Submission No 8

EXAMINATION OF THE AUDITOR-GENERAL'S PERFORMANCE AUDIT REPORTS OCTOBER 2016 – MAY 2017

Organisation: Department of Justice

Date Received: 5 February 2018

Office of the Secretary
Justice Precinct Offices
Locked Bag 5111, Parramatta NSW 2124
www.justice.nsw.gov.au

Mr Bjarne Nordin
Public Accounts Committee
GPO Box 12
SYDNEY NSW 2001

Your Ref: D17/35209
Our Ref:D18/22311/DJ

Dear Mr Nordin

Auditor-General's performance audit report: Implementation of the NSW Government's program evaluation initiative

I am pleased to advise that the Department of Justice ('the department') accepted and implemented the following recommendations identified in the above report, namely:

- an evaluation centre of excellence
- a Master List of programs
- a process for cluster wide program identification based on agency and cluster strategic planning process which aligns to State priorities and objective processes to prioritise cluster agency programs

The department has partially implemented the following recommendation:

- an objective process to prioritise cluster agency programs across the whole cluster, taking into account the capability and capacity to conduct evaluations.

Attachment A details the department's actions against each of the recommendations.

The department has also been working to enhance its evaluative capability beyond these recommendations. An Evaluation Framework was developed and distributed in 2017 (Attachment B). Other initiatives, including an evaluation hub, strengthened evaluation governance, and evaluation workshops, will be rolled out in 2018. The department also participates in the NSW Government Evaluation Working Group, to share information with evaluation experts from other agencies.

Should the Committee require additional information, please do not hesitate to contact [redacted]

Yours sincerely

Andrew Cappie-Wood
Secretary
- 5 FEB 2018

<table>
<tr><th>Recommendation</th><th>Accepted or rejected</th><th>Actions to be taken</th><th>Due date</th><th>Status (completed, on-track, delayed) and Comment</th><th>Responsibility</th></tr>
<tr><td>1. Have a centre of excellence for evaluation</td><td>Accepted</td><td>Ongoing review of resourcing</td><td>July 2018</td><td>Completed.

The NSW Bureau of Crime Statistics and Research (BOCSAR) is recognised as a centre of excellence in evaluation for the Department of Justice in the Auditor General's Report. Evaluation officers are also located in the Corrective Services NSW and Juvenile Justice divisions. All evaluation officers work separately to program areas.

The department's Performance and Assurance Group will continue to monitor evaluation activity and requirements to determine if evaluation resources remain sufficient.</td><td>BOCSAR

Performance and Assurance Group</td></tr>
<tr><td>2. Develop a master list of eligible cluster agency programs with tier ranking and linkage to NSW Government priorities</td><td>Accepted</td><td>Ongoing review</td><td>Quarterly</td><td>Completed.

The department has further refined the master list to incorporate cluster programs that are eligible for evaluation. Programs on the master list are derived from strategic planning processes and documented in Corporate and Division Business Plans. The Performance and Assurance Group also review Cabinet documents, budget papers, annual reports, media releases, and internal briefings to identify other potential inclusions, in conjunction with program areas.

The format of the Justice master list is based on the 2017-18 Treasury Evaluation Schedule template. It can be filtered by numerous variables, including tier ranking, NSW Government priority, stage(s) of justice, internal or external delivery, findings, lessons learned, and funding. This ensures that significant programs can be prioritised for assessment of effectiveness and value for money.

The master list is updated as new information is received. The Performance and Assurance Group will undertake a full review of the master list each year, prior to the start of the annual evaluation schedule process.</td><td>Performance and Assurance Group</td></tr>
<tr><td>3. Ensure that agency and cluster strategic planning processes align programs and program evaluations to NSW Government priorities.</td><td>Accepted</td><td>Ongoing</td><td>na</td><td>Completed.

The bulk of justice programs undergoing evaluation in 2017-18 are related to government priorities including reducing violent crime, reducing reoffending and reducing domestic violence.

<table>
<tr><th colspan="2">Government priorities</th><th>2017-18 evaluations*</th></tr>
<tr><td rowspan="3">Government priorities – Justice led</td><td>Reducing violent crime</td><td>2</td></tr>
<tr><td>Reducing adult reoffending</td><td>7</td></tr>
<tr><td>Reducing domestic violence</td><td>11</td></tr>
<tr><td>Government priorities – Justice contributing</td><td>Improving government services</td><td>5</td></tr>
<tr><td>Departmental priorities</td><td></td><td>10</td></tr>
<tr><td></td><td>Total</td><td>35</td></tr>
</table>
*Depicts most relevant priority only. Some programs being evaluated contribute to more than one priority.</td><td>Performance and Assurance Group</td></tr>
</table>

