Effective Professional Development Design in a Civil Service Context – Executive summary

alma economics
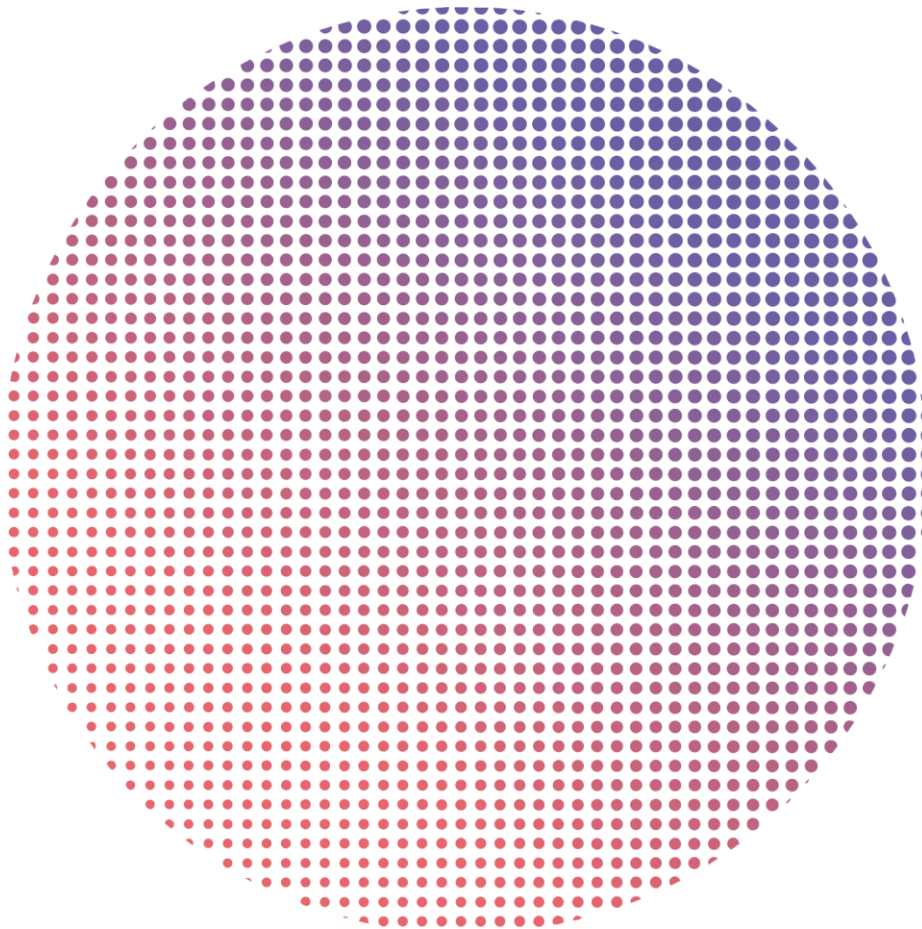
Executive Summary

Introduction

To ensure that investments made by the UK civil service in professional development are as effective as possible, extensive and high-quality evidence synthesis work is necessary to support future decision-making. Previous systematic reviews in the professional development area define professional development as facilitated learning opportunities for qualified professionals that aim to enhance their knowledge and skills in ways relevant to their application in practice (Filges et al. 2019; Sims et al. 2021).

The existing evidence base primarily consists of research on the design features and mechanisms of effective professional development elsewhere in the public sector, as evident in comprehensive guides for the teaching and medical professions (e.g., Education Endowment Foundation 2021; AMEE 2024). Mechanisms are the "active ingredients" of an intervention—key components that drive effectiveness and would reduce its impact if removed (Sims et al. 2021).

No equivalent evidence reviews have been conducted regarding the civil service. Existing research is fragmented, of varying quality, and often drawn from international contexts (Elliott 2023). Chapman (1987) has also highlighted the lack of a unified framework of concepts within public administration teaching, as well as a wide gap between academic research and practitioners.

To address these evidence gaps, Government Skills has launched an integrated programme comprising three projects to explore the relationship between professional development, workforce skills, and productivity. As part of this initiative, this systematic review examines the key characteristics, design features, and mechanisms of professional development that enhance productivity-related skills within the civil service context.

This document is one of three outputs from the project and serves as an executive summary of the main findings in a clear, non-technical format. The second output is a detailed technical report outlining the full evidence base and methodology. The third is a practitioners' guide, presenting a tailored taxonomy of mechanisms of change to support the effective design, delivery, and evaluation of professional development in the civil service.

Objectives

This review aims to synthesise existing literature to identify the characteristics of effective professional development in the civil service and adjacent contexts. It seeks to:

(i) explore the overall effectiveness of these interventions in driving improvements in knowledge, skills, networks, work performance, and productivity;

(ii) examine the mechanisms, design features, and forms (clusters of mechanisms) of professional development associated with the greatest impact on driving these improvements; and

(iii) identify factors that support the successful implementation of professional development programmes in the civil service and adjacent contexts.

To achieve this, four research questions are addressed:

1) What are the characteristics of the studies and interventions in the experimental impact evaluation literature on professional development design in the civil service context and adjacent contexts?

2) Overall, how effective are professional development interventions in the civil service context at improving knowledge, skills, networks, work performance, and productivity?

   a) Does this vary based on study characteristics (features of the evaluation not specific to the intervention itself)?

   b) Does the effectiveness of the interventions vary based on the types of workers and the target outcomes?

3) Which design features, mechanisms, and forms of professional development are associated with the greatest impact?

   a) Which design features (e.g., online versus face-to-face; longer duration versus shorter duration) are associated in the literature with the greatest impact on skills, knowledge, networks, work performance, and productivity?

   b) Which mechanisms and forms (clusters of mechanisms) do we observe in the literature?

   c) Which mechanisms or forms (combinations of mechanisms) are associated in the literature with the greatest impact on skills, knowledge, networks, work performance, and productivity?

4) What supports the successful implementation of professional development interventions targeted at driving improvements in knowledge, skills, networks, work performance, and productivity in the civil service context and adjacent contexts?

What did we find?

This review found that professional development interventions tend to have a positive impact, particularly on skills, across civil service and adjacent contexts. However, the evidence base is limited in both size and consistency, with significant variation and gaps in reporting and design, making it challenging to draw firm conclusions about what works best. Strengthening the quality and quantity of future research will be essential to inform the design of effective professional development.

We present a more detailed account of our findings by research question below.

Research Question 1: What are the characteristics of the studies and interventions in the experimental impact evaluation literature on professional development design in the civil service context and adjacent contexts?

We found a total of 27 studies, of which 20 were independent from each other. The studies varied widely in multiple ways.

- **Setting and design:** The majority of papers (19) used randomised controlled trials (RCTs), while eight employed quasi-experimental designs. Designs varied in complexity, with some studies including multiple treatment groups. The

geographic focus was largely on North America (USA, Canada), with 14 papers set in the region.

- **Participants:** A wide range of roles related to the civil service professions were studied. Operational delivery professions (e.g., probation officers, call centre workers) were overrepresented, likely due to the availability of performance metrics that can be readily quantified. Other studies examined tax professionals, project managers, and policymakers. Participant demographics varied; women made up an average of 66% of samples in cases where gender was reported. The age and tenure of study participants varied widely.

- **Interventions:** The systematic review covers a wide range of professional development interventions across different sectors, from job-specific training for probation officers and call centre employees to broader leadership and career development programs. The interventions vary in format and duration, including online courses, structured workshops, and long-term training programs, reflecting the varied nature of professional development efforts.

- **Outcomes:** Most studies focused on skill acquisition, with fewer examining work performance, productivity, or knowledge outcomes, and none focused on networks. Operational delivery roles were more likely to have measurable productivity or performance indicators. Assessment methods also varied widely, including self-reported surveys, administrative data, and behavioural coding.

Research Question 2: Overall, how effective are professional development interventions in the civil service context at improving knowledge, skills, networks, work performance, and productivity?

Overall effectiveness of professional development interventions

We found that most studies reported positive effects of professional development interventions on knowledge, skills, work performance, or productivity. We found that this was true for 15 out of 16 independent studies. However, variation in study characteristics prevented a quantitative synthesis that summarises the magnitudes of these effects.

Variation in effectiveness by study characteristics, type of outcome, and worker

We found no clear patterns indicating that factors such as sample size or risk of bias scores affected the results of the interventions (i.e., direction or statistical significance of intervention effects).

Across outcomes, interventions most consistently improved skills, with the majority of significant positive results found in this category. Skill acquisition was often an immediate target of training and measured through structured assessments, making effects easier to detect. In contrast, work performance and productivity outcomes were less consistently studied. Significantly fewer studies presented results on these outcomes, and nearly all of them stemmed from employees in operational delivery roles (such as call centres and probation services), where measurement of work performance and/or productivity is more straightforward. Only one out of six studies that included these outcomes found positive and statistically significant results. This is a key finding, especially since developing specific skills does not automatically guarantee better job performance or productivity. This limitation in the evidence base is understandable: it is difficult to measure the output of knowledge-intensive roles like

those common in the civil service. It is, however, an important limitation because a key reason for upskilling the workforce is to improve workplace performance or productivity.

Further, findings from individual studies reveal that women benefit more from transformational leadership training when controlling for prior leadership. There was also an indication that public sector workers experienced slightly better outcomes in response to the intervention compared to those in the private sector, although the difference was not statistically significant. Additionally, call centre workers who feel a sense of mutual commitment to their employer exhibit stronger responses to training, highlighting the importance of worker characteristics.

Research Question 3: Which design features, mechanisms, and forms of professional development are associated with the greatest impact?

Influence of intervention design on effectiveness

The synthesis revealed some trends in the design features of professional development interventions, although limitations in reporting and a lack of causal evidence make it difficult to draw definitive conclusions.

We observed no clear pattern between the type of engagement (online, face-to-face, or mixed) and training dosage on the results. However, we observed a pattern between the type of professional development and the results. Most studies combined different professional development methods, with the most common being a mix of lecture and seminar sessions, as well as low-tech simulation-based learning. Notably, all studies that produced statistically significant results used multiple types of professional development, suggesting that combining different approaches may lead to more effective outcomes.

Several individual studies examined how design features impact the effectiveness of professional development interventions. A few studies found that incorporating more intensive or interactive elements into professional development interventions yields better outcomes. Peer coaching, combining training methods, and adding motivational interviewing to existing programmes all significantly improved outcomes. Workshops were more effective than passive information sharing, and training intensity, such as providing advanced support or combining different approaches, produced stronger results.

Mechanisms and forms underpinning professional development interventions

The most common mechanisms in the interventions were behavioural practice/rehearsal, instruction on how to perform a behaviour, and feedback on behaviour. Social support and goal setting also played a significant role in many interventions. However, differences were observed based on profession. Social support and habit formation were more common in interventions for operational delivery workers, while interventions for other professions focused more on feedback mechanisms and action planning. We observed no clear pattern between the number of mechanisms and the direction or statistical significance of results.

Research Question 4: What supports the successful implementation of professional development interventions targeted at driving improvements in knowledge, skills, networks, work performance, and productivity in the civil

service context and adjacent contexts?

The systematic review revealed that few studies explicitly reported on elements related to implementation and process evaluation. Limited or inconsistent adaptations to interventions were often made without assessing their impact. Several studies provided insights into the factors that influenced successful implementation, whether positively or negatively. Organisational factors, such as financial stability, supportive environments, and smaller size, supported effective implementation, while time constraints, competing demands, and cognitive overload were significant barriers.

How did we generate these results?

We conducted a systematic review of evidence on professional development in the civil service and adjacent contexts. A systematic review is ideally suited to providing an unbiased and reliable assessment of existing evidence, using transparent, clearly defined, and replicable procedures.

What did we search for? We searched for English-language RCTs and quasi-experimental studies from OECD countries that examined professional development interventions for the most common UK civil service professions, both within the UK civil service sector and the broader public and private sectors. Eligible studies compared interventions to a comparator group and assessed their impact on knowledge, skills, networks, work performance, and productivity. Based on these eligibility criteria and scoping, we developed a set of search strings around methodology, population, and intervention.

Where did we search? We searched a range of academic bibliographic databases (e.g., Web of Science, Scopus) and grey literature repositories and also conducted a call for evidence. In addition, we screened the backward and forward citations of the studies included through the searches.

Following our search, how did we confirm which studies to include in our analysis? Following our search, we confirmed the inclusion of studies through a structured selection process. After storing all records in a reference management software (Zotero), we removed duplicates and screened titles and abstracts using an online tool (Rayyan). This allowed us to apply our eligibility criteria and assess which studies should proceed to full-text review. We then proceeded to full-text screening, including only studies that met all eligibility criteria, such as geographical focus, population, intervention, outcomes, methodology, and comparator.

How did we extract information from the included studies? Following the selection of the included studies, we extracted information from each study, such as bibliographic information, geographical focus, research design and characteristics of population, intervention, outcomes, and the quantitative data needed for the synthesis. Reviewers first discussed each data element to ensure a clear understanding and piloted the extraction process to check if further guidance was needed. An independent double extraction was performed on a 10% sample of studies to ensure consistency. Any discrepancies found during this process were discussed in meetings. Data extraction continued once all discrepancies were resolved. This process allowed us to gather the necessary study characteristics and outcome data for our analysis.

How did we synthesise this information to answer the research questions? We

first assessed the risk of bias for each study using the revised JBI critical appraisal tools for both RCTs and quasi-experimental studies (Barker et al. 2024; 2023). This evaluation helped us understand the reliability of the study findings by identifying potential limitations in the study design or execution that could affect the accuracy of the reported results. Although no studies were excluded based on this assessment, it allowed us to interpret the findings with appropriate caution, particularly where estimates may have been influenced by methodological weaknesses.

Further details on how we synthesised the results from the included studies to address each research question are provided below.

- **RQ1:** We produced a systematic evidence map, where we coded the studies based on specific characteristics. This included the setting, study design, type of participants (e.g., operational delivery professionals, tax professionals, mixed group), the type of intervention, and outcomes.

- **RQ2:** Although we intended to conduct a meta-analysis, this approach was not feasible due to the small number of independent studies identified as well as the wide range of outcome types and measurement approaches among the included studies. Instead, we conducted a **narrative synthesis** and also used **vote counting** by effect direction. Narrative synthesis involves summarising and interpreting study findings in a descriptive way to identify patterns and insights across the studies. Vote counting involves tallying the number of studies that show positive or negative effects of an intervention, without comparing the magnitude of those effects. To illustrate trends, we also created **effect direction plots,** which visually summarise the balance of positive and negative findings. Namely, we produced effect direction plots based on overall sample size, risk of bias assessment scores, target outcome, and type of worker.

- **RQ3:** To capture the specific mechanisms used in professional development interventions, we developed a tailored **taxonomy** based on the Behaviour Change Techniques (BCT) taxonomy ('BCT Taxonomy (v1): 93 Hierarchically-Clustered Techniques', n.d.). This taxonomy comprises 29 mechanisms, including 27 from the original taxonomy and two new additions. Additionally, two existing mechanisms were amended for clarity and examples for each mechanism were tailored to the professional development context. These updates enhance the applicability of the taxonomy to the study's focus. Due to the limited number of included studies relative to the identified mechanisms, we adopted the same approach as for RQ2. We created effect direction plots based on dosage of training, type of engagement, type of training, and the number of mechanisms. We also produced bar plots to display the number of mechanisms per included paper, the most common mechanisms and forms used, and whether these differences vary depending on the type of worker the interventions target.

- **RQ4:** We analysed studies that included an implementation and process evaluation. Since these evaluations varied in approach and were largely qualitative, we used a **thematic analysis** to identify key factors affecting implementation.

Limitations and implications

Limitations of evidence

A key limitation identified is the lack of studies that meet the eligibility criteria. This can be attributed to the difficulties of conducting experimental research within the civil service and the complexity of measuring productivity outcomes in the public sector. Additionally, there is significant variability in outcome measures, making comparisons difficult and preventing a meaningful meta-analysis. The quality of reporting also presents issues—many abstracts lack clear details on key study components, potentially leading to relevant studies being overlooked. Furthermore, studies often fail to provide sufficient information on how the interventions were conducted, making it harder to assess the risk of bias. Gaps in reporting participant characteristics, such as demographics, also limit the ability to evaluate how interventions work for different groups. Similarly, details on intervention design, duration, and implementation are often insufficient, making it difficult to determine which components drive success. Lastly, process and implementation evaluations, which provide critical insights into how interventions function in practice, are frequently underreported or lack depth. These limitations highlight the need for more rigorous, transparent, and detailed reporting in future research.

Limitations of the review process

The review process itself also faces some limitations, which arose due to necessary methodological decisions. Some relevant studies may have been missed due to the selection of search criteria, limited database access, and a focus on titles and abstracts during the initial screening step. The eligibility criteria prioritised OECD countries for their relevance to the UK civil service and excluded softer outcomes, such as job satisfaction. By focusing on RCTs and quasi-experimental studies, valuable qualitative research may have been omitted. Our synthesis methods—vote counting and narrative synthesis—provided useful insights but did not enable precise comparisons of effect sizes. Additionally, while we used the BCT as a framework, some professional development approaches may not align fully with it. Despite these trade-offs, our approach ensured a focused and meaningful synthesis of the available evidence.

Implications for practice and policy

The review provides strong evidence that professional development interventions have a positive impact on skill development and, to a lesser extent, work performance and productivity.

To bridge the gap between skills and knowledge development and productivity, training programmes should not only clearly define the specific skills they aim to improve but also establish robust methods for evaluating their impact. This includes designing interventions that generate evidence on how skill development translates into measurable improvements in performance or productivity.

When designing professional development interventions, it is essential to establish evaluation frameworks that not only measure overall impact but also identify which specific elements were most effective and why. Additionally, the design of professional development interventions should consider how training integrates with employees'

existing workloads, ensuring that participation is both feasible and sustainable. While further research is needed, practitioners should consider incorporating a mix of training methods, interactive elements, and ongoing support, as this may enhance effectiveness.

Implications for further research

Several insights were identified for further research. We found a clear need for more high-quality research on professional development in the civil service to inform the design of effective interventions. Given the limited availability of evidence specific to the civil service, developing a framework for how best to transfer and apply insights generated in the private sector could help bridge this gap. Additionally, promoting standardised outcome measures would improve comparability across studies and strengthen the evidence base. Improving reporting standards is also crucial; abstracts should systematically include key PICO elements, and studies should provide detailed descriptions of interventions, populations, and experimental designs. Implementation and process evaluations should be incorporated to better understand how interventions function in practice. Furthermore, a structured approach to documenting mechanisms, such as integrating the taxonomy into intervention design and evaluation frameworks, would enhance transparency and comparability of results. As more studies emerge, the taxonomy can be refined to better reflect the nuances of professional development in the public sector.

References

AMEE. 2024. 'AMEE Guides'. 2024. https://amee.org/amee-guides/

Barker, Timothy H., Nahal Habibi, Edoardo Aromataris, Jennifer C. Stone, Jo Leonardi-Bee, Kim Sears, Sabira Hasanoff, et al. 2024. 'The Revised JBI Critical Appraisal Tool for the Assessment of Risk of Bias for Quasi-Experimental Studies'. *JBI Evidence Synthesis* 22 (3): 378–88. https://doi.org/10.11124/JBIES-23-00268

Barker, Timothy H., Jennifer C. Stone, Kim Sears, Miloslav Klugar, Catalin Tufanaru, Jo Leonardi-Bee, Edoardo Aromataris, and Zachary Munn. 2023. 'The Revised JBI Critical Appraisal Tool for the Assessment of Risk of Bias for Randomized Controlled Trials'. *JBI Evidence Synthesis*, February. https://doi.org/10.11124/JBIES-22-00430

'BCT Taxonomy (v1): 93 Hierarchically-Clustered Techniques'. n.d. https://digitalwellbeing.org/wp-content/uploads/2016/11/BCTTv1_PDF_version.pdf

Chapman, R.J.K. 1987. 'Teaching Public Administration: Education or Training?' *Australian Journal of Public Administration* 46 (4): 363–75. https://doi.org/10.1111/j.1467-8500.1987.tb02579.x

Education Endowment Foundation. 2021. 'Effective professional development'. https://educationendowmentfoundation.org.uk/education-evidence/guidance-reports/effective-professional-development

Elliott, I.C., K.A. Bottom, and K. O'Connor. 2023. 'The Status of Public Administration Teaching in the UK'. *Journal of Public Affairs Education* 29 (2): 262–74. https://doi.org/10.1080/15236803.2023.2202609

Filges, T., C. Torgerson, L. Gascoine, J. Dietrichson, C. Nielsen, and B.A. Viinholt. 2019. 'Effectiveness of Continuing Professional Development Training of Welfare Professionals on Outcomes for Children and Young People: A Systematic Review'. *Campbell Systematic Reviews* 15 (1): 1060. https://doi.org/10.1002/cl2.1060

Sims, S., H. Fletcher-Wood, A. O'Mara-Eves, S. Cottingham, C. Stansfield, J. Herwegen, and J. Anders. 2021. 'What Are the Characteristics of Teacher Professional Development That Increase Pupil Achievement? A Systematic Review and Meta-Analysis'. *Education Endowment Foundation*. https://educationendowmentfoundation.org.uk/education-evidence/evidence-reviews/teacher-professional-development-characteristics