are particularly successful or unsuccessful.

Our approach to impact evaluation of ADS

As FLS, SLS, and DLP all have similar outcomes, a similar data collection approach was adopted for all three programmes. This uses pre and post questionnaires for applicants using validated and previously tested questions to assess self-reported changes in systems leadership, networking, and collaboration. This approach was adopted across all three ADS programmes, starting with the cohorts recruited in 2022 and beginning the programme in early 2023.

We have used a slightly different evaluation design for DLP, due to the much smaller cohort size for the programme, and due to our intent to carry out a more in depth process evaluation of this programme in its first, pilot year. We have produced a separate evaluation update report for DLP, which can be found here [\[insert hyperlink\]](#). Where relevant, and as we used the same outcome measures, we also talk in this update report about some of the DLP interim findings and lessons learned.

For FLS and SLS, we wanted a rigorous evaluation design which helped us to understand whether the programme actually caused any changes in outcome we observed. Questionnaires were provided to all applicants to the programmes to complete. This allowed us to construct a ‘treatment’ group of those who are accepted onto the programme, and a control group of those who were not accepted. Results for treatment and control groups could then be compared using a specialist quasi-experimental evaluation design called difference-in-difference analysis. The difference-in-difference analysis approach compares the trends we would have expected for two comparable groups in our key outcomes (how far their systems leadership, networking and collaboration improved), to what was actually seen. We can then estimate the effect of taking part in the programme by comparing the outcomes in the treatment group to those in the control group. We can also assess the results for statistical significance to see whether any change is likely or not to be due to chance variation.

It is important to have a sufficiently large sample when carrying out difference-in-difference analysis. The achieved sample size for FLS was sufficient for a full difference-in-difference analysis in the first year.² The achieved sample size for SLS was just sufficient for an indicative

² The FLS sample size for the treatment group was 248 for the pre-questionnaire and 112 for the post-questionnaire, and for the control group was 543 in the pre and 164 in the post.





difference-in-difference analysis,³ which provides a very approximate view of whether the findings on impact appeared to be in line with those for FLS.

Impact measures

Our surveys for each programme included a wide variety of statements to measure potential changes in different aspects of systems leadership, networking, and collaboration. To make reporting easier, and better understand the relationship between these many statements, we used a method called factor analysis to combine these multiple statements into a smaller number of grouped outcomes (or factors). This analysis identified 14 factors, each of which is then grouped into one of six overarching themes:⁴

- Networking: Using a network to achieve positive results; putting time and effort into building a network.
- Systems leadership: Seeing the bigger picture; interpersonal and communication skills; discussing and improving service provision.
- General leadership: Building credibility and taking responsibility; reflecting upon and trying new leadership practices; delivering high quality and cost-effective projects; level of satisfaction with current work and level of interest in learning.
- Collective leadership: Developing self and others; focusing self and others towards an objective; setting specific targets and measuring progress.
- Leadership and Impact: Considering the impact of work on the public.
- Collaboration: Collaboration.

Measuring change using a before and after measurement

Our difference-in-difference impact analysis assessed change over the course of the programme in two ways. We first assessed whether there was a statistically significant change for each of the factors.⁵ Where there was a statistically significant change, we then assessed

³ The SLS sample size for the treatment group was 67 for the pre-questionnaire and 54 for the post-questionnaire, and for the control group was 24 in the pre and 17 in the post.

⁴ These factors represent the patterns in this year's data. When we include future data, the factors may or may not be the same.

⁵ A simple definition of statistical significance is that a result is statistically significant if the results in the data are unlikely to be due to random chance. Our analysis here uses a p-value, and takes the generally standard approach of taking a result as statistically significant if the p-value is less than 0.05. A result being statistically significant does not mean that it is necessarily practically meaningful.





the amount of change by measuring the effect size,⁶ and placing these into one of four categories:⁷

- Negligible (an effect size or less than 0.02);
- Small (an effect size of at least 0.02);
- Medium (an effect size of at least 0.15);
- Large (an effect size of at least 0.35).

This approach therefore allows us to say for each of the factors whether there is evidence there was a change (statistical significance) and, if so, whether we can say it was a small, medium, or large change (effect size).

Limitations

Our difference-in-difference analysis approach gives us considerably more rigorous results than have been previously provided for SLS and FLS. There are, however, certain limitations to our approach.

The main limitation of our approach is that, as we do not have administrative data on organisational results such as retention or staff absence, we rely on self-reported measures relating to system leadership, networking, and collaboration. Therefore, we don't know whether participants' performance in their roles changed using an independent measure (such as a performance rating), nor whether this influenced performance and delivery in the short or long term. We are also only assessing the impact on leaders, and not on their teams. Nonetheless, the self-reported questions we used have been tested in other public sector contexts for their validity, so we can have confidence in these measures.

Difference-in-difference analysis relies on the assumption that the two groups would have had the same change in outcomes if there had not been an intervention - the parallel trends assumption. Evidence to substantiate this assumption can be obtained by seeing whether the treatment and control groups have similar changes in outcome measures over multiple time points before the programme or intervention starts. We only have data for one time point before SLS or FLS started (our pre-survey data), meaning we cannot be certain that the parallel

⁶ The effect size is a standardised measure of the size of any change, which allows you to easily compare the size of change across different variables which may use different scales (e.g. height vs weight). Our analysis here uses Partial Cohen's f^2 , a common way to measure the change in a multiple regression model where both independent and dependent variables are continuous.

⁷ The key thresholds for these categories (0.02, 0.15, and 0.35) and the categorisations of small, medium, and large are taken from Cohen. J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.).





trends assumption holds prior to the intervention. To address this limitation we examined whether there were different results between treatment and control groups depending on the exact point when people answered the FLS survey (between April and May 2023). This showed there were no significant differences in trends across the two groups over this limited period.

As participants were not randomised into the treatment or control group we cannot be sure that they are the same type of people, and different post-questionnaire data approaches for each programme may have affected the profile of respondents.⁸ Nonetheless, comparing successful and unsuccessful candidates, rather than comparing people who took part in the scheme with non-applicants, is much more likely to create a group which is more comparable on hidden, relevant features like motivation and engagement. As noted, the SLS results are indicative and likely to be affected by the small sample size, particularly for non-participants.

While these limitations are noted, the overall impact analysis approach is considerably more rigorous than the vast majority of training evaluations carried out in Government and more widely to date. The Next Steps section later in this document outlines how we will build on this initial analysis to provide more rigorous information in future.

Main Impact Findings

Our difference-in-difference impact analysis found that FLS had statistically significant and positive changes for 7 of the 14 factors, but for 5 of these 7 significant factors the magnitude of the change was negligible. The remaining factors, relating to seeing the bigger picture (in the Systems Leadership theme) and reflecting upon and trying new leadership practices (in the General Leadership theme) both had a small positive effect (0.02).

A slightly different pattern was seen in the indicative analysis for SLS, with statistically significant and positive changes being seen for 6 of the 14 factors, with all 6 having a small effect size, namely:

- Networking: Using a network to achieve positive results (0.06); Putting time and effort into building a network (0.09)
- Systems leadership: Seeing the bigger picture (0.14)
- General leadership: Reflecting upon and trying new leadership practices (0.07); Delivering high quality and cost-effective projects (0.09)
- Leadership and impact: Considering the impact of work on the public (0.06)

⁸ For FLS, the post-questionnaire was primarily completed in-person at the final event, and via e-mail. For SLS, the post-questionnaire was entirely completed via email.





There were no statistically significant changes in this indicative analysis for the other 8 factors.⁹ It is important to emphasise that due to the small sample size, these results may overestimate the extent of change and are likely to look different when we repeat the tests with a larger sample size next year.

G7 and G6 comparison analysis

We were interested in whether the outcomes and levels of progress made by participants differed according to their grade. We undertook exploratory difference-in-difference impact analysis on FLS participants to assess whether there was a different impact on each factor for G7s than G6s. Splitting the FLS sample into these two groups reduced the sample size, making it difficult to identify significant changes. This analysis showed that G7s tended to have greater increases than G6s, but these were not significant (possibly due to the small sample size), with the only significantly greater increase for G7s than G6s being for the Leadership and Impact factor. These findings are therefore inconclusive, and we will continue to look at this question as the evaluation continues.

Research into motivation

Examining motivation is an important area of analysis. Evidence from meta-analyses shows that motivations can be important in explaining individual variance in training outcomes ([Colquitt et al, 2000](#)).

We included an amended (reduced) version of the Adult Education Motivation Scale (AEMS) in our pre-course questionnaires with the 2023/24 cohorts.

Descriptive analysis was conducted to examine three separate research questions:

- What motivates participants to apply for FLS and SLS?
- Are FLS and SLS applicants with certain types of motivation more likely to be accepted onto the programme?
- Are participants particularly motivated by being accepted on to a programme as opposed to participating in a programme?

⁹ These were: *Systems Leadership*: Discussing and Improving Service Provision, Interpersonal and Communication Skills; *General Leadership*: Building Credibility and Taking Responsibility, Dissatisfaction with Current Work and Lack of Interest in Learning. *Collective Leadership*: Developing Self and Others, Focusing Self and Others Towards an Objective, Setting Specific Targets and Measuring Progress. *Collaboration*: Collaboration





Limitations

Participants were directed to the motivation questionnaire immediately after completing their application. It is possible that having just completed the application form may have affected responses. It is also possible that respondents self-censored and did not include responses which they believed may be less socially acceptable, such as amotivation.

Analysis on types of motivation

We analysed the extent that participants for FLS and SLS fitted into each of six different types of motivation, as identified in responses to our amended version of the Adult Education Motivation Scale (AEMS):

- Amotivation: participants don't see a relation between taking part in a programme and relevant outcomes.
- External regulation: participants take part to gain a reward, such as promotion or better pay.
- Introjected regulation: participants are motivated by internalised pressures, such as gaining self-esteem.
- Identified regulation: participants take part voluntarily as they feel it is important.
- Integrated regulation: participants take part as they feel the training fits with their value system and needs.
- Intrinsic regulation: participants take part because of the enjoyment and satisfaction they get.

What motivates participants to apply for FLS and SLS?

There were clear patterns in why people were motivated to apply for SLS or FLS. Participants were primarily either intrinsically motivated (to learn new things in general, or new things that interest them) or through intrinsic regulation, because they felt it was important (to do better at work, or in their career). Participants tended not to be externally motivated (to get a better job, or salary), with almost no participants being amotivated (not knowing why they are wanting to do the training, or thinking it is a waste of time).

This presents a positive pattern for both programmes, given the link identified by Van Den Broeck ([2021](#)) between intrinsic motivation and employee well-being, attitudes and behaviour, and between identified regulation and performance and organisational citizenship behaviour.

There were no notable differences in motivation between those applying for FLS and those applying for SLS, with the sole exception of slightly higher





scores for those applying for FLS on statements relating to an introjected motivation, namely being motivated by internalised pressures - wanting to be on the course so they can show themselves what they can accomplish or that they can complete the course. Van Den Broeck (2021) suggests introjected motivation is linked to both distress, commitment, and performance, indicating “that introjected people may perform well by pressuring themselves or striving to feel better about themselves, but with some well-being price to pay”.

Are applicants with certain types of motivation more likely to be accepted?

Although none of the ADS programmes use motivation as a criterion during the assessment process, it is possible that the motivation of participants could affect whether or not they are accepted onto a programme.

We used a technique called logistic regression analysis to assess whether participants with certain types of motivation were more or less likely to be accepted onto the programme. As with the impact analysis, above, the larger sample size for FLS means we can be more confident in our findings, whereas the small SLS sample sizes means these results are only exploratory.

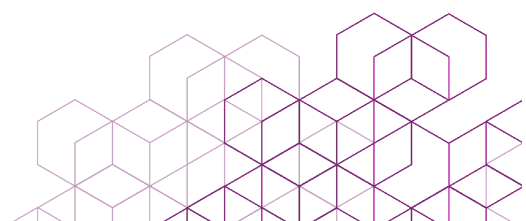
Our analysis suggests that the type of motivation that participants had for applying did not affect whether they were successful or not in being accepted onto either programme. There was no statistically significant difference in types of motivation between those who were accepted onto FLS or SLS and those who were not.

Are participants motivated by being accepted on to a programme as opposed to participating?

The revised AEMS scale used for this research also included two statements designed to assess the hypothesis that participants may be particularly motivated by being accepted on to the programme. This is linked to a hypothesis that some people may apply for programmes primarily for the ‘badge’ of being accepted on to FLS or SLS, rather than primarily for outcomes relating to the course content.

This hypothesis was tested by including two regular statements relating to different sub-categories of extrinsic motivation and relating to the training itself, with two variants relating to simply being accepted on to the course.

- Extrinsic motivation: Identified





- A; Because this training will help me in the rest of my career
- B: Because just being able to show I have been accepted on this training will help me in the rest of my career
- Extrinsic motivation: Introjected
- C: By successfully completing this training, I will show myself what I can accomplish
- D: Because just by being accepted on this training I can show myself what I can accomplish

Analysis showed that applicants were considerably more likely to self-report being motivated by the training itself, than by simply being accepted on to the course, with this particularly the case for the statements relating to identified motivation. Applicants were more likely to be motivated to be accepted onto the course if being accepted showed them what they could achieve (statement D) than if they felt it would help them in the rest of their career (statement B). The relatively low endorsement for this latter statement, suggests that only a minority of participants are motivated by the ‘badge’ of being accepted on to the programme, with these participants also having other motivations.

Next Steps

These are not conclusive findings, as we have not yet completed our evaluation. It is too early to be able to be fully clear about the extent of any change (particularly for SLS), and to assess whether and why we may have found limited change for several intended outcomes of these programmes, nor why we may have found positive change for some of them. We may find that these results look different next year with a larger sample and, potentially, with adapted analytical approaches. Therefore, the most important thing for us to do next is to complete our evaluation, and analyse the final data. These interim results, however, have already been provided to staff working on ADS Reform, who are using these as one source of evidence to help develop future iterations of the programme. Several changes have already been implemented for the FLS and SLS programmes due to start in 2025, and we will explain these fully in our final evaluation report, as they will be relevant in interpreting the results.

We are using the same outcome measures in questionnaires for the current FLS, SLS, and DLP cohorts, the second year of our evaluation. This will provide additional data, potentially allowing more rigorous impact analysis with larger sample sizes (particularly for SLS), and more rigorous research into motivation. Depending on future sample sizes, this may include analysis to examine the clusters of motivations held by programme applicants and whether different





clusters lead to different outcomes. We will publish an evaluation report at the end of the current evaluation.

We will be designing a new evaluation for the reformed ADS as these programmes are designed and launched. We will aim to keep the outcomes we measure as similar as possible to those in the pre-reform programmes, given the intention of the current evaluation to provide a baseline for the reformed programme. We will, however, make adaptations that ensure the evaluation questions also remain relevant to current and future priorities. We will continue to apply lessons learned from wider evidence reviewing and other Government Skills evaluations of learning and development into our evaluation design for the ADS. Our evaluation plans and findings will be uploaded to the Evaluation Registry, and will be published.

In parallel, we are reviewing wider evaluation evidence on leadership development programmes, and collaborating with external academic experts, to help us develop explanations for why some outcomes may, or may not be, being developed through these programmes, and to inform the reform process.

This note was prepared by Government Skills. It represents a snapshot of the evidence as of August 2024. To learn more about this work, please contact:

gscu.comms@cabinetoffice.gov.uk

This evaluation is registered on the [Government Evaluation Registry](#), and can be located by searching for “Senior Leaders Scheme”, and “Future Leaders Scheme”.