4. Programs are objectively prioritised across the cluster for evaluation and inclusion in evaluation schedules, taking account of the department's capacity and capability to conduct evaluation	Accepted	Governance mechanisms to be agreed and fully implemented	July 2018	On track. The department is establishing governance mechanisms, including a reference group, to set the forward agenda for evaluation, provide recommendations to Executive on implementation and quality assurance of evaluation activity, and consideration of cluster capacity and capability to conduct evaluations. Governance will be supported by resources, including analysis tools, that collate, organise and disseminate evaluation findings for action throughout the cluster. Evaluation capability will be enhanced via workshops conducted by the department's centre of excellence, BOCSAR. These will be progressively rolled out through 2018.	Performance and Assurance Group

NSW GOVERNMENT | Justice

Program Evaluation

THIS PAGE INTENTIONALLY LEFT BLANK

Table of Contents

1. Introduction **4**

Policy context 4

Scope 4

Framework design 5

2. Key concepts **5**

‘Program’ 6

‘Program evaluation’ 6

2. Principles of program evaluation **11**

3. When to evaluate **14**

4. Setting priorities for evaluation **15**

Evaluation Schedule 15

5. Data considerations **16**

6. Disseminating evaluation findings **16**

7. Governance, roles and responsibilities **17**

References **18**

Appendix 1: Glossary **19**

Appendix 2: Program Logic **21**

Appendix 3: Evaluation plan **22**

Appendix 4: Program scale & evaluation **24**

Appendix 5: Evaluation Report - briefing **25**

1. Introduction

Our department delivers a diverse range of programs to support people and communities across NSW. It is important to know whether or not our programs are delivering what we thought they may – e.g. increasing accessibility of the justice system, reducing reoffending, or enhancing support to vulnerable citizens.

Evaluation is a way that we can assess our programs. This Evaluation Framework[1] documents the principles and processes to guide rigorous and transparent program evaluations, to inform decision-making, planning and practice within our department.

Policy context

The Evaluation Framework has been developed in response to the *NSW Government Program Evaluation Guidelines* and the *Circular C2016-01 Program Evaluation*, issued by the Department of Premier and Cabinet. These documents reflect a whole-of-government commitment to return quality services through evidence-based policy and decision making.

This Framework is also informed by the NSW Auditor-General's Report *Implementation of the NSW Government's program evaluation initiative*.[2] The Auditor-General made several recommendations to strengthen evaluation practice in our department, including ensuring that strategic planning processes align program evaluations to NSW Government priorities, objectively prioritising programs across the cluster for evaluation, and creating centres of excellence for evaluation.[3]

This Framework is being delivered as Financial Management Transformation (FMT) is being rolled out across NSW Government. FMT is expected to drive better outcomes by introducing a strong focus on results and performance of Government expenditure. The Evaluation Framework complements FMT's focus on measuring program performance.

Scope

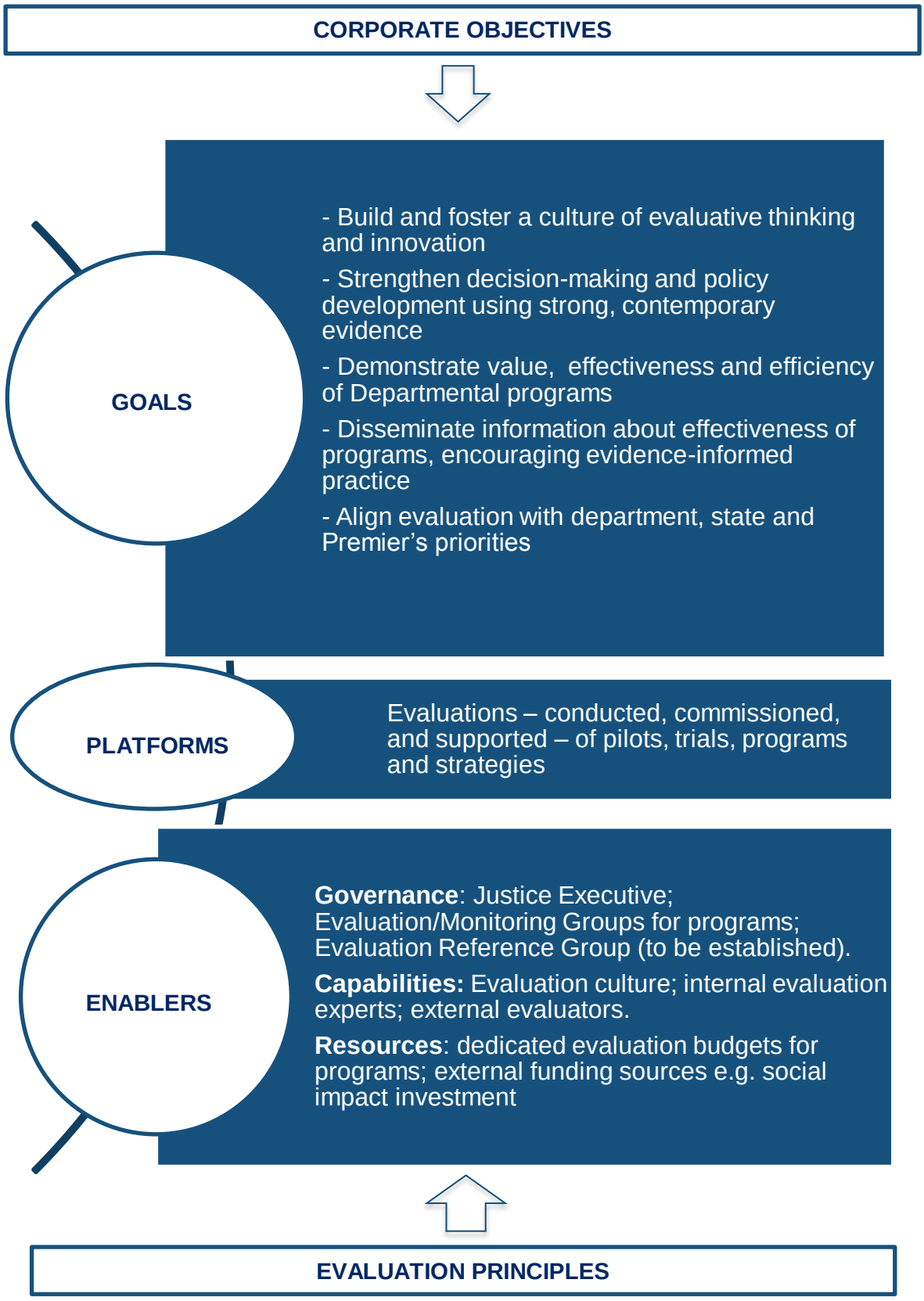
This Framework applies to all program evaluation activity in the department.[4] This includes evaluation activity related to improving service delivery, crime prevention, reoffending and treatment initiatives, as well as innovative pilots designed to address long standing issues. The Framework applies to evaluation activity undertaken internally or commissioned by the department and conducted externally.

---

[1] To ensure currency, the Framework will be reviewed 12 months from its date of effect.
[2] New South Wales Auditor General, Implementation of the NSW Government's program evaluation initiative,
[3] Ibid.
[4] Program evaluation is only one element of evaluation. We also conduct policy evaluations, for example, when Governments pass laws changing the penalties for various offences.

Framework design

A high level overview of the Framework is provided below.



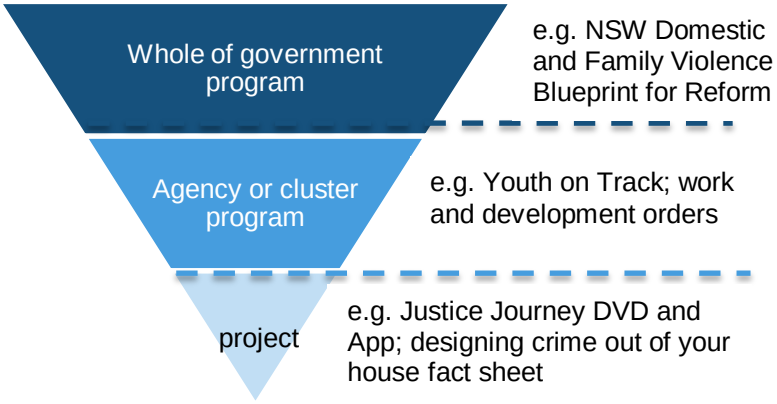
2. Key concepts

‘Program’

For the purposes of this Framework, ‘program’ is broadly defined as a set of activities managed together over a sustained period of time that aim to deliver an outcome/outcomes for a client or client group.[5] Program is sometimes used interchangeably with policy, project, intervention, initiative, or strategy.

Programs vary in size and nature. They can be large, representing a whole-of-government initiative involving multiple agencies, such as the *NSW Domestic and Family Violence Blueprint for Reform 2016-2021, Countering Violent Extremism*, and *Keep Them Safe*. Often these large programs are broken down into smaller projects for evaluation. Programs may also be smaller community-level projects, designed to address a particular issue within a local area, for example funding to reduce the risk of graffiti near a community centre.

Figure 1. Justice programs at different levels



‘Program evaluation’

Program evaluation is defined as ‘a systematic and objective process to make judgments about the merit or worth of one or more programs, usually in relation to their effectiveness, efficiency and appropriateness.’[6]

Evaluation is a core component of improving the quality of evidence which we use to make decisions. It helps us to assess what we do. It lets us know what is working, in which context and for whom. Evaluation may also alert us to where we could make changes to our programs to deliver better outcomes. For these reasons, where available, evaluation findings should be used to inform decisions to continue, expand, amend or discontinue programs.

Evaluation differs from other types of research, monitoring and assessment that are routinely carried out within our department, such as program reviews, policy analysis, internal reporting, and audits. Some of these are similar to evaluation in terms of data collection and analysis; however their purpose and level of analysis may differ significantly. Further information about each of these can be found in the NSW Program Evaluation toolkit.[7]

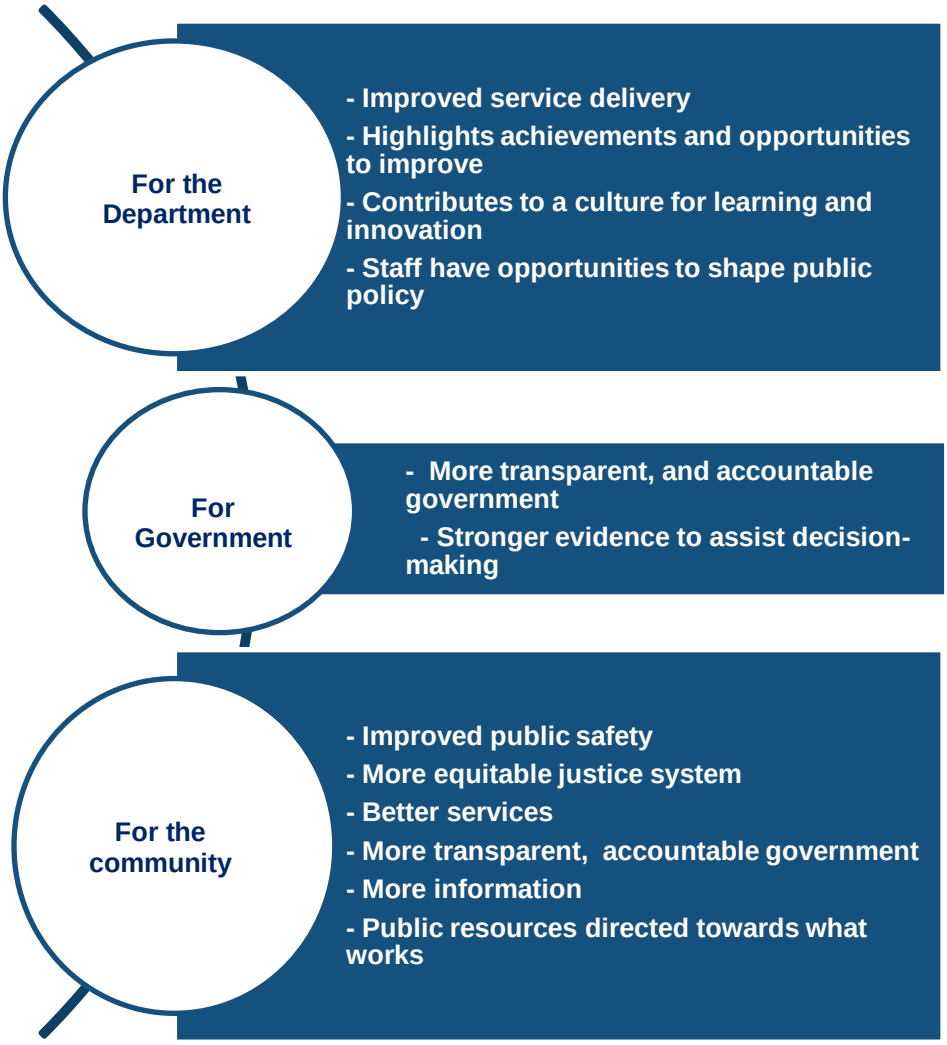
---

[5] Adapted from Department of Premier and Cabinet (2016) *NSW Government Program Evaluation Guidelines*, Sydney.
[6] Department of Premier and Cabinet (2016) *NSW Government Program Evaluation Guidelines*, Sydney, p 4.
[7] Department of Premier and Cabinet (2016) *NSW Government Program Evaluation Guidelines*, Sydney.

Benefits of evaluation

Without evaluation we have no way of knowing whether our policies and programs are achieving their stated objectives or whether they are the most cost-effective ways of doing so. Evaluation can also be a core part of contributing to **innovation** in government, by allowing us to test alternative ways of delivering or designing a program for different target groups. Evaluations are also a good way to **share knowledge** about our work with other agencies and the public, and to show that we are using public funds on programs that are fit-for-purpose. Figure 2 identifies other potential benefits.

Figure 2: Potential benefits of evaluation



The purpose, audience and significance of a program will inform the scope and type of evaluation chosen to assess the impact of a program. This Framework covers three key types of program evaluation: process, outcome and economic.

Process evaluation

Process evaluation considers how a program is delivered, describing current operating conditions and identifying any issues that are stopping the program from being implemented as intended.[8] A process evaluation may consider which aspects of a program are working well, and which aspects could be improved to ensure the best outcomes.

Process evaluation may include methods such as document reviews, stakeholder interviews, administrative program data analysis, observation and surveys.

Examples of process evaluation questions
Was the program implemented as intended?
Are there any barriers to program delivery?
To what extent is the program reaching intended recipients?
Are the outputs of the program as expected?

Example - process evaluation

In 1999 NSW Police introduced a new program ("Operation VENDAS") designed to increase the crime clear-up rate by making more effective use of forensic (e.g. DNA, fingerprint) evidence. To evaluate the program, the NSW Bureau of Crime Statistics and Research compared the trend in clear-up rates in sites where VENDAS was operating with the trend in sites where it was not operating. No difference was found in the crime clear up rates in the test and control sites. A process evaluation revealed that this is because there were only modest increases in the quantity of forensic evidence collected at crime scenes in the test sites.[9] You can find the report here.

Outcome evaluation

Outcome evaluation seeks to determine whether programs and policies are achieving their objectives.[10] We conduct these evaluations when we are trying to determine whether or not specific results, impacts or changes occurred because of our program.

Outcome evaluation may also be able to help us identify for whom a program works best, and in what circumstances. Ideally, it will also identify any unintended consequences for participants and stakeholders. Outcome evaluation is best conducted once a program has been bedded down and is operating smoothly. Particularly complex or large programs may need to be in operation for a number of years before a quality outcome evaluation can be conducted.

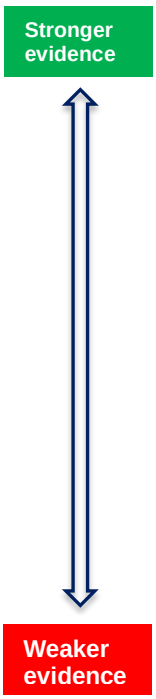
To conduct an outcome evaluation, you will need to establish *the counterfactual*: what would have happened if you did not introduce the program? Control groups and baselines are two common examples of a counterfactual.

Examples of outcome evaluation questions
Did the program achieve its stated objectives?
To what extent can changes be attributed to the program?
Does the program work better for some groups than others?
Are there any unintended or undesirable consequences for participants or stakeholders?

A range of research designs can be used to conduct an outcome evaluation:

[8] Department of Premier and Cabinet (2016) *NSW Government Program Evaluation Guidelines*, Sydney.
[9] Jones, C. & Weatherburn, D. (2004) *Evaluating Police Operations (1): A Process and Outcome Evaluation of Operation Vendas,* Sydney: NSW Bureau of Crime Statistics and Research.
[10] Department of Premier and Cabinet (2016) *NSW Government Program Evaluation Guidelines*, Sydney.

Experimental: the strongest method for demonstrating a causal relationship between program activities and outcomes.[11] It measures changes in the outcome for people randomly allocated to either an 'intervention' group (those in the program) or a 'control' group (people not in the program)**. e.g.** randomised controlled trials.[12] Where the groups are large enough and there are no systematic differences between them prior to the program, any differences between the two groups after the program operates must be due to the program.



Quasi-experimental: typically used when experimental designs are not feasible, due to high costs or ethical concerns. These compare outcomes for program participants, either against a non-random control group or at different phases of the program. **e.g.** pre and post studies with a control group, cohort studies, multiple baseline design. If you wish to employ a control group, you will need to consider whether or not you can create a group who are identical to the people in the program, e.g. people eligible for a program but who do not participate in it because there are limited places available.

Observational: these studies do not use a control group but examine changes pre and post program implementation, or employ statistical controls or use qualitative data only **e.g.** stakeholder interviews, expert opinion, pre and post studies with no control group. The key risk with observational studies is that differences in outcomes between those who participate in a program and those who don't may be due to factors other than the program.

It is important to conduct both process and outcome evaluations for complex or priority programs. A process evaluation should be conducted first, followed by the outcome evaluation. If an outcome evaluation yields negative results for a program, but no process evaluation has been conducted, it cannot be known if the negative results are due to poor implementation of the program, or the program not being fit-for-purpose in the first place.[13]

Example – outcome evaluation
In 2011, the NSW Bureau of Crime Statistics and Research (BOCSAR) conducted a study for the NSW Drug Court to see whether more intensive supervision reduced drug use among Drug Court participants.[14] Participants entering the Drug Court Program were randomly allocated to a supervision as usual (SAU) group (1 meeting a week with the Drug Court Judge for three months) or an intensive judicial supervision (IJS) group (2 meetings a week with the Drug Court Judge for six months). At follow up, the results showed that the IJS group had fewer positive tests for illicit drug use, lower self-reported frequency of drug use and fewer sanctions than the SAU group. You can find the report here.

Economic evaluation

Economic evaluation identifies, measures, and values a program's economic costs and benefits.[15] Economic evaluation can only be conducted once the costs of a program are known, and reliable data about the program's outcomes exists (i.e. after an outcomes evaluation). There are two key types:

[11] Department of Premier and Cabinet (2016) *NSW Government Program Evaluation Guidelines*, Sydney.
[12] For further information, see Weatherburn, D (2009) *Policy and program evaluation: recommendations for criminal justice policy analysts and advisors,* Sydney, pp 2-3.
[13] Weatherburn, D (2009) *Policy and program evaluation: recommendations for criminal justice policy analysts and advisors,* Sydney, p 3.
[14] Jones, C. (2011) *Intensive judicial supervision and drug court outcomes: Interim findings from a randomised controlled trial*, Sydney: NSW Bureau of Crime Statistics and Research.
[15] Department of Premier and Cabinet (2016) *NSW Government Program Evaluation Guidelines*, Sydney.

- **Cost-benefit analysis (CBA)**: determines the net benefits of a program to society and whether the benefits outweigh the costs of providing the program. CBA is the preferred method of economic evaluation for NSW Government programs.[16]
- **Cost-effectiveness analysis (CEA)**: determines which of two or more strategies, programs, or policies produces a specified outcome at the lowest cost. CEA does not substitute for CBA of NSW Government programs. It should be used only as a supplementary approach.[17]

Examples of economic evaluation questions
Do the benefits of program A outweigh its costs?
What is the marginal cost associated with a program? (i.e. how much does it cost to put each extra person on the program?)
Could resources be allocated more efficiently?
What components of the program are most costly?
Is program A more cost-effective than program B in reducing domestic violence re-offending?

Example – economic evaluation

In 2012, BOCSAR compared the cost-effectiveness of Youth Justice Conferences (YJCs) to the Children's Court in reducing juvenile re-offending.[18] Costs of processing and rates of re-offending were calculated using two matched samples of juveniles; one of which had their cases dealt with by the NSW Children's Court and the other of which had their cases dealt with at a YJC. The results of an earlier study had shown no difference in rates of re-offending among these matched samples. The costing analysis, however, revealed that the average cost of a YJC case disposal was approximately 18 per cent less than that of a court disposal—making YJCs considerably more cost-effective than the Children's Court for the kinds of cases that YJCs deal with. You can find the report here.

[16] NSW Treasury (2017) *NSW Government Guide to Cost-Benefit Analysis*, Sydney, p ii. https://www.treasury.nsw.gov.au/sites/default/files/2017-03/TPP17-03%20NSW%20Government%20Guide%20to%20Cost-Benefit%20Analysis%20-%20pdf.pdf

[17] NSW Treasury (2017) *NSW Government Guide to Cost-Benefit Analysis*, Sydney, p ii. https://www.treasury.nsw.gov.au/sites/default/files/2017-03/TPP17-03%20NSW%20Government%20Guide%20to%20Cost-Benefit%20Analysis%20-%20pdf.pdf

[18] Webber, A. (2012) *Youth Justice Conferences versus Children's Court: A comparison of cost-effectiveness*, Sydney: NSW Bureau of Crime Statistics and Research.

2. Principles of program evaluation

These principles should underpin the planning and conduct of an evaluation of a Justice program. They are derived from the NSW Government program evaluation guidelines.[19]

1. Plan evaluation early

Evaluation should be built into the design of a program. Planning an evaluation early will help ensure that the program has clear aims and objectives, and a strong rationale. Articulating evaluation questions from the outset will help to identify data needs and sources, which can inform ongoing monitoring and evaluation. Planning evaluation early can also significantly increase the types and robustness of evaluations that may be conducted.

An important early step is to develop a **program logic**. This is a diagram showing how the program is intended to work. Program logic links activities with outputs, impacts and outcomes, and aims to show causal links for a program. It can act as a single source of truth for program managers and stakeholders about the issues a program will address, the outcomes the program is intended to achieve, and how it is expected to achieve them.

Appendix 2 provides an example of a program logic model.

Obtain early approval to develop and/or use key datasets. Doing this early on will establish baseline data to compare against future data, and help to measure the program's impact.

Ensure that all key stakeholders are given an opportunity to comment on an evaluation design before it is finalised. This helps avoid criticism of the design if the evaluation indicates the program failed to produce its intended outcomes.

Document this work in an **evaluation plan**. This plan should accompany the business case.

Appendix 3 outlines the information that should be included in an evaluation plan.

2. Appropriately resource evaluations

Consider the resources and timeframe needed to conduct an evaluation of the program (or specific aspects of the program), and document it in your business case and evaluation plan. Be realistic. Time and resource availability will impact the design and scale of an evaluation. Focus on the most relevant evaluation questions to keep your evaluation manageable.

3. Be as rigorous as you can

Evaluations should be methodologically sound and replicable in accordance with the program size, risk, priority and significance.[20] Where possible, use methods which will produce reliable findings and enable sound conclusions to be reached about a program's effectiveness.

A robust evaluation of a large program may include the use of existing administrative datasets, surveys of clients and interviews with key stakeholders. Smaller evaluations are likely to be conducted internally, by staff skilled in data collection and analysis. Use the *NSW Department of Premier and Cabinet's Evaluation Toolkit* to guide the design of your evaluation.

[19] Department of Premier and Cabinet (2016) *NSW Government Program Evaluation Guidelines*, Sydney.
[20] Department of Premier and Cabinet (2016) *NSW Government Program Evaluation Guidelines*, Sydney.

4. Expertise

Evaluation teams should include relevant expertise. In most cases this will mean involving people with strong statistical skills and substantial experience in program evaluation.

5. Independence

Good practice dictates that the evaluator should be independent from the program manager.

In deciding whether to conduct an evaluation internally or externally, consider whether the project is a priority, the technical expertise required, and whether additional resources are required to ensure timely delivery.

If an evaluation is to be conducted externally, it is essential to comply with NSW Government and departmental procurement policies, procedures and guidelines. Please visit the Justice Intranet for further information.

In most cases involving criminal justice policies or programs, the **NSW Bureau of Crime Statistics and Research** should be commissioned to conduct the evaluation. Where necessary (e.g. where economic evaluation is a requirement), the Bureau will work with a suitable partner.

6. Be informed and guided by stakeholders

Stakeholders are vital to the evaluation process. Stakeholders are people or agencies that have an investment in the evaluation. Stakeholders can include primary intended users such as program managers, other agencies assisting with delivery, and program participants. Potential program users, who represent unmet need, may also be stakeholders.

Evaluations should foster stakeholder input throughout the whole process. In the planning stage, stakeholders should inform the definition of activities, outputs and expected outcomes. They can also assist in identifying what can be measured to indicate a program's impact. When the evaluation is underway, they can support data collection. At the end stage, they are key to helping the findings influence policy.

The department contracts a number of non-government organisations to deliver services, such as Youth on Track. Where applicable, funding agreements with non-government organisations should specify that service providers are to support evaluations of department programs.

7. Ethics

Evaluation can present potential risks to participants, in terms of loss of privacy, damage to vulnerable groups, and physical or mental harm.

Staff and other parties engaged to conduct evaluations must do so in accordance with the Department's Code of Ethics and Conduct. All evaluations conducted by and/or for the Department must comply with the ethical principles set out in the Guidelines for the Ethical Conduct of Evaluation, produced by the Australasian Evaluation Society. Any research or evaluation that involves Corrective Services staff or offenders/inmates must also comply with the requirements of the Corrective Services Ethics Committee.

Consider whether your program evaluation is likely to involve vulnerable or distinct cultural groups, such as Aboriginal people or refugees. You will need to consider this in your planning and design, in terms of culturally appropriate methods of data collection, and provision of feedback to the community. You may wish to consult the Australian Institute of Aboriginal and Torres Strait Islander Studies Guidelines for Ethical Research in Australian Indigenous Studies when designing program evaluations involving this group. You may also consider engaging a consultant from the community to assist with the planning and conduct of the evaluation.

8. Be open and transparent

As per BOCSAR practice, key stakeholders should always be given a briefing on initial findings before an evaluation report is drafted. They should also be given an opportunity to comment on any draft report before it is finalised, although responsibility for the final content of the report lies with evaluators. Consideration should also be given to seeking independent expert review of the draft report before it is finalised.

Make sure the final report clearly explains the aim of the program being evaluated, the means by which the program is intended to achieve its outcomes, the justification for the methods chosen to evaluate the program, the assumptions underpinning those methods, the results and any conclusions drawn from those results. This will enhance accountability and credibility, and promote evidence-based policy development.

Evaluation reports must be publicly released in a timely manner, unless there is an overriding public interest against disclosure, in line with the *Government Information (Public Access) Act 2009* (GIPAA).[21]

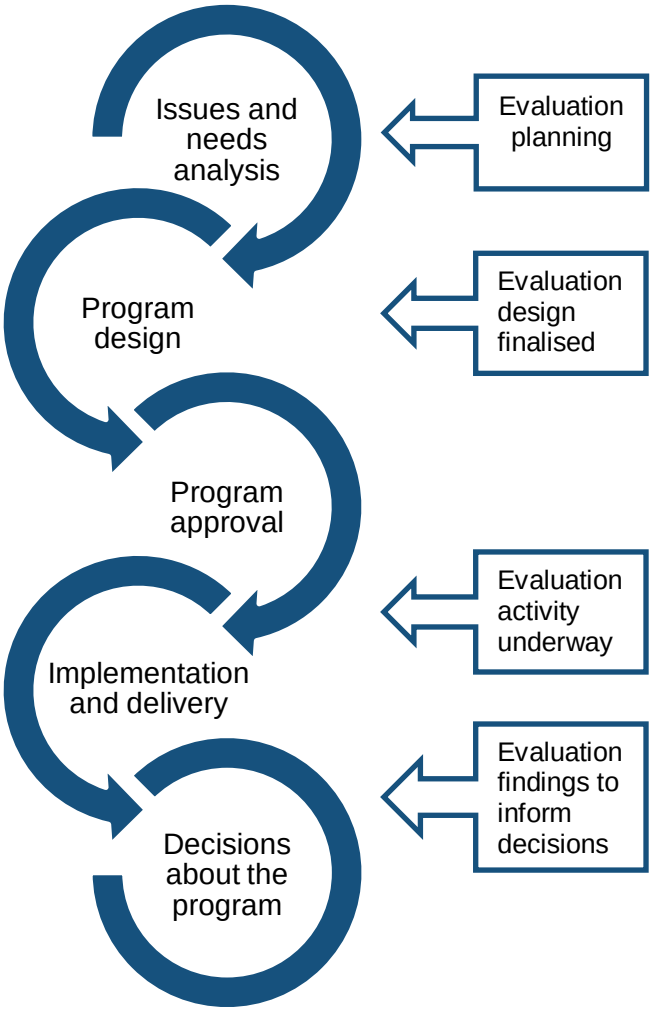
[21] Department of Premier and Cabinet. (2016) *Program Evaluation*, Circular 2016-01, Sydney,

3. When to evaluate

Evaluation can take place across the lifecycle of a program, in the phases of design, implementation, delivery and conclusion. Evaluation has an equally important role to play in testing the impact of new programs, as it does in assessing whether existing programs continue to be relevant, are delivered as intended, and are resulting in effective outcomes.

Outcome evaluations should be carried out once a program is stable, and ideally before it is expanded. Where possible, link the timing of an evaluation to when decisions about the program's continuation are going to be made. If you are evaluating a time-limited program, remember to plan your evaluation so that the findings are available several months prior to the end of the program's funding. That way, the findings can inform future business cases.

Figure 3. Process overlaps: program development and evaluation



Routinely engaging in evaluative thinking at these stages will help to embed evaluation in our activities. It also has a number of benefits for the Department. In the planning stage, it will help us to be more forward looking, because we need to anticipate the outcomes of programs we propose. Similarly, using evaluation to inform decisions will help improve our accountability and responsiveness, particularly if we use the information to inform decisions about a program's amendment, continuation, expansion or discontinuation.

4. Setting priorities for evaluation

The NSW Government Program Evaluation Guidelines provide direction on how to prioritise programs for evaluation. All major programs of strategic importance and/or financial investment must incorporate an evaluation strategy into their business case. Other programs should be prioritised for evaluation based on the following criteria (not ranked in order of importance):

- **Size**: the larger the government investment, the higher the priority for evaluation to assess if the program is achieving the intended outcomes and is value for money.
- **Strategic significance**: programs included in the Department's corporate and strategic plans, or those which address the Premier's priorities, are more important to evaluate.
- **Degree of risk**: programs that pose a high risk to government, the Department, or the community, should be evaluated to assess their effectiveness and justify investment.
- **Existing evidence base**: where there is a limited evidence base, because the program is innovative, a trial or is being transferred to a new setting/new group of clients, an evaluation should be conducted to assess its progress against intended outcomes.

The information at *Appendix 4* can help you to determine the scale of your program, and whether or not an evaluation is necessary.

A master list of Justice programs is being developed to map past, current and future evaluation activity across the Department and cluster, against priorities. This will provide further insight into programs which may benefit from evaluation.

Evaluation Schedule

The Department is required to prepare an annual evaluation schedule for approval by the Cabinet Expenditure Review Committee (ERC) in September each year. It includes:

- a list of programs to be evaluated in the current financial year, and program evaluations completed last year
- program funding details
- expected program and evaluation completion dates
- type of evaluation (process, outcome and/or economic)
- who will conduct the evaluation (agency, cluster or contractor)
- for completed evaluations, major findings and how we are using them.

If your program is undergoing evaluation, please ensure it is reported on this schedule. Simply send an email with the above information to evaluation@justice.nsw.gov.au.

5. Data considerations

Evaluation involves collecting and analysing relevant, reliable data. When planning your evaluation, consider what you want to know, what you need to measure and the data needed to be able to do this. There may be existing baseline data which you can use to compare to future data. You should also consider any new data that should be collected to support your evaluation.

There are many different types of data that could be analysed when conducting an evaluation. You may need to analyse program data, court data, offender records, police records, population data and/or gather data from program users/managers by way of surveys or interviews.

Involve evaluation experts early on in the process who can assist you to determine the data you may need to collect and analyse. For criminal justice programs, please approach the Bureau of Crime Statistics and Research in the first instance. For other programs, please email evaluation@justice.nsw.gov.au for assistance.

You can use the **Program Implementation Checklist** to guide your thinking about the types of data that may need to be collected. You can find it in the Evaluation Hub on the intranet.

6. Disseminating evaluation findings

The primary purpose for conducting an evaluation is to inform justice programs, for the benefit of the people of NSW. Findings should be communicated. This will help provide accountability for past investment in programs, and guide future program decisions. There are three main audiences for our evaluations:

- Program staff/managers who will use the evaluation findings to inform day-to-day program decisions.
- High-level policy-makers who will use the evaluation to inform funding and policy design decisions. Consider how to relay the information to Justice Executive, Ministers, and Treasury.
- Community, including other practitioners, academics, media and program participants.

Findings can be disseminated in various ways:

- Place the final report online
- Evaluators can hold feedback sessions with stakeholders
- If suitable, coordinate with the Media Unit to communicate to the media
- Summarise the findings for relevant Committees and Groups
- Conference papers/academic journals.

The branch responsible for the program must brief relevant Justice Executive and Ministers on evaluation findings/reports, **prior** to the public release. Ