

Northern Territory Government Program Evaluation Framework: Process Evaluation Report

Department of Treasury and Finance

Program Evaluation Framework



Document title	Northern Territory Government Program Evaluation Framework: Process Evaluation Report
Contact details	Department of Treasury and Finance
Approved by	Senior Management Group
Date approved	
Document review	
TRM number	

Version	Date	Author	Changes made
1.0	September 2022	Program Evaluation Unit	First version
1.1	April 2023	Program Evaluation Unit	Incorporating feedback from Department of Treasury and Finance executives

Acronyms	Full form
BRS	Budget Review Subcommittee of Cabinet
DTF	Department of Treasury and Finance
ICT	information and communications technology
NTPS	Northern Territory Public Sector
PECoP	Northern Territory Government Program Evaluation Community of Practice
PEF	Northern Territory Government Program Evaluation Framework
PEU	Department of Treasury and Finance Program Evaluation Unit

Contents

1. Executive summary	4
2. Background	6
2.1. Treasurer's Direction Evaluation and Review 1993–2001	6
2.2. Program Evaluation Framework	6
3. The Program Evaluation Framework	8
3.1. Developing the Program Evaluation Framework	8
3.2. Overview of the Program Evaluation Framework	8
4. Evaluation methods	12
4.1. Data collection methods	12
4.2. Ethical considerations	13
4.3. Limitations	13
5. Timeline of key program evaluation activities	14
6. Findings	15
6.1. To what extent has the PEF been implemented?	15
6.2. Are the guidance tools and templates being used by agencies?	24
6.3. How user-friendly and appropriate have the guidance tools been for agencies?	25
6.4. How ready is the PEF for an outcome evaluation?	27
6.5. Are there any adjustments to the implementation approach that need to be made?	40
7. Recommendations	41
8. Conclusion	42
9. Appendix A – PEF program logic	43
10. Appendix B – Data matrix	44
11. Appendix C – Evaluation maturity	49

1. Executive summary

The Northern Territory Government Program Evaluation Framework (PEF) was released in May 2020 to support Territory Government agencies to achieve the best outcomes within allocated budget and build a contextualised evidence-base of what works in the Territory.

This process evaluation assesses the initial implementation of the PEF, identifies early issues regarding program delivery and sets the baseline for the scheduled 2024-25 outcome evaluation. The key evaluation questions for the PEF process evaluation were:

1. To what extent has the PEF been implemented?
2. Are the guidance tools and templates being used by agencies?
3. How user-friendly and appropriate have the guidance tools been for agencies?
4. How ready is the PEF for an outcome evaluation?
5. Are there any adjustments to the implementation approach that need to be made?

The evaluation included a desktop review of Cabinet submissions, agency program master lists and evaluation schedules, the evaluation register, Department of Treasury and Finance (DTF) website analytics and internal Program Evaluation Unit (PEU) records. It also included an employee survey¹ capturing awareness of the PEF and toolkit engagement with the PEU, evaluation perceptions, evaluation maturity and individual evaluation capability.

1. To what extent has the PEF been implemented?

The PEF included six key reforms, all of which have been implemented:

1	Annual schedule of evaluations	First annual evaluation schedule approved by the Budget Review Subcommittee of Cabinet in 2021. Subsequently embedded into business as usual processes.
2	Evaluation register	Established in January 2020. As at April 2023, there were 115 Territory Government monitoring and evaluation reports and 122 other relevant resources. DTF regularly updates the register as part of business as usual processes.
3	Evaluation overview and sunset clause requirements in Cabinet submissions	Cabinet submission template updated to include evaluation requirements in August 2020.
4	Evaluation guidance tools	Toolkit published online in November 2020. DTF implements updates as required, including responding to user feedback.
5	Evaluation advice and feedback to agencies	Since 2020, DTF has been requested to provide evaluation advice and feedback on 54 programs across a range of government agencies. Also embedded into business as usual Cabinet submission review processes.
6	Evaluation capacity development	Since 2019, DTF has delivered a total of 40 evaluation-related presentations to 1211 attendees. The Program Evaluation Community of Practice (PECoP) membership has steadily increased – as at April 2023, the PECoP consisted of 179 members.

¹ A total of 41 staff responded, including 35 Program Evaluation Community of Practice members (response rate of 23%) and 6 non-members.

2. Are the guidance tools and templates used?

The PEU webpage has had 11,700 unique page views since May 2020 (the release of the PEF). There were 5,579 unique page views in the year to September 2022 and 5,312 in the year to September 2021. The PEF has been downloaded from the website 1,329 times, with 1,238 unique downloads.

3. How user-friendly and appropriate have the guidance tools been for agencies?

Of the survey respondents who were aware of the PEF and toolkit², the majority:

- had referred to it in the last year (82.5%),
- agreed it is relevant to their work (97.5%)
- agreed the guidance is user-friendly (97.5%).

4. How ready is the PEF for an outcome evaluation?

DTF is taking a staged approach to implementing the PEF (as reflected in the program logic at Appendix A). The PEU's focus in 2022 is on continuing to build awareness of the PEF across the Northern Territory Public Sector (NTPS), ensuring new programs have evaluation strategies and helping agencies prioritise existing programs for evaluation.

As more evaluations are completed, the focus will shift to the PEF's medium-term outcomes of quality evaluations and using evaluation findings. DTF has drafted evaluation standards and will begin to assess evaluation quality from 2023. Initial data is available for each of the seven outcome areas as outlined in the data matrix at Appendix B.

5. Are there any adjustments to the implementation approach that need to be made?

The implementation approach has been adjusted since commencement of the PEF and respondents to the staff survey suggested several areas for further improvement, which form the basis for the recommendations in this report. Overall, feedback from the staff survey was overwhelmingly positive.

Recommendations

Key recommendations include strengthening the toolkit, improving awareness of the PEF and the role of DTF, actively following up with agencies to ensure awareness of their responsibilities and promoting professional development opportunities. Two recommendations have been included to ensure outcome data on evaluation quality standards and improved coordination will be captured for use in the outcome evaluation of the PEF.

² 40 (97.5%) survey respondents were aware of the PEF and toolkit.

2. Background

The Territory Government has implemented two separate whole of government evaluation approaches in the past 30 years:

- The Treasurer's Direction Evaluation and Review 1993–2001.
- The Program Evaluation Framework (2020).

2.1. Treasurer's Direction Evaluation and Review 1993–2001

The first evaluation approach was established in 1993 with the Treasurer's Direction Program Evaluation and Review. This Treasurer's Direction was in effect for eight years and required agencies to review all their functions at least once every three years. A brief summary of the results of any evaluation needed to be published in the agency's annual report.

A Performance Management System audit in 2000³ investigated how well program evaluation supported organisational performance under the Treasurer's Direction. The audit found:

- there was inconsistency in how agencies recognised and used program evaluation
- no central agency was providing clear leadership on how and why to carry out program evaluation
- there was general compliance with the Treasurer's Direction on Evaluation and Review but it appeared unlikely that program evaluation would be conducted in a structured and formal way if the government requirement did not exist.

Overall, the Treasurer's Direction Evaluation and Review had created a tick-the-box mindset rather than a genuine evaluation culture. The audit concluded that a program evaluation culture would be evident if:

- program evaluation was recognised in training programs and management practices
- Cabinet submissions on major policy issues included an evaluation strategy
- evaluation summaries in annual reports demonstrated an understanding of how evaluation was used.

The Treasurer's Direction Evaluation and Review was withdrawn in 2001 (the year following the audit) with the intention to develop "a revised evaluation and review policy that more accurately reflects the new outputs-based environment...in the Northern Territory"⁴.

2.2. Program Evaluation Framework

The PEF was developed by DTF following the establishment of the PEU in 2018 and recommendations arising from the 2019 Fiscal Strategy Panel's final report.

2.2.1. Establishing the Program Evaluation Unit

In 2017, DTF commissioned a capability review for an independent assessment of DTF's strengths and areas for improvement. The review found that DTF is a high performing agency and noted there was an opportunity to "lead the development of program evaluation capability to make evidence-based decisions

³ Auditor General for the Northern Territory February 2000 Report to the Legislative Assembly https://ago.nt.gov.au/reports/2000_February.pdf.

⁴ Department of Treasury and Finance website https://treasury.nt.gov.au/data/assets/pdf_file/0004/481549/TD-EG-P3S4.pdf.

and support other agencies to do the same". The review noted the commitment to build the evidence base was supported by other agencies consulted as part of the review.

The review recommended DTF elevate the importance of evaluation through:

- developing a deeper understanding of agency programs
- building a robust evidence-based approach to program design, implementation and evaluation that supports high quality advice to government.

In response to this recommendation, DTF established the PEU in October 2018, initially comprising one staff member (Senior Administrative Officer 1) with an additional staff member joining the team in July 2020 (Administrative Officer 7) and further support provided by a rotating graduate position.

The PEU was established as a sub-unit within the Budget Development team to make the most of the synergies between the work units and to provide resource flexibility and surge capacity. A later restructure resulted in the development of a key senior role in DTF (Senior Director Budget Development and Evaluation) with both the Assistant Director Program Evaluation and Director Budget Development reporting to this role.

In addition to implementing the PEF, the PEU is also responsible for coordinating reviews under the Organisational Review Framework⁵ and analysing relevant Cabinet submissions from an evaluation perspective.

2.2.2. A plan to fix the budget

Soon after the PEU was established, the need for a whole of government approach to program evaluation was highlighted in the 2019 Fiscal Strategy Panel's final report: [A plan for budget repair](#). The report made four evaluation-related recommendations and noted:

"In comparison to other jurisdictions, the Territory lacks a robust whole of government program evaluation framework to support consistent and regular assessment of agency activities to determine their effectiveness and efficiency."

The Territory Government's response to the report, [A plan to fix the budget](#), accepted the four evaluation-related recommendations, being:

- develop a whole of government program evaluation framework
- update the Cabinet submission template to include program evaluation requirements
- develop an annual schedule of evaluation of existing programs for approval by the Budget Review Subcommittee of Cabinet
- mandate the use of sunset clauses in new programs.

The PEU in DTF was responsible for implementing these program evaluation reforms.

⁵ The Organisational Review Framework establishes a rolling program of organisational reviews to ensure agency expenditure is aligned to government priorities and that services are being provided efficiently https://treasury.nt.gov.au/_data/assets/pdf_file/0003/1018344/Attachment-G-Northern-Territory-Government-Agency-Organisational-Review-Framework.pdf.

3. The Program Evaluation Framework

3.1. Developing the Program Evaluation Framework

In developing the PEF, DTF utilised existing evaluation expertise across the NTPS by establishing a PECoP and considered lessons learnt from other jurisdictions and expert feedback.

3.1.1. Establishing the Program Evaluation Community of Practice

DTF identified that there was existing evaluation expertise within agencies and established the PECoP in October 2019 to provide relevant officers the opportunity to give feedback on the proposed program evaluation reforms. More broadly the PECoP was established to:

- share evaluation knowledge, expertise and experiences across agency boundaries
- promote the role of evaluation in delivering public value
- integrate evaluation into program design and policy development
- improve evaluation skills across the NTPS.

The PECoP actively engaged in the development of the PEF, with over 40 members providing officer-level feedback.

3.1.2. Lessons learnt from other jurisdictions and expert feedback

DTF consulted with the central government agency program evaluation units in Western Australia, New South Wales and the Australian Capital Territory as well as the then Department of Industry, Innovation and Science within the Commonwealth Government to draw on lessons learnt from the establishment and operation of program evaluation initiatives in these jurisdictions.

Feedback from other jurisdictions indicated it takes a sustained effort over a number of years to embed an effective whole of government monitoring and evaluation system. It was also clear from the discussions that a key factor for successful implementation of the PEF would be to embed evidence-based policy and evaluative thinking across the NTPS and shift perceptions of evaluation as a compliance exercise to being an integral part of program design and policy development.

Evaluation experts at Charles Darwin University and Menzies School of Health Research also provided positive and constructive feedback on the draft PEF.

3.2. Overview of the Program Evaluation Framework

The PEF was released in 2020 and is publicly available on the DTF website⁶. The PEF covers Territory Government-funded programs (that is, excludes programs that are wholly externally funded) and, as a general rule, does not apply to infrastructure and information and communications technology (ICT) projects, which are covered by separate review processes.

⁶ <https://treasury.nt.gov.au/dtf/financial-management-group/program-evaluation-unit>.

The PEF integrates evaluation into Territory Government policy and budget development processes and aims to improve transparency and accountability, and encourage better use of Territory Government funds by:

- ensuring new programs and extensions to existing programs have identified goals and objectives that are achievable and measurable, or include actions to develop measurement as part of the program
- ensuring new programs and extensions to existing programs have an evaluation strategy
- incorporating sunset provisions in new programs, linked to evaluation outcomes
- establishing a rolling schedule of evaluations to ensure existing programs are evaluated over time
- providing a clear mandate for agencies to evaluate their programs and target investment
- outlining expected evaluation principles and standards
- providing government with clear advice about the costs and benefits of evaluation (including data collection and analysis) to help inform evaluation decisions
- establishing a protocol for policy and program officers to plan for evaluation across the program lifecycle (with a step-by-step guide in the online evaluation toolkit)
- establishing a tiered system of evaluations to ensure evaluation is proportionate to the cost, risk and complexity of a program
- describing how the Territory Government can build evaluation capability within the NTPS and foster a culture of continuous improvement
- outlining how the Territory Government will measure progress in implementing the framework.

3.2.1. Roles and responsibilities

Central oversight is critical to developing a strategic whole of government approach to evaluation and strengthening evaluation culture⁷. The PEF includes a centralised approach to program evaluation that aims to support:

- a consistent standard of evaluation across agencies
- an ability to identify systemic issues across government
- capacity to set strategic priorities for and identify gaps in evaluation
- accountability for multi-agency and whole of government programs
- coordinated capability building, resourcing, data collection, reporting and evaluative effort
- a centralised repository of evaluations to enhance continuous learning and quality improvement.

The PEF sets out the evaluation roles and responsibilities as part of the Territory Government's centralised approach to program evaluation (see Table 1). Under this approach, evaluation is primarily undertaken by individual agencies (this may include using external experts commissioned by the agency) to maintain a close link between the evaluation and the program area with relevant subject matter knowledge and experience.

Evaluation activity is overseen, coordinated and supported by DTF as well as being supported by the Department of the Chief Minister and Cabinet, the Office of the Commissioner for Public Employment and the Department of Corporate and Digital Development.

⁷ Bray, R., Gray, G. 't Hart, P, Evaluation and learning from failure and success, Australia and New Zealand School of Government, 2019.

Table 1: Territory Government program evaluation roles and responsibilities

Agency	Role
Department of Treasury and Finance	<ul style="list-style-type: none"> • promote the use of evaluation in government decision-making • lead the PECoP⁸ • support agencies to complete evaluation strategies as part of their Cabinet submissions • coordinate the rolling schedule of evaluations • strengthen evaluative capacity across the NTPS • prepare an annual whole of government summary of evaluations • maintain a database of completed evaluations and share findings to inform future policy and program design
Department of the Chief Minister and Cabinet	<ul style="list-style-type: none"> • promote the use of evaluation in government decision-making • collaborate on evaluations through appropriate forums • advise on agency evaluation priorities and schedules • advise on the development and enhancement of the evaluation toolkit to support capability building
Office of the Commissioner for Public Employment	<ul style="list-style-type: none"> • promote the use of evaluation in government decision-making • support capability building and evaluation training opportunities
Department of Corporate and Digital Development	<ul style="list-style-type: none"> • facilitate open data, and data sharing and linkage initiatives
Line agencies	<ul style="list-style-type: none"> • conduct or commission evaluations • identify evaluation priorities for the rolling schedule of evaluations (to be determined by the Budget Review Subcommittee of Cabinet) • incorporate lessons learned from previous evaluations into program and policy design • include evaluation strategies as part of Cabinet submissions • report to the PEU on evaluations undertaken and managed • upskill relevant staff to build evaluation capability and capacity • provide feedback on the evaluation toolkit

3.2.2. Key components of the PEF

The key components of the PEF and how they link to intended outcomes are outlined in the program logic in [Appendix A](#).

The key outputs (assessed as part of this process evaluation) are:

- an annual schedule of evaluations
- an evaluation register
- evaluation overview and sunset clause in the Cabinet submission template
- evaluation guidance and tools (toolkit and templates)
- evaluation advice and feedback provided to agencies by DTF
- evaluation capacity development activities and information sharing by DTF.

⁸ A Community of Practice is a group of people with a common interest who share knowledge and best practices.

The key short-term outcomes (to be evaluated in 2024-25) are:

- improved coordination across government and reduced duplication of evaluative effort and evidence of prioritisation
- timely evidence for key decisions and decision-makers
- evaluation and data collection are proportional to the cost, risk and complexity of a program
- increased proportion of evaluations considered at the program planning stage
- increased evaluation maturity across the Territory Government including:
 - awareness and understanding of the value of evaluation
 - evaluation capabilities
 - demand for evaluation and its use.

The medium term outcomes (to be evaluated in 2027-28) are:

- evaluations meeting quality standards
- evaluation findings used by decision-makers and program designers, and recommendations implemented

While this process evaluation looks at whether outcome data is being collected and sets the baseline, the short and medium-term outcomes will be considered in more detail as part of the outcome evaluations scheduled in 2024-25 and 2027-28.

3.2.3. Transition and review

The final section of the PEF considers transition and review. It notes that successful implementation of the PEF requires cultural change across the NTPS including training and development to improve evaluation capability and quality. It sets out the hierarchy of policy objectives and indicative performance measures (reproduced below).

Objective	Indicative performance measures
Build a culture of enquiry, learning and continuous improvement	Percentage of staff who are aware of the program evaluation framework. Percentage of staff who perceive evaluation as an opportunity to improve government services.
Increase government evaluation capability	Whole of government evaluation training available. Increase in the number of staff who have undertaken evaluation training. Improvement in whole of government evaluation maturity ⁹ Increased number of evaluations meeting quality standards.
More informed government decision-making	Increased number of evaluation recommendations implemented.
Higher quality government services	Increased number of programs that use the results from previous evaluations to inform design and implementation.

These objectives and performance measures are reflected in the program logic (at Appendix A) and data matrix (at Appendix B) in this report. Measurement of evaluation maturity is outlined at Appendix C.

⁹ As per Table 5 in the PEF, reproduced at Appendix C.

4. Evaluation methods

This process evaluation has been conducted internally by DTF staff, noting the outcome evaluation in 2024-25 will be externally commissioned.

4.1. Data collection methods

This evaluation has been informed by a desktop review and data received through an employee survey. To ensure all elements of the PEF have been captured, key indicators have been developed against the program logic (outcomes and outputs) and key evaluation questions. The data matrix at [Appendix B](#) outlines the full list of indicators, baseline, targets and data sources.

Desktop review

The desktop review involved review of:

- Cabinet submissions and DTF comments
- agency program master lists and evaluation schedules
- DTF evaluation register
- DTF website analytics
- DTF internal records.

Employee survey

An employee survey was undertaken in September 2022 to collect information on awareness and perceptions relating to evaluation across the NTPS. The survey provided feedback on the implementation of the PEF and informed recommendations for improvement, as well as establishing a baseline for several outcome indicators.

The survey was distributed through the PECoP with a 23% response rate (35 of 154 members at time of survey). Members were asked to forward the survey to relevant staff outside of the PECoP. A total of 41 staff completed the survey, including 6 staff who were not members of the PECoP.

The survey questions captured:

- awareness of the PECoP
- awareness and use of the PEF and toolkit
- evaluation engagement with DTF
- evaluation perceptions
- evaluation maturity
- individual evaluation capability
- open feedback.

The data matrix ([Appendix B](#)) outlines the full list of output and outcome indicators that use the survey as a data source.

4.2. Ethical considerations

As this process evaluation did not seek input from the community or high risk groups, no formal ethics approval was required. An evaluation work plan was developed including a list of key stakeholders and a risk mitigation plan. All data collected from the employee survey was anonymous.

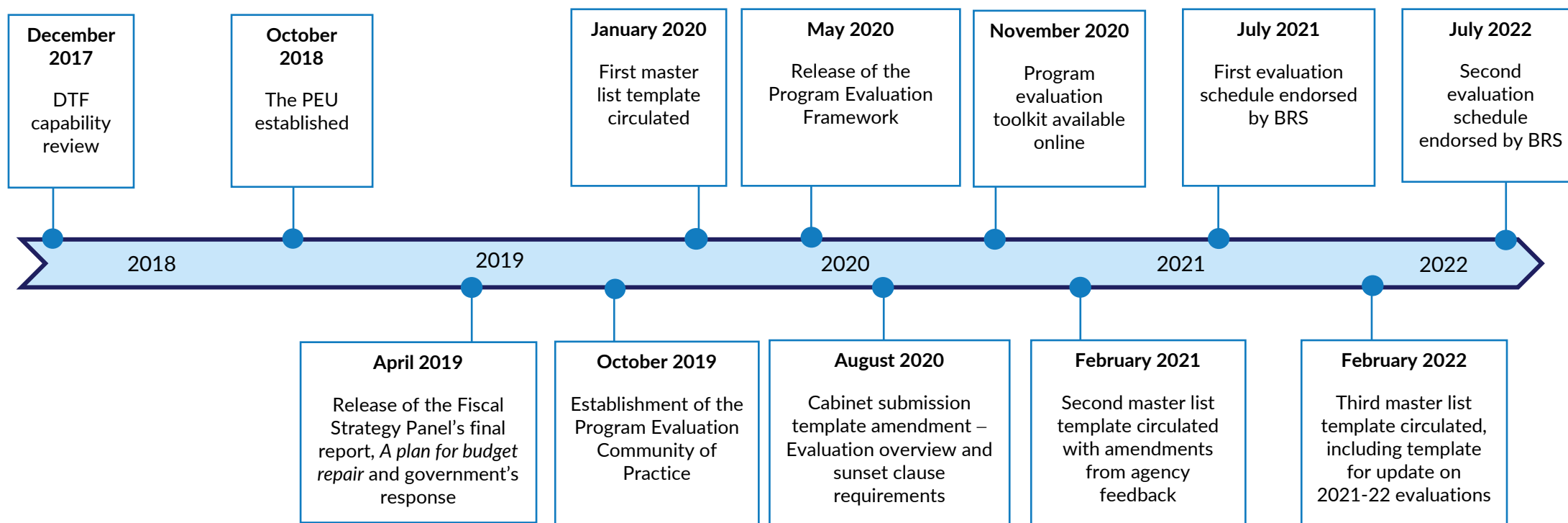
4.3. Limitations

The two main limitations of this process evaluation were the survey sample size and the difficulty measuring evaluation capability and maturity.

The survey was circulated primarily to existing PECoP members and their networks. While this cohort is the key target group of the PEF, it is not necessarily a representative sample of all relevant NTPS staff for questions such as awareness of the PEF.

There is a degree of subjectivity in determining evaluation capability and maturity. The survey captured individuals' perceptions of the level of maturity at a point in time. It is possible that staff perceptions may change as awareness, knowledge and skills increase over time (that is, what they initially determined as a high level of maturity might become a lower classification as their knowledge base increases). This could be addressed in future evaluations by triangulating perception data with other measures, such as changes in the quality of evaluation reports over time against objective standards.

5. Timeline of key program evaluation activities



6. Findings

This section of the report sets out the findings in relation to each of the evaluation questions:

1. To what extent has the PEF been implemented?
2. Are the guidance tools and templates being used by agencies?
3. How user-friendly and appropriate have the guidance tools been for agencies?
4. How ready is the PEF for an outcome evaluation?
5. Are there any adjustments to the implementation approach that need to be made?

6.1. To what extent has the PEF been implemented?

The PEF included six key reforms, all of which have been completed:

1	Annual schedule of evaluations	First annual evaluation schedule approved by the Budget Review Subcommittee of Cabinet in 2021. Subsequently embedded into business as usual processes.
2	Evaluation register	Established in January 2020. As at April 2023, there were 115 Territory Government monitoring and evaluation reports and 122 other relevant resources. PEU regularly updates the register as part of business as usual processes.
3	Evaluation overview and sunset clause requirements in Cabinet submissions	Cabinet submission template updated to include evaluation requirements in August 2020.
4	Evaluation guidance tools	Toolkit published online in November 2020. PEU implements updates as required, including responding to user feedback.
5	Evaluation advice and feedback to agencies	Since 2020, DTF has been requested to provide evaluation advice and feedback on 54 programs across a range of government agencies. Also embedded into business as usual Cabinet submission review processes.
6	Evaluation capacity development	Since 2019, DTF has delivered a total of 40 evaluation-related presentations to 1211 attendees. The PECoP membership has steadily increased - as at April 2023, the PECoP consisted of 179 members.

A Performance Management System Audit by the Auditor-General in March 2022 assessed the actions taken to address the recommendations in the Fiscal Strategy Panel's final report: [A plan for budget repair](#) and did not identify any issues regarding the implementation of the program evaluation reforms. The [Treasurer's Direction – Organisational performance and accountability](#), released in November 2022, establishes the minimum standards for Territory Government agencies to plan, review and report their organisational performance and accountability, including evaluation.

6.1.1. Output 1: Annual schedule of evaluations

Program stage	Indicator	Baseline	Target	Result
Output 1: Annual schedule of evaluations coordinated by DTF (prioritised using agency program master lists)	Each agency has a program master list to prioritise program evaluations	Program master lists and evaluation schedules did not previously exist	All agencies submit a program master list to DTF each year	Yes
	Number of programs on the whole of government program master list		Number of programs on the whole of government program master list remains steady or increases each year	The 2022-2023 list had an additional 50 programs compared to the previous year (increasing from 266 to 316 programs)
	A whole of government annual schedule is submitted and endorsed by the Budget Review Subcommittee of Cabinet (BRS)		Evaluation schedule submitted to, and endorsed by, BRS each year	Yes

Program master list

The PEF requires agencies to review their existing stock of programs and develop evaluation strategies for priority programs over time. Monitoring and evaluation requires the commitment of resources so it is necessary to balance the cost of evaluation and the risk of not evaluating, noting that sometimes monitoring will be sufficient.

Building on lessons learnt from other jurisdictions¹⁰, the annual program master list is designed to capture all Territory Government-funded programs that *could* be evaluated to help prioritise evaluations. It identifies the extent to which existing programs have been evaluated and the proposed timing of any future evaluation.¹¹

The first master list template was adapted from a template shared by a government agency from the New South Wales. The initial Territory Government template required agencies to provide:

- the program name and key contacts
- a brief description of the program
- alignment with Territory Government priorities
- strategic alignment (links to associated Territory Government or agency strategic plans)
- program tier (based on the tier classification in the PEF)

¹⁰ In 2016, the New South Wales Auditor-General undertook a [performance audit of the NSW Government's program evaluation initiative](#). The audit set out the good practice model expected from each agency to prepare an evaluation schedule including a master list of all current agency programs with their tier ranking and linkage to government priorities. NSW Auditor-General's Report to Parliament, Implementation of the NSW Government's program evaluation initiative, 2016 accessed October 2020.

¹¹ Australian Public Sector Review, Evaluation and learning from failure and success; and 2016 NSW Auditor-General's Report to Parliament, Implementation of the NSW Government's program evaluation initiative.

- whether the funding is time-limited or ongoing
- if and when the program was last evaluated
- if and when there are plans to evaluate the program in future
- program budget per financial year.

DTF circulated the first master list template as part of a 2021 Budget Circular in January 2020 and received program lists from each agency in June 2020. The individual agency program lists were collated into a whole of government program master list and cross checked by DTF against annual reports, Cabinet decisions, grant programs, government awarded contracts and press releases to identify potential gaps.

The task of completing a full program master list for the first time presented a number of challenges for agencies. DTF identified gaps and inconsistencies in the draft program lists and timelines for completion exceeded initial estimates. DTF received feedback on the useability and content within the template and several common pain points were identified.

The lessons learnt from the first year informed the development of a revised program master list template for 2021-22, with the following changes:

- streamlined and simplified – when printed, it fits on one page.
- less of an emphasis on budget – the detailed budget columns were replaced with a single column for approximate annual budget.
- increased emphasis on evidence – under ‘evaluation activity’ agencies were encouraged to consider the overall evidence base for a program including research, reviews and audits.
- strategic alignment – added diagrams and examples in the instruction tabs to help agencies align their programs with overarching strategies
- updated the ‘Purpose’, ‘Definitions’ and ‘What to include’ tabs with extra information.

The second amended program master list template was circulated in February 2021 and agencies commented that the new version was easier to work with.

DTF collates the agency-level program lists into a whole of government program master list and continues to work with agencies to improve the comprehensiveness of the whole of government program master list, noting the 2022-23 list had an additional 50 programs compared with the previous year (increasing from 266 to 316 programs).

Whole of government evaluation schedule

At an agency level, the program lists are used to identify which programs should be prioritised for evaluation. Programs planned for evaluation in the coming financial year make up the proposed agency-level evaluation schedule (that is, the agency’s evaluation schedule is a subset of the agency’s program list).

A whole of government evaluation schedule is compiled by DTF by collating the agency-level evaluation schedules. The proposed evaluation schedule is submitted to BRS for endorsement, along with an annual summary of evaluation findings for the previous year.

The first whole of government evaluation schedule was collated for 2021-22 and endorsed by BRS in June 2021. The second evaluation schedule (for 2022-23) was endorsed by BRS in July 2022 along with an update on the progress and outcomes of the evaluations scheduled for 2021-22. The update on completed evaluations includes a DTF summary of the findings (including comments on the quality of the evaluation where relevant) and an agency response to the evaluation findings and recommendations.

6.1.2. Output 2: Evaluation register

Program stage	Indicator	Baseline	Target	Result
Output 2: Evaluation register	Evaluation and monitoring reports added to the evaluation register as they are completed (either on the DTF website or on the internal register)	Evaluation register did not previously exist	Number of evaluation and monitoring reports on the DTF website remains steady or increases each year	As at April 2023, there were 115 Territory Government monitoring and evaluation reports and 122 other relevant resources (such as evaluations of Territory based programs by non-government organisations or the Commonwealth).

DTF is responsible for maintaining a register of all completed Territory Government evaluations. This central repository of evaluations enables systemic assessment of lessons learnt, whether positive or negative, helping improve future programs and strengthening policy design across agencies.

All publicly available reports have been collated and are regularly updated on the DTF website¹² to inform future program and policy planning. As at April 2023, there were 115 Territory Government monitoring and evaluation reports and 122 other relevant resources (such as evaluations of Territory-based programs by non-government organisations or the Commonwealth). DTF also maintains an internal register of evaluation reports that are available by request pending permission of the host agency

Of the 13 evaluations completed from the evaluation schedule in 2021-22, DTF has received a copy of 8 evaluation reports as at April 2023. Requests for the remaining reports will be followed up as part of the next program master list process.

6.1.3. Output 3: Cabinet submission template includes evaluation overview and sunset clause considerations

Program stage	Indicator	Baseline	Target	Result
Output 3: Cabinet submissions template includes evaluation overview and sunset clause considerations	Cabinet submission template updated with evaluation overview template and sunset clause considerations	Original Cabinet submission template	Cabinet submission template updated	Yes

DTF made amendments to the Cabinet handbook and Cabinet submission template in 2020 to include evaluation overviews and sunset clause requirements.

¹² https://treasury.nt.gov.au/dtf/financial-management-group/program-evaluation-unit#ntg_published_evaluations

Evaluation overview

The Cabinet submission template and handbook changes require an evaluation overview for all new programs seeking additional funding that exceeds \$1 million in a financial year¹³. The evaluation overview should be a concise summary of the key outcomes the program is trying to achieve and how and when success will be measured. A full evaluation work plan must be submitted to DTF if the program is approved to proceed.

The evaluation overview template prompts agencies to outline:

- what the program is aiming to achieve
- how the program will achieve this – at a minimum by stating the outputs and related outcomes, or by attaching a program logic
- what external factors may also influence the program's outcomes (such as other programs trying to achieve the same outcome)
- how the program's success will be measured – indicators, baselines, targets and data sources
- what evaluations will be required in the first five years of the program, including the timing and cost of the evaluations.

Sunset clauses

Information about sunset clauses has been included in the Cabinet handbook and the Cabinet submission template prompts the inclusion of a sunset clause date. A sunset clause is a built-in decision point for government. Unless otherwise directed by Cabinet, funding for new programs (or extensions of existing programs) that exceeds \$1 million in a financial year are subject to an initial five-year sunset clause. This aims to ensure ongoing funding for programs is informed by evaluation.

Originally, programs subject to a sunset clause were defined as 'funded for a finite period, with the decision for further funding (either wholly or in part) informed by an evaluation'. Once the Treasurer's Direction – Organisational Performance and Accountability came into effect in November 2022, the PEF and the Cabinet Handbook were updated with a more flexible definition: 'a specified period after which funding for a program is reviewed or ceases'.

6.1.4. Output 4: Evaluation guidance and tools

Program stage	Indicator	Baseline	Target	Result
Output 4: Evaluation guidance tools (toolkit and template)	Whole of government program evaluation toolkit and templates published on the DTF website	Toolkit and templates did not previously exist	Toolkit and templates published	Yes

DTF has developed a program evaluation toolkit to support Territory Government employees who are responsible for policy development, program design and program development. It provides guidance, resources and templates to help managers integrate evaluation planning into program design, prepare a program for evaluation, commission and manage an external evaluation or lead an internal evaluation.

¹³ Excluding infrastructure and ICT programs.

The toolkit is structured around the program evaluation cycle and guides managers through completing the two key templates: the program evaluation overview (part of the Cabinet submission template) and the full evaluation work plan. The toolkit and the templates have been adapted over time in response to stakeholder feedback.

The PEF and toolkit is publicly available via the PEU webpage on the DTF website.

6.1.5. Output 5: Evaluation advice and feedback provided to agencies by DTF

Program stage	Indicator	Baseline	Target	Result
Output 5: Evaluation advice and feedback to agencies by DTF	Number of programs on which DTF has provided evaluation advice or feedback	Previously there was not a central agency providing evaluation advice and feedback to agencies	Number of programs on which DTF has provided evaluation advice or feedback remains steady or increases each year	Territory Government agencies have requested DTF evaluation advice and feedback on 54 different programs. DTF has responded to each of these requests
	Number and proportion of survey respondents who agree evaluation guidance or feedback had been useful and relevant to their team		Number and proportion of survey respondents who agree evaluation guidance or feedback had been useful and relevant to their team remains steady or increases each year	24 (58.5%) respondents had engaged with DTF for evaluation guidance or feedback. Of these respondents, 100% strongly agreed or agreed the support had been useful and relevant

As at April 2023, Territory Government agencies had requested DTF evaluation advice and feedback on 54 different programs. DTF has responded to each of these requests – sometimes with high level feedback on an appropriate evaluation approach and sometimes with detailed feedback on evaluation plans, including program logics and data matrices.

DTF has been requested to provide evaluation advice on programs across a range of government priorities including domestic and family violence, tourism, economic stimulus, Aboriginal justice, youth justice, health, education and sustainability.



Employee survey results

"It would be great for the Program Evaluation Unit to increase its visibility across agencies"

"If we could clone them so they have capacity to work with all of us directly, that'd be ideal!"

"the staff are approachable and helpful."

"...I would suggest that [DTF's] profile and ability to reach more teams and agencies is only a feature of the resources and time versus [its] very proactive and helpful approach"

- 24 (58.5%) respondents had engaged with DTF for evaluation guidance or feedback.
- Of these respondents, 100% strongly agreed or agreed the support had been useful and relevant to their team

"Great team and much needed"

"Engaging with the team is easy and always insightful"

"Great source of knowledge and advice"

"DTF PEU has provided one-on-one briefings which is valued and they've also helped with drafting documents, which is greatly appreciated"

6.1.6. Output 6: Evaluation capacity development activities and information sharing by DTF

Program stage	Indicator	Baseline	Target	Result
Output 6: Evaluation capacity development activities and information sharing by DTF	Number of evaluation capacity development activities coordinated or promoted by DTF. This indicator is from the PEF	Previously, there was not a central agency sharing evaluation capacity and development activities	Number of evaluation capacity development activities coordinated or promoted by DTF remains steady or increases each year	DTF promotes about 100 evaluation presentations or workshops each year
	Number of DTF evaluation presentations and number of attendees		Number of DTF presentations and number of attendees remains steady or increases each year	Since 2019, DTF has delivered a total of 40 evaluation-related presentations to 1211 attendees
	Number and proportion of survey respondents who have attended evaluation training promoted by DTF		Number and proportion of survey respondents who have attended evaluation training promoted by DTF remains steady or increases each year	25 (61%) respondents had attended evaluation training promoted by DTF
	Percentage of survey respondents who agreed DTF shares relevant information in PEU emails and seminars		>80% of survey respondents agreed DTF shares relevant information	Of the 24 respondents that had engaged with DTF for evaluation guidance and feedback, 100% agreed that DTF shared relevant information in PEU emails and seminars

DTF has been encouraging evaluation capacity development by organising evaluation-related presentations, sharing relevant information with the PECoP, and organising and promoting evaluation training and professional development.

Since 2019, DTF has delivered a total of 40 evaluation-related presentations to 1,211 attendees. Target groups have included a range of Territory Government agencies (executive and officer level), non-government sector, Australian Evaluation Society members, Charles Darwin University (CDU) Territory Government Agreement Executive Committee, Public Sector Management Program students, CDU Masters of Public Policy students, and Territory Government graduates. In addition, DTF promotes about 100 evaluation presentations or workshops each year.

The PECoP is a key mechanism for building evaluation awareness and capacity. PECoP membership has steadily increased since establishment – as at April 2023, the PECoP consisted of 179 members.

DTF worked with the Office of the Commissioner for Public Employment to include evaluation training providers in the across government training panel contract to make it easier to access training.



Employee survey results

Training and professional development

- 25 (61%) respondents had attended evaluation training promoted by DTF
- Respondents identified the following areas for future training and professional development opportunities:
 - program logic and development of evaluation questions
 - bias in data and statistics
 - developing evaluation criteria
 - systems theory
 - working with consultants and commissioning evaluations
 - process evaluation and implementation science
 - end to end policy development, including evaluation planning
 - how to retrofit evaluation to existing programs that were not designed for evaluation
 - more Australian Evaluation Society workshops
 - more university level opportunities.

Information sharing

Of the 24 respondents that had engaged with DTF for evaluation guidance and feedback, 100% agreed that DTF shared relevant information in PEU emails and seminars.

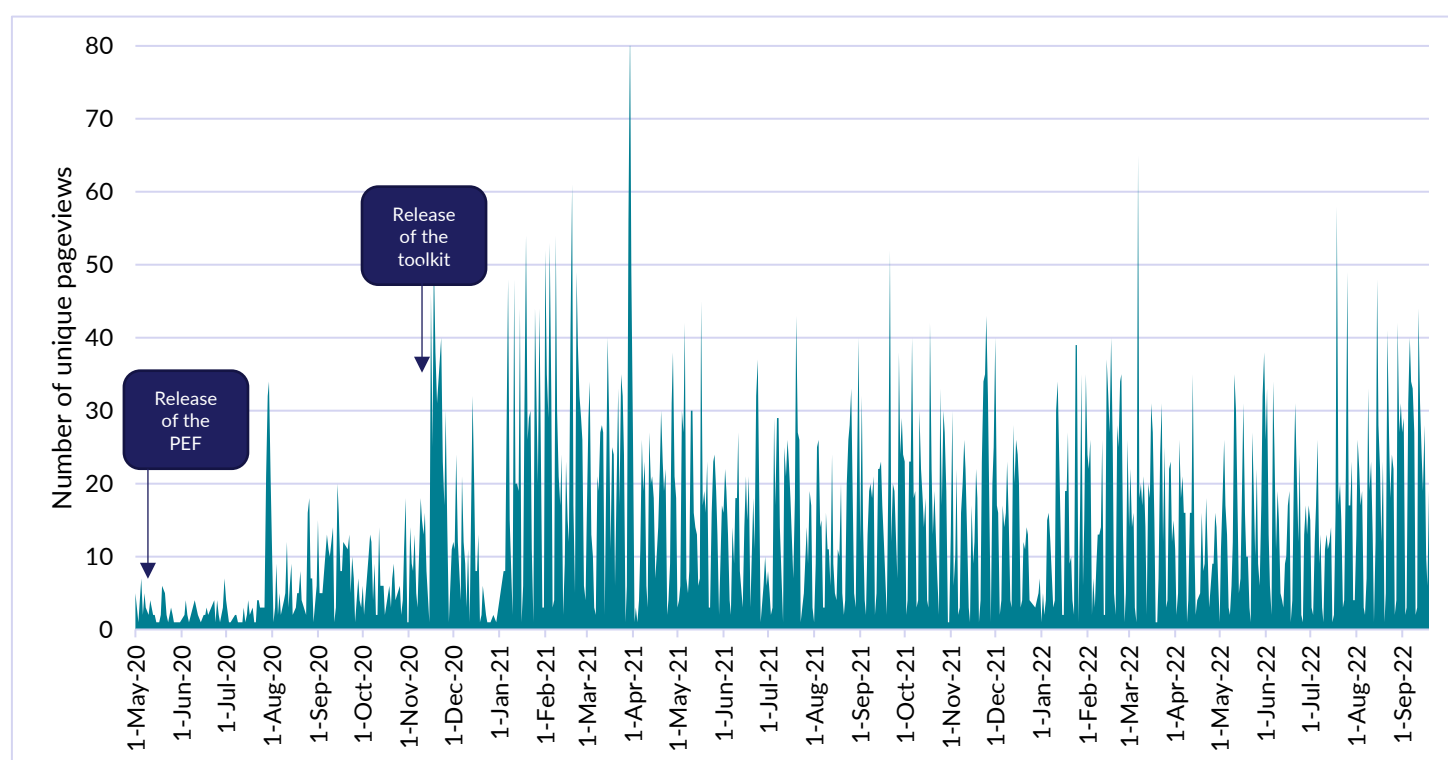
6.2. Are the guidance tools and templates being used by agencies?

Program stage	Indicator	Baseline	Target	Result
Output 4: Evaluation guidance tools (toolkit and template)	Number of DTF PEU webpage unique page views	Webpage and PEF did not previously exist	Number of DTF PEU webpage unique page views remains steady or increases each year	The DTF PEU webpage has had 11,700 unique page views since May 2020 (the release of the PEF). In the year to September 2022
	Unique downloads of the PEF		Number of unique downloads of the PEF remains steady or increases each year	1,238 unique downloads of the PEF
	Percentage of survey respondents who have referred to the PEF and toolkit in the last year		>80% of survey respondents have referred to the PEF and toolkit in the last year	33 respondents (82.5%) have referred to the PEF in the last year

The DTF PEU webpage has had 11,700 unique page views since May 2020 (the release of the PEF). In the year to September 2022, there were 5,579 unique page views compared to 5,312 in the year to September 2021. The PEF has been downloaded from the website 1,329 times, with 1,238 unique downloads.

PEF webpages with the most views included:

- section 2: Complete the evaluation work plan
- section 2.5: Evaluation methodology
- section 1: Complete the evaluation overview
- templates, acronyms and glossary.

Graph 1: Number of unique DTF PEU website page views from May 2020 to September 2022

Forty (97.5%) survey respondents were aware of the PEF and toolkit. Of those who were aware of the PEF and toolkit:


- 33 (82.5%) referred to it in the last year
- 39 (97.5%) strongly agreed or agreed the PEF and toolkit is relevant to their work.

6.3. How user-friendly and appropriate have the guidance tools been for agencies?

Program stage	Indicator	Baseline	Target	Result
Output 4: Evaluation guidance tools (toolkit and template)	Stakeholder feedback on how user-friendly and appropriate the toolkit and templates are	Toolkit and templates did not previously exist	>80% of survey respondents agree or strongly agree that the toolkit and templates are user friendly	39 respondents (97.5%) strongly agreed or agreed that the guidance in the PEF and toolkit is user-friendly
	Toolkit and template are adapted over time in response to stakeholder feedback		Toolkit and template are adapted over time in response to stakeholder feedback	Yes

A Territory Government agency-level evaluation strategy report¹⁴ stated that some staff found the evaluation guidance documents inaccessible, describing them as ‘dense’, ‘complex’ and ‘technical’. To understand whether this was a widespread concern, DTF undertook the stakeholder survey so relevant Territory Government staff had the opportunity to give anonymous feedback and suggestions for improvement.

The survey reported an overwhelmingly positive response to the guidance documents (see box below) indicating the comments from the agency-level evaluation strategy report may be more of a reflection of evaluation maturity rather than the quality of the guidance documents.

 **Employee survey results**

"The resources are easy to access, easy to use"

"[the PEF and toolkit] is very wide-ranging which is great. I think it walks the line well between too much information and sufficient information without going off on tangents"

- 40 (97.5%) survey respondents were aware of the PEF and toolkit. Of those who were aware of the PEF and toolkit:
 - 33 (82.5%) have referred to it in the last year
 - 39 (97.5%) strongly agreed or agreed the PEF and toolkit is relevant to their work
 - 39 (97.5%) strongly agreed or agreed that the guidance in the PEF and toolkit is user-friendly

¹⁴ This report was commissioned by a Territory Government agency to help fulfil its evaluation responsibilities under the PEF. The report considered the current state of evaluation at the agency, the future state for evaluation the agency needed to achieve and how evaluation can most effectively be embedded across the agency. The report assessed the agency at the beginning or having developing levels of evaluation maturity.

6.4. How ready is the PEF for an outcome evaluation?

To assess how ready the PEF is for an outcome evaluation, it was necessary to consider whether the appropriate data was being collected for an outcome evaluation and set the baseline, where available, so change can be measured over time.

The short-term outcomes for the 2024-25 outcome evaluation are:

- improved coordination across government, and reduced duplication of evaluative effort and evidence of prioritisation
- timely evidence for key decisions and decision-makers
- evaluation and data collection are proportionate to the cost, risk and complexity of a program
- increased proportion of evaluations considered at the program planning stage
- increased evaluation maturity across the Territory Government including:
 - awareness and understanding of the value of evaluation
 - evaluation capabilities
 - demand for evaluation and its use.

The medium-term outcomes for the 2027-28 outcome evaluation are:

- evaluations meeting quality standards
- evaluation findings used by decision-makers and program designers, and recommendations implemented.

6.4.1. Outcome 1: Improved coordination across government and reduced duplication of evaluative effort

Program stage	Indicator	Baseline	Target	Source
Outcome 1: Improved coordination across government and reduced duplication of evaluative effort	Proportion of staff who agree evaluation is coordinated across the Territory Government	25%	30% of survey respondents agree evaluation is coordinated across the Territory Government by 2024	Employee survey

With central oversight of evaluation activity across government, DTF is able to identify duplication and facilitate coordination between Territory Government agencies. For example, through the master list process, DTF identified a school-based program involving two agencies that had plans to conduct two separate internal evaluations in subsequent years. DTF highlighted this with both agencies and recommended the establishment of an agreed program logic and combined evaluation approach.

DTF has previously shared the evaluation schedules of programs related to the Royal Commission into the Detention and Protection of Children with the National Indigenous Australians Agency to support a coordinated approach to evaluation between the Territory Government and the Commonwealth.

To encourage a One Territory Government approach, it is useful for agencies to have visibility of planned evaluations across government. Following endorsement of the 2022-23 evaluation schedule, DTF shared the proposed schedule with evaluation coordinators for each agency and the Auditor-General's Office.

At present, DTF does not have a system for monitoring and recording instances of improved coordination. The establishment of a simple monitoring spreadsheet will allow better data capture in preparation for the 2024-25 outcome evaluation (see recommendation 6).

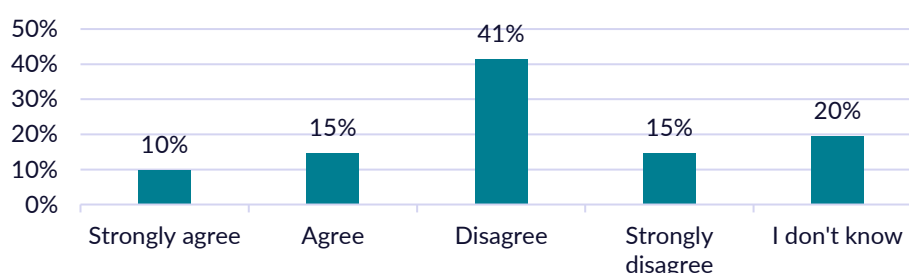


Employee survey results

Baseline results 2022

- Only 25% of respondents agreed or strongly agreed evaluation is coordinated across the Territory Government.

Graph 4: Staff perceptions - "Evaluation is coordinated across the Territory Government"



6.4.2. Outcome 2: Timely evidence for key decisions

Program stage	Indicator	Baseline	Target	Source
Outcome 2: Timely evidence for key decisions	Proportion of staff who agree evaluations are conducted in a timely manner to inform decision-making	22% of respondents agreed evaluations are conducted in a timely manner to inform decision-making	30% of survey respondents agree evaluations are conducted in a timely manner to inform decision-making by 2024	Employee survey
	Number and proportion of evaluations completed as per the evaluation schedule	13 of 56 (23%)	50% of evaluations completed as per the evaluation schedule by 2024	PEU records

The inaugural evaluation schedule was endorsed by BRS in July 2021, including 56 planned evaluations across 10 agencies. In June 2022, when the next annual evaluation schedule was submitted to BRS, only 13 of the evaluations planned for 2021-22 had been completed, 19 were underway and 24 had not started. From a whole of government perspective, a key lesson learnt was that agencies generally intend to evaluate programs but are sometimes distracted by other priorities arising throughout the year.

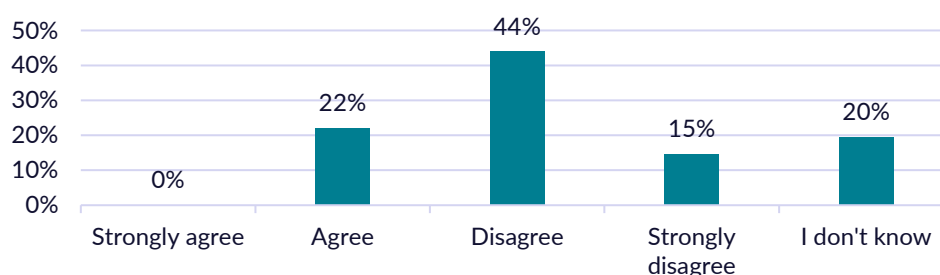


Employee survey results

Baseline results 2022

Staff perception data will also be used as a proxy for measuring this outcome over time. Only 22% of respondents agreed evaluations are conducted in a timely manner to inform decision-making.

Graph 5: Staff perceptions – "Evaluations are conducted in a timely manner to inform decision-making"



6.4.3. Outcome 3: Evaluation and data collection are proportional to the cost, risk and complexity of a program

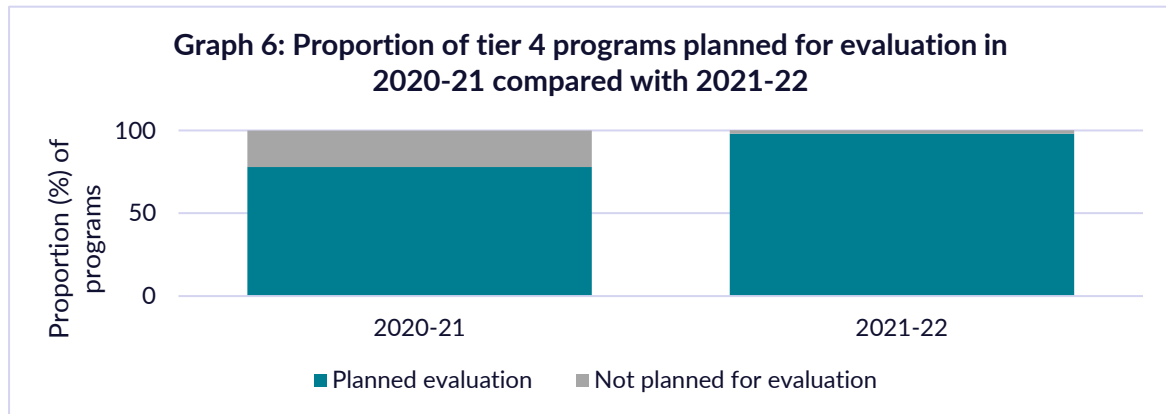
Program stage	Indicator	Baseline	Target	Source
Outcome 3: Evaluation and data collection are proportional to the cost, risk and complexity of a program	Proportion of total programs on the master list with planned evaluations by tiering	2020-21 master list Tier 4 – 78% (40 of 51) Tier 3 – 62% (43 of 69) Tier 2 – 48% (57 of 118) Tier 1 – 54% (29 of 54)	Higher tiered programs should be prioritised for outcome evaluations, lower tiered programs should undergo process evaluations or monitoring	Territory Government Master List
	Cost of external commissioned evaluations by tier	Established in 2023-24	Will be set once the baseline is established	Territory Government Quotations and Tenders Online
	Proportion of staff who agree current evaluation activity in their agency is proportional to the priority and risk profiles of policies and programs	39%	50%	Employee survey

The program master list requires agencies to categorise their programs against the tiers. A tier 4 program is one that may be high risk, has significant government funding, a strategic priority for government or lacks a current evidence base. For tier 4 programs, evaluation is mandatory. For tier 3 programs, evaluation is expected, and tier 2 and 1 programs it is at the agency's discretion.

To assess whether evaluations are proportionate to cost, risk and complexity of a program across government, DTF compared the proportion of planned evaluations against all programs by tier. DTF will work with agencies to improve their prioritisation approach, noting there is some inconsistency in the categorisation of tiering that also needs to be addressed. DTF monitors the number of programs against each tier over time.

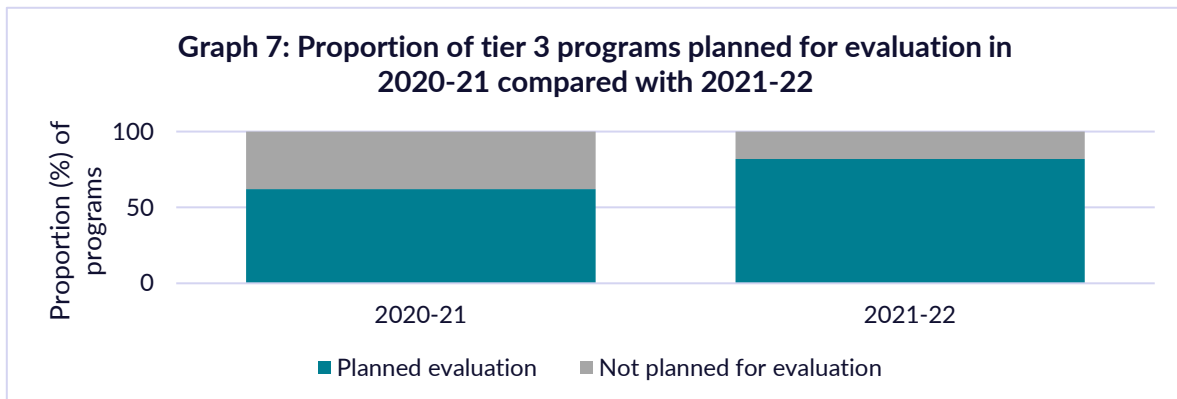
Mandatory evaluation – tier 4 programs

Graph 6 shows 78% of tier 4 programs were planned for evaluation in the baseline year 2020-21 compared with 98% (45 of 46) tier 4 programs in 2021-22.



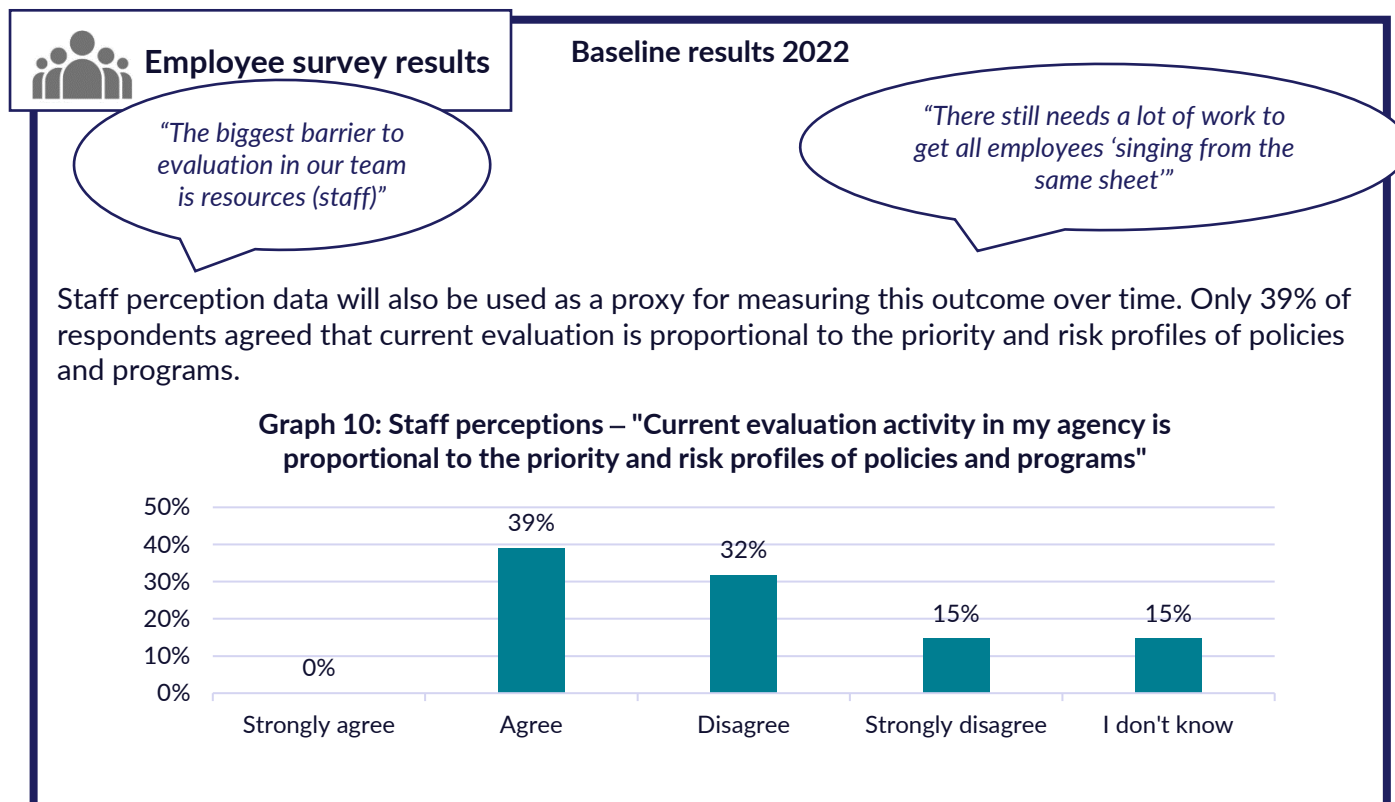
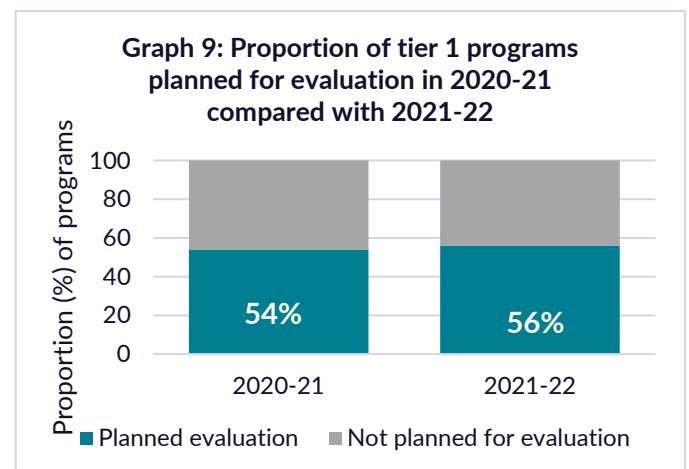
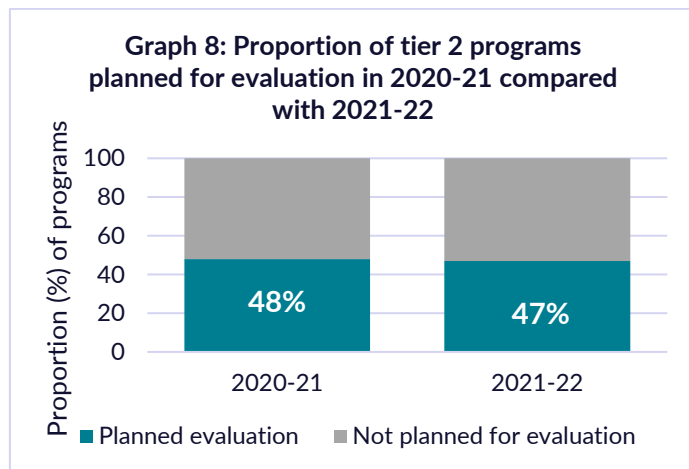
Expected evaluation – tier 3 programs

Graph 7 shows 62% tier 3 programs were planned for evaluation in 2020-21 compared to 82% tier 3 programs in 2021-22.



Agency's discretion – tier 1 and 2 programs

A lower proportion of tier 1 and 2 programs have evaluation planned, as is appropriate under the PEF.



Cost of external evaluations

In March 2022, an across government panel contract was established for professional services, including program evaluation. This provides a more streamlined procurement process where selected contractors have already broadly demonstrated their experience and capability in program evaluation.

The majority of externally commissioned evaluations will now be captured through this panel contract and will allow for central oversight of the cost of external evaluations by tiering. The 2024-25 PEF outcome evaluation will review total funding against tiers.

6.4.4. Outcome 4: Increased proportion of evaluations considered at the program planning stage

Program stage	Indicator	Baseline	Target	Source
Outcome 4: Increased proportion of evaluations consider at the program planning stage	Proportion of evaluation work plans submitted to PEU within 6 months of Cabinet approval	47%	100%	PEU records

Agencies' compliance with completing an evaluation overview as part of a Cabinet submission and submitting a full work plan within six months of approval is one data source for measuring outcome 4.

As DTF has taken a staged approach in rolling out the requirements of the PEF, the baseline data for this outcome will be captured following December 2022 when the first evaluation work plans are due to DTF.

6.4.5. Outcome 5: Increased evaluation maturity across the Territory Government

Program stage	Indicator	Baseline	Target	Source
Outcome 5: Increased evaluation maturity across the Territory Government, including: <ul style="list-style-type: none"> • awareness of the value of evaluation • evaluation capabilities • demand for evaluation and its use 	Proportion of staff aware of the PEF? This indicator is from the PEF – we will use the PECO membership as a proxy.	179	Remains steady or increase	PEU records
	Proportion of relevant staff who perceive evaluation as an opportunity to improve government services. This indicator is from the PEF – we will use number and proportion of survey respondents who perceive evaluation as an essential part of policy development and program management.	100%	100%	Stakeholder survey
	Staff perceptions of individual evaluation capability Proportion of survey respondents who feel confident in: <ul style="list-style-type: none"> • developing program logic • developing evaluation questions and performance measures their data analysis skills	Developing program logic – 83% Developing evaluation questions and performance measures – 80% their data analysis skills – 80%	Staff perceptions of individual evaluation capability remain steady or increases each survey	Stakeholder survey

Northern Territory Government Program Evaluation Framework: Process Evaluation Report

Program stage	Indicator	Baseline	Target	Source
	Staff perceptions of evaluation maturity at the team and agency level	Team: <ul style="list-style-type: none"> Beginning – 23% Developing – 43% Embedded – 26% Leading 9% Agency level: <ul style="list-style-type: none"> Beginning – 40% Developing – 54% Embedded – 3% Leading – 3% 	Staff perceptions of evaluation maturity at the team and agency level remain steady or increase each survey	Stakeholder survey
	Level of whole of government evaluation maturity. This indicator is from the PEF – will be assessed using the evaluation maturity matrix in the PEF.	Evaluation maturity not previously measured	Assessment of whole of government evaluation maturity increases with each PEF evaluation	Stakeholder survey/focus group/interview, agency-level evaluation maturity assessments

Evaluation maturity refers the extent to which an agency or team has the necessary capability, capacity and culture to embed evaluation into practice to help drive evidence base practice and continuous quality improvement.



Employee survey results

Baseline results 2022

"The PEF (and the work of the [DTF] PEU team more broadly) is driving improvements in processes"

Perceptions of individual capability:

The survey asked respondents to consider their individual evaluation capability in three key areas: developing program logics, developing evaluation questions and performance measures and their data analysis skills.

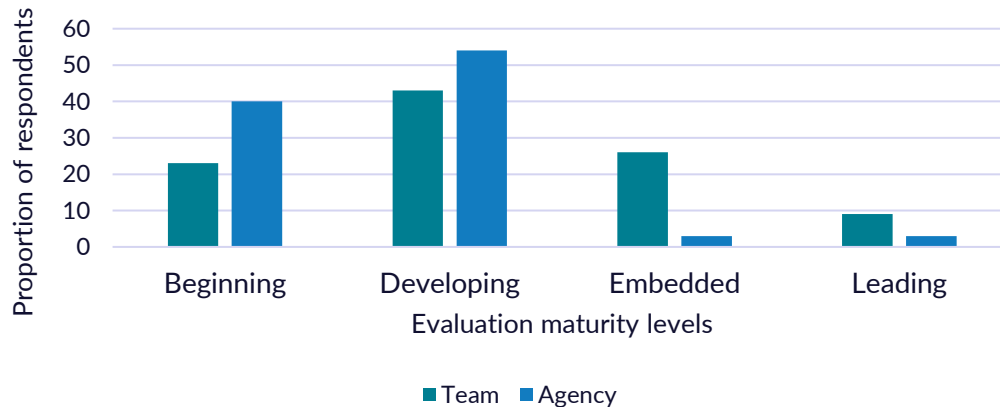
- 34 (83%) respondents feel confident in developing program logics
- 33 (80%) respondents feel confident in developing evaluation questions and performance measures
- 33 (80%) respondents feel confident in their data analysis skills
- 29 (71%) respondents felt confident across all three areas

Understanding the value of evaluation

- 100% of respondents perceive evaluation as an essential part of policy development and program management.

Staff perceptions of evaluation maturity

Graph 11: Staff perceptions of evaluation maturity within respondent's teams vs agency level



- In general, respondents rated their team's evaluation maturity higher than the rest of the agency (the agency bars are higher in the beginning and developing categories and lower in the embedded and leading categories).



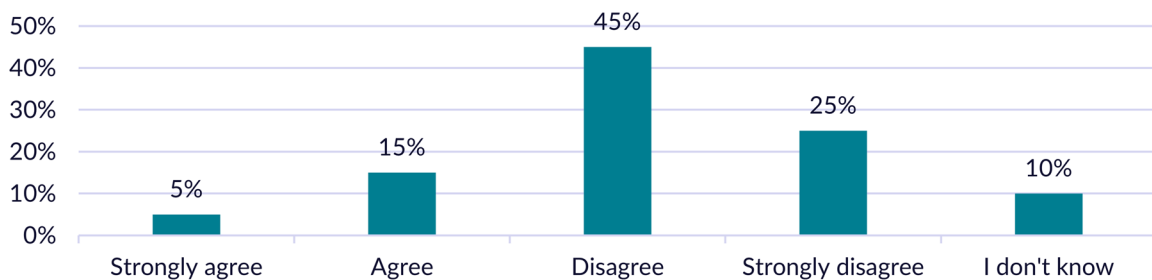
Employee survey results

Baseline results 2022

Management support for evaluation

- Only 20% of respondents agreed or strongly agreed that management within their agency supported and promoted evaluations

Graph 12: Staff perceptions - "Management in my agency supports and promotes the conduct of evaluation"



"There needs to be more emphasis from the top. I feel like there has been a lot of pushing up to get evaluation on the agenda and not a lot of genuine knowledge about quality evaluation and the PEF amongst senior leaders."

The PEF (Table 5) (reproduced at Appendix C) outlines four levels of evaluation maturity: beginning, developing, embedded and leading maturity. DTF has worked with the Office of the Commissioner for Public Employment to specifically include evaluation in policy officer job descriptions and in the revised Capability Framework.

Based on the evaluation maturity matrix and staff perception data, DTF has estimated the Territory Government evaluation maturity for each of the components in the maturity matrix as at September 2022.

Table 2: Assessment of evaluation maturity as at September 2022

Component	Estimated evaluation maturity	Rationale
Culture	Developing maturity Widespread awareness of the benefits of evaluation	Although the survey respondents overwhelmingly perceived evaluation as an integral part of policy development, respondents were concerned about the level of evaluation support from agency management.
Capacity	Developing maturity Targeted training and recruitment is used to develop staff skills. Formal evaluation policies and structures are in place.	Evaluation has been included in relevant policy officer job descriptions and the revised Capability Framework. Evaluation training is included in the across government training panel contract and evaluation training is regularly promoted. The PEF has integrated evaluation into the budget development process.
Planning	Beginning maturity Evaluation planning occurs for some programs, mainly after implementation. No, or very basic, evaluation plans.	The first evaluation work plans for programs approved in the 2022 Budget process are due to DTF in December 2022.
Strategy	Beginning maturity Programs with identified problems are prioritised.	The program master list has encouraged agencies to prioritise large and high risk programs. As most programs have not been designed for evaluation (and therefore don't have baseline data) there has been a focus on process evaluations, with few outcome or impact evaluations completed so far.
Conducting	Developing maturity Priority programs are evaluated.	Agencies intend to evaluate but sometimes get distracted by other priorities throughout the year.
Using	Beginning maturity Evaluation findings disseminated within the agency. Significant recommendations are implemented.	The annual evaluation memo to BRS helps to close the loop on evaluations so that agencies need to explain their response to the recommendations. More evaluations will need to be completed before this component of evaluation maturity can increase.

6.4.6. Outcome 6: Evaluations meeting quality standards

Program stage	Indicator	Baseline	Target	Source
Outcome 6: Evaluations meeting quality standards	Proportion of DTF summaries in the annual evaluation memo to BRS that note concerns with the evaluations	Of the 8 completed evaluations with a report on DTF file, DTF had concerns with the findings in 4 of the reports	Less than 50%	PEU records
	The proportion of evaluations submitted to the PEU meeting quality standards (see Appendix C). This indicator is from the PEF	Established in 2023-24	Will be set once the baseline is established	PEU records

In line with the program logic at Appendix A, evaluations meeting quality standards is a medium-term outcome that is expected in 2027-28. However, an early indication of evaluations meeting quality standards is whether DTF had concerns with the findings of the completed reports submitted to DTF as part of the annual evaluation summary to BRS.

In 2021-22, of the eight completed evaluations with a report on DTF file, DTF had concerns with the findings of four reports, including:

- methodological concerns (for example, subjective weightings that were not well explained)
- conclusions that are not fully consistent with the evidence (for example, concluding a program has been successful at achieving an outcome without taking into consideration the other programs in the community aimed at improving that outcome)
- insufficient detail (for example, a one-page case study in an annual report that mentions how the program has been adjusted over time to respond to lessons learnt but does not consider whether alternative investments could have been more effective).

In addition, DTF has drafted evaluation standards based around the 10 best practice evaluation principles in the PEF. From 2023, evaluations will be assessed against this checklist.

6.4.7. Outcome 7: Evaluation findings used and recommendations implemented

Program stage	Indicator	Baseline	Target	Source
Outcome 7: Evaluation findings used and recommendations implemented	The proportion of accepted recommendations from evaluation reports implemented. This indicator is from the PEF	Established in 2023-24	Will be set once the baseline is established	Agency updates on the evaluation schedule
	Increased number of programs that use the results from previous evaluations to inform design and implementation. This indicator is from the PEF, measured by: <ul style="list-style-type: none"> proportion of relevant Budget Cabinet submissions citing previous evaluations DTF comments informed by evaluations, where appropriate 	<p>In 2022, 58% of relevant Budget Cabinet submissions cited previous evaluations</p> <p>In 2022, DTF comments on all relevant Budget Cabinet submissions were informed by evaluations</p>	<p>100% of relevant submissions</p> <p>DTF comments informed by evaluations, where appropriate</p>	Budget Cabinet submissions and comments (Content Manager)
	Staff perceptions that evaluation findings and recommendations are communicated widely within their agency	In 2022, 15% of respondents agree evaluation findings and recommendations are communicated widely within their agency	>80% of survey respondents agree evaluation findings and recommendations are communicated widely within their agency in 2027-28	Stakeholder survey
	Staff perceptions that evaluation findings and recommendations are being used to improve service delivery	In 2022, 43% of respondents agreed or strongly agreed evaluation findings and recommendations are being used to improve service delivery	>80% of survey respondents agree or strongly agree evaluation findings and recommendations are being used to improve service delivery in 2027-28	

There are several short-term outcomes that need to be in place before there is likely to be traction in terms of evaluation findings being used and recommendations implemented. This is therefore a medium term outcome that is expected to be observable by 2027-28.

Proportion of Budget Cabinet submissions citing previous evaluations

In 2022, 58% of the Cabinet submissions reviewed by the PEU referenced previous evaluations, reviews or data to support their proposal. The quality of the evidence base has not been assessed as part of this process evaluation but the data is available on the Territory Government record system (Content Manager) for later assessment during the outcome evaluation.

DTF comments informed by evaluations, where appropriate

DTF reviews the existing evidence base of all Budget Cabinet submissions where evaluation requirements apply to inform the DTF comment. DTF comments on relevant Cabinet submission in 2022 referenced relevant reviews, highlighted evidence gaps or recommended evaluation approaches to support the Cabinet decision-making process.

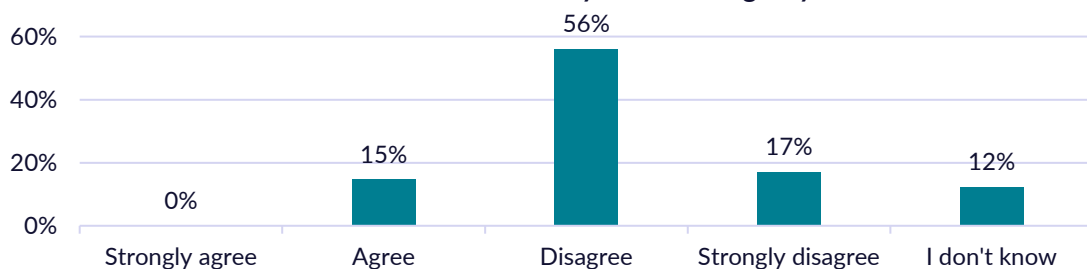


Employee survey results

Baseline results 2022

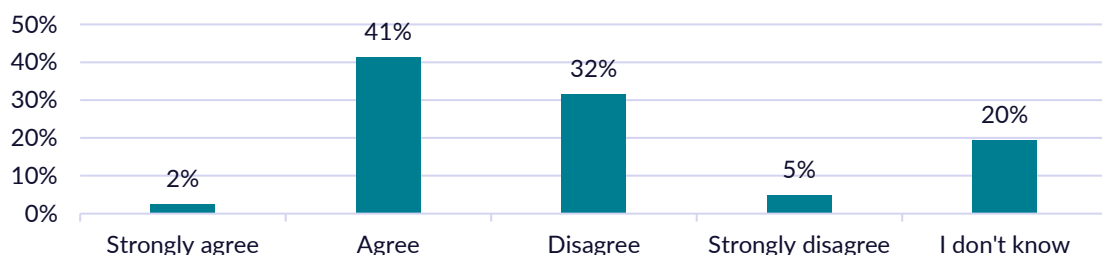
- Only 15% of respondents agree evaluation findings and recommendations are communicated widely within their agency.

Graph 2: Staff perceptions – "evaluation findings and recommendations are communicated widely within the agency"



- Only 43% of respondents agreed or strongly agreed evaluation findings and recommendations are being used to improve service delivery.

Graph 3: Staff perceptions – "evaluation findings and recommendations are being used to improve service delivery"



6.5. Are there any adjustments to the implementation approach that need to be made?

Respondents provided the following suggestions on how the PEF and toolkit could be improved:

- **improve templates and examples:**
 - include templates for the three types of evaluations in the framework
 - include terms of reference and evaluation report templates
 - provide online forms that could assist with individual and systems level reporting and capability development
 - interactive template for co-constructing program logics
 - include de-identified examples or copies of past documentation following completion of evaluations (via request if not able to publish)
 - provide examples of good and bad evaluation frameworks.
- **improve awareness of the PEF:**
 - provide more presentations and opportunities to ask questions
 - links to the framework and toolkit should be on Territory Government intranet
- **improve the content:**
 - more information around systems thinking and behavioural science
 - There is still ambiguity about what constitutes a 'program'
 - program logic template does not include activities and need to clarify indicators and measures

Respondents suggested DTF could improve evaluation support to agencies through:

- more presentations on the framework (noting these will be required on a regular basis to allow for staff turnover)
- workshops using real examples
- proactively engaging rather than following up post-project
- improved clarity around what is required to be submitted to DTF, what is optional and what is unnecessary
- clearer separation of informal information sharing with the community of practice from more formal communication
- feedback on draft Territory Government evaluations
- assisting agency staff to establish program evaluation frameworks.¹⁵

¹⁵ DTF does provide feedback on draft Territory Government evaluations and evaluation frameworks. These responses indicate DTF needs to clarify its role and what support is available to agencies through further communication, including updating the website.

7. Recommendations

Improving implementation

1. Strengthen the toolkit

- DTF to develop templates and standards for the three types of evaluations (process, outcome and impact)
- DTF to provide examples of good quality evaluation work plans
- DTF to liaise with DCDD on the feasibility of creating online forms that could assist with reporting and interactive program logic templates.

2. Improve awareness of the PEF and the role of DTF

- DTF to record a presentation about the PEF for new staff
- DTF to offer additional face-to-face or online presentations and opportunities for staff to ask questions
- DTF to investigate linking the PEF on the Territory Government intranet.

3. Actively follow up

- DTF to proactively remind agencies of work plan and reporting due dates
- DTF to proactively contact relevant program and policy areas where the Territory Government has publicly committed to evaluating strategies and programs.

4. Training and professional development opportunities

- DTF to promote available training on:
 - developing program logics, evaluation questions and evaluation criteria
 - data analysis
 - systems theory
 - process evaluation and implementation science
 - commissioning evaluations.

Strengthening outcome measures

5. Measuring evaluation quality standards

- DTF to finalise the evaluation standards in collaboration with the PECoP and publish the standards on the PEU webpage
- DTF to measure evaluation reports against standards to provide feedback to agencies and to inform DTF Cabinet submission comments.

6. Measuring improved coordination

- DTF to develop a simple monitoring spreadsheet to capture instances where the program master list and evaluation schedule have been used to identified duplication and facilitated coordination across agencies.

8. Conclusion

The PEF consisted of six key reforms, all of which have been implemented. The guidance tools are widely used and are considered user-friendly, however survey respondents made several suggestions to improve the implementation approach.

Since 2019, DTF has delivered a total of 40 evaluation-related presentations to 1,211 attendees. The PECoP membership has steadily increased – as at April 2023, the PECoP consisted of 179 members.

Since 2020, agencies have requested DTF evaluation advice and feedback on 44 programs across a range of government priorities. DTF has responded to each of these requests, providing evaluation advice on programs including domestic and family violence, tourism, economic stimulus, Aboriginal justice, youth justice, health, education and sustainability.

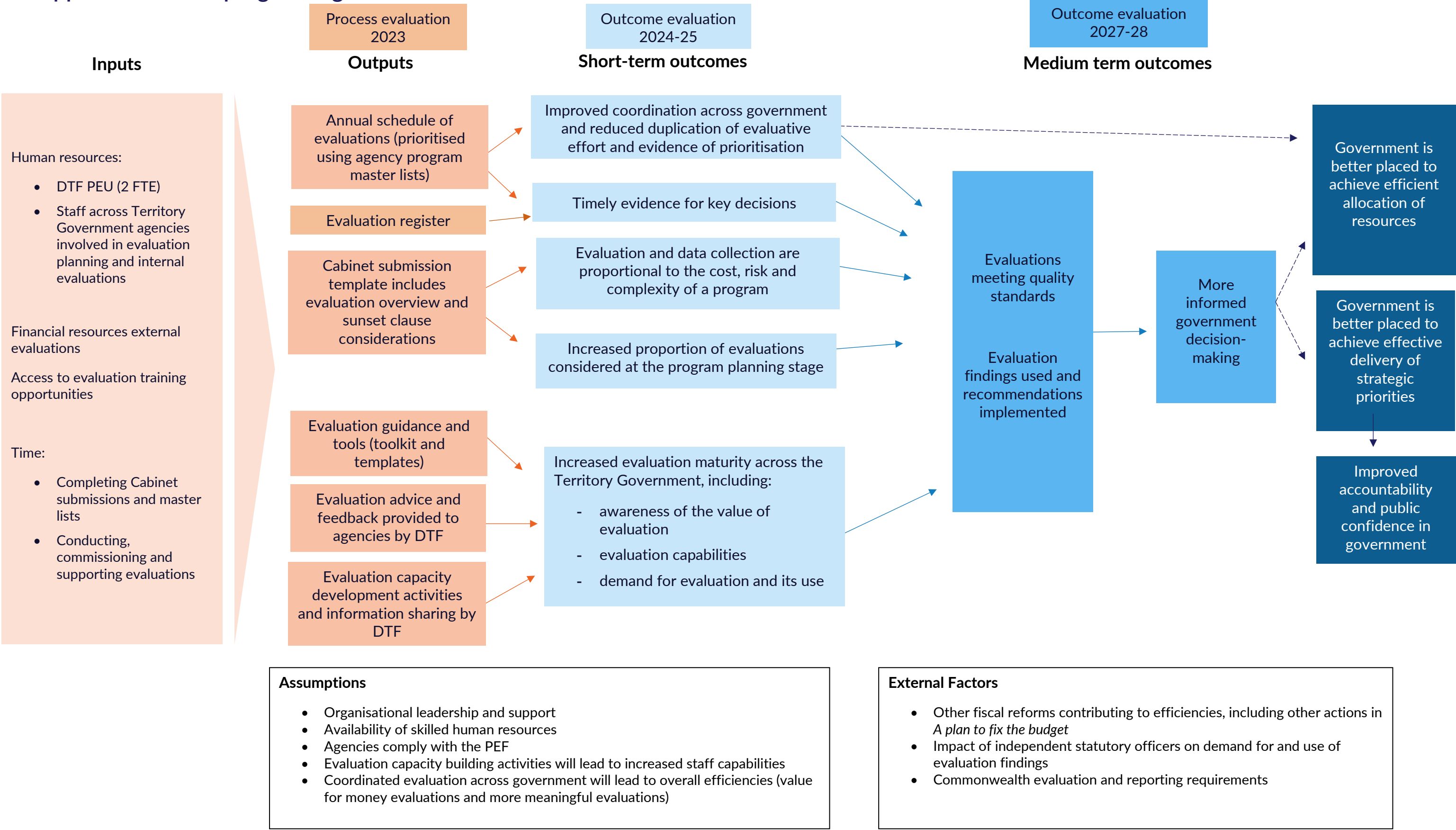
In contrast to the previous evaluation approach established under the 1993 Treasurer's Direction Evaluation and Review, there is now a central agency providing guidance on how and why to carry out evaluations, evaluation training is available and promoted, and Cabinet submissions on major policy issues include evaluation overviews.

It is encouraging that agencies are using the templates and engaging with DTF for evaluation advice and feedback. The initial success of the PEF implementation may be linked to embedding the PEU within the Budget Development team at DTF and integrating evaluation into the budget development process – this creates a clear incentive for agencies to plan for evaluations as they are developing the Cabinet submissions.

DTF's follow up after the budget development process, such as reviewing evaluation work plans and reports and summarising the results (including highlighting concerns with the evaluations) for BRS appears to motivate agencies to undertake evaluations and ensure the methodology and scope are appropriate. However, the low baseline for the key outcomes measurable at this stage (such as coordinated evaluations across government, evaluations meeting quality standards and evaluation findings being used and recommendations implemented) indicate there is still much work to be done.

Ironically, there are not many publicly available evaluations of evaluation reforms. Therefore, it is not yet clear whether the PEF will be able to achieve the intended outcomes even if it is well implemented. Feedback from other jurisdictions has indicated that it can take substantial time to create an evaluation culture (around 10 years). Given the lack of evidence in this area, it will be important to ensure the planned outcome evaluations for the PEF are completed as scheduled (2024-25 and 2027-28).

9. Appendix A – PEF program logic



10. Appendix B – Data matrix

Key evaluation questions	Program stage	Indicator	Baseline	Target	Data source
Outputs – Process evaluation 2023					
<ul style="list-style-type: none"> To what extent has the PEF been implemented? Are the guidance tools and templates used by agencies? How user-friendly and appropriate have the guidance tools been for agencies? How ready is the PEF for an outcome evaluation? Are there any adjustments to the PEF implementation approach that need to be made? 	Output 1: Annual schedule of evaluations coordinated by DTF (prioritised using agency program master lists)	– Each agency has a program master list to prioritise program evaluations	Program master lists and evaluation schedules did not previously exist	Agencies submit a program master list to DTF each year	BRS memo PEU records
		– Number of programs on the whole of government program master list		Number of programs on the whole of government program master list remains steady or increases each year	
		– A whole of government annual schedule is submitted and endorsed by BR		Evaluation schedule submitted and endorsed each year	
	Output 2: Evaluation register	– Evaluation and monitoring reports added to the evaluation register as they are completed (either on the DTF website or on the internal register)	Evaluation register did not previously exist	Number of evaluation and monitoring reports on the DTF website remains steady or increases each year	PEU records
		– Proportion of completed evaluations with a copy the report on DTF file in the annual evaluation summary to BRS	Evaluation reports were not previously provided to DTF	Proportion of completed evaluations with a copy the report on DTF file in the annual evaluation summary to BRS remains steady or increases each year	BRS memo
	Output 3: Cabinet submissions template includes evaluation overview and sunset clause considerations	– Cabinet submission template updated with evaluation overview template and sunset clause considerations	Original Cabinet submission template	Cabinet submission template updated	Cabinet submission template on the Territory Government intranet
	Output 4: Evaluation guidance tools (toolkit and template)	– Whole of government program evaluation toolkit and templates published on the DTF website	Toolkit and templates did not previously exist	Toolkit and templates published	DTF website
		– Stakeholder feedback on how user-friendly and appropriate the toolkit and templates are	Toolkit and templates did not previously exist	>80% of survey respondents agree or strongly agree that the toolkit and templates are user friendly	Stakeholder survey Written feedback Territory Government agency-level evaluation strategy report
		– Toolkit and template are adapted over time in response to stakeholder feedback		Toolkit and template are adapted over time in response to stakeholder feedback	
		– Number of DTF PEU webpage unique page views	Webpage and PEF did not previously exist	Number of DTF PEU webpage unique page views remains steady or increases each year	DTF website analytics

Key evaluation questions	Program stage	Indicator	Baseline	Target	Data source
		– Unique downloads of the PEF		Number of unique downloads of the PEF remains steady or increases each year	Stakeholder survey
		– Percentage of survey respondents who have referred to the PEF and toolkit in the last year		>80% of survey respondents have referred to the PEF and toolkit in the last year	
	Output 5: Evaluation advice and feedback to agencies by DTF	– Number of programs that DTF has provided evaluation advice or feedback on	Previously, there was not a central agency providing evaluation advice and feedback to agencies	Number of programs that DTF has provided evaluation advice or feedback on remains steady or increases each year	PEU records
		– Number and proportion of survey respondents who agree evaluation guidance or feedback had been useful and relevant to their team	Previously, there was not a central agency providing evaluation advice and feedback to agencies	Number and proportion of survey respondents who agree evaluation guidance or feedback had been useful and relevant to their team remains steady or increases each year	Stakeholder survey Written feedback
	Output 6: Evaluation capacity development activities and information sharing by DTF	– Number of evaluation capacity development activities coordinated or promoted by DTF. This indicator is from the PEF	Previously, there was not a central agency sharing evaluation capacity and development activities	Number of evaluation capacity development activities coordinated or promoted by DTF remains steady or increases each year	PEU records Territory Government agency-level evaluation strategy report
		– Number of DTF evaluation presentations and number of attendees		Number of DTF presentations and number of attendees remains steady or increases each year	
		– Number and proportion of survey respondents who have attended evaluation training promoted by DTF		Number and proportion of survey respondents who have attended evaluation training promoted by DTF remains steady or increases each year	Stakeholder survey
		– Percentage of survey respondents who agreed DTF shares relevant information in PEU emails and seminars		>80% of survey respondents agreed DTF shares relevant information	

Outcomes – Outcome evaluation in 2024-25					
<p>Update on the key evaluation questions from the process evaluation plus:</p> <ul style="list-style-type: none"> To what extent are evaluations systematically used to inform decision-making across the Territory Government? To what extent are evaluations meeting evaluation standards? To what extent has there been an improvement in coordination across government and reduced duplication of evaluative effort? To what extent are evaluations scheduled to provide timely evidence for key decisions? To what extent has there been an increase in evaluation maturity across the Territory Government? Have there been any unintended consequences, positive or negative as a result of the introduction of the NT Program Evaluation Framework? 	Outcome 1: Improved coordination across government and reduced duplication of evaluative effort	– Proportion of staff who agree evaluation is coordinated across the Territory Government	25%	30% of survey respondents agree evaluation is coordinated across the Territory Government by 2024-25	Employee survey
	Outcome 2: Timely evidence for key decisions	– Proportion of staff who agree evaluations are conducted in a timely manner to inform decision-making	22% of respondents agreed evaluations are conducted in a timely manner to inform decision-making.	30% of survey respondents agree evaluations are conducted in a timely manner to inform decision-making by 2024-25	Employee survey
		– Number and proportion of evaluations completed as per the evaluation schedule	13 of 56 (23%)	50% of evaluations completed as per the evaluation schedule by 2024-25	PEU records
	Outcome 3: Evaluation and data collection are proportional to the cost, risk, and complexity of a program	– Proportion of total programs on the master list with planned evaluations by tiering	2020-21 master list Tier 4 – 78% (40 of 51) Tier 3 – 62% (43 of 69) Tier 2 – 48% (57 of 118) Tier 1 – 54% (29 of 54)	Higher tiered programs should be prioritised for outcome evaluations, lower tiered programs should undergo process evaluations or monitoring	Territory Government Master List
		– Cost of external commissioned evaluations by tier	Established in 2023-24	Will be set once the baseline is established	Territory Government Quotations and Tenders Online
		– Proportion of staff who agree current evaluation activity in their agency is proportional to the priority and risk profiles of policies and programs	39%	50%	
	Outcome 4: Increased proportion of evaluations consider at the program planning stage	– Proportion of evaluation work plans submitted to PEU within 6 months of Cabinet approval	47%	100%	PEU records
	Outcome 5: Increased evaluation maturity across the Territory Government, including: – awareness of the value of evaluation – evaluation capabilities – demand for evaluation and its use	– Proportion of staff aware of the PEF. This indicator is from the PEF – we will use the PECoP membership as a proxy.	179	Remains steady or increase	PEU records
		– Proportion of relevant staff who perceive evaluation as an opportunity to improve government services. This indicator is from the PEF – we will use number and proportion of survey respondents who perceive evaluation as an essential part of policy development and program management.	100%	100%	Stakeholder survey

		<ul style="list-style-type: none"> – Staff perceptions of individual evaluation capability – Proportion of survey respondents who feel confident in: <ul style="list-style-type: none"> ○ developing program logic ○ developing evaluation questions and performance measures ○ their data analysis skills 	1. Developing program logic – 83% 2. developing evaluation questions and performance measures – 80% 3. their data analysis skills – 80%	Staff perceptions of individual evaluation capability remain steady or increases each survey	Stakeholder survey
		<ul style="list-style-type: none"> – Staff perceptions of evaluation maturity at the team and agency level 	Team: <ul style="list-style-type: none"> • beginning – 23% • developing – 43% • embedded – 26% • leading 9% Agency level: <ul style="list-style-type: none"> • beginning – 40% • developing – 54% • embedded – 3% • leading – 3% 	Staff perceptions of evaluation maturity at the team and agency level remain steady or increase each survey	
		<ul style="list-style-type: none"> – Level of whole of government evaluation maturity. This indicator is from the PEF – will be assessed using the evaluation maturity matrix in the PEF. 	Evaluation maturity not previously measured	Assessment of whole of government evaluation maturity increases with each PEF evaluation	Stakeholder survey/focus group/interview, agency-level evaluation maturity assessments

Outcomes – Outcome evaluation in 2027-28					
<p>Update on the key evaluation questions from the process evaluation and the 2023-24 outcome evaluation plus:</p> <ul style="list-style-type: none"> To what extent are evaluations meeting evaluation standards? To what extent are evaluations systematically used to inform decision-making across the Territory Government? 	Outcome 6: Evaluations meeting quality standards	– Proportion of DTF summaries in the annual evaluation memo to BRS that note concerns with the evaluations	Of the 8 completed evaluations with a report on DTF file, DTF had concerns with the findings in 4 of the reports	Less than 50%	PEU records
		– The proportion of evaluations submitted to the PEU meeting quality standards (see appendix C). This indicator is from the PEF	Established in 2024-25	Will be set once the baseline is established	PEU records
	Outcome 7: Evaluation findings used and recommendations implemented	– The proportion of accepted recommendations from evaluation reports implemented. This indicator is from the PEF	Established in 2024-25	Will be set once the baseline is established	Agency updates on the evaluation schedule
		Increased number of programs that use the results from previous evaluations to inform design and implementation. This indicator is from the PEF, measured by: <ul style="list-style-type: none"> – proportion of relevant Budget Cabinet submissions citing previous evaluations – DTF comments informed by evaluations, where appropriate 	<p>In 2022, 58% of relevant Budget Cabinet submissions cited previous evaluations</p> <p>In 2022, DTF comments on relevant Budget Cabinet submissions were informed by evaluations</p>	<p>100% of relevant submissions</p> <p>DTF comments informed by evaluations, where appropriate</p>	Budget Cabinet submissions and comments (Content Manager)
		– Staff perceptions that evaluation findings and recommendations are communicated widely within their agency	In 2022, 15% of respondents agree evaluation findings and recommendations are communicated widely within their agency	>80% of survey respondents agree evaluation findings and recommendations are communicated widely within their agency in 2027-28	Stakeholder survey
		– Staff perceptions that evaluation findings and recommendations are being used to improve service delivery	In 2022, 43% of respondents agreed or strongly agreed evaluation findings and recommendations are being used to improve service delivery	>80% of survey respondents agree or strongly agree evaluation findings and recommendations are being used to improve service delivery in 2027-28	

11. Appendix C – Evaluation maturity¹⁶

	Beginning maturity	Developing maturity	Embedded maturity	Leading maturity
Culture	Evaluation awareness is low and is as a response to identified problems.	Widespread awareness of the benefits of evaluation.	Evaluation perceived as an integral component of sound performance management.	Demonstrated commitment to continuous learning and improvement across government.
Capacity	Evaluation skills are limited. No formal evaluation procedures and structures are in place.	Targeted training and recruitment is used to develop staff skills. Formal evaluation policies and structures are in place.	General evaluation skills are widespread. Relevant staff have higher order skills and experience, which is leveraged by the agency. Evaluation systems, structures and procedures are robust, integrated and of proven effectiveness.	The government is recognised for its evaluation expertise and innovative procedures and systems.
Planning	Evaluation planning occurs for some programs, mainly after implementation. No, or very basic, evaluation strategy.	Programs have well defined objectives and performance indicators as a baseline for future evaluation. Evaluation activity is coordinated and an evaluation strategy is in place.	Evaluation planning is an integral component of policy development.	Evaluation plans are in place for most programs.
Strategy	Programs with identified problems are prioritised.	Large and high risk programs are prioritised.	Guidelines for prioritising and scaling evaluation activity are used.	Evaluations are prioritised and scaled to provide the most useful evidence for the least cost.
Conducting	Evaluation occurs but is infrequent and ad hoc.	Priority programs are evaluated.	Evaluation is widespread and conforms to this framework.	Evaluation is almost universal and best practice.
Using	Evaluation findings disseminated within the agency. Significant recommendations are implemented.	Evaluation findings routinely inform decision making and are often disseminated outside the agency.	Evaluation findings are widely disseminated and used to improve performance.	Findings are used to optimise service delivery and improve accountability and transparency.

¹⁶ Reproduced from page 21 of the PEF – adapted from the ACT Government Evaluation Policy 2010.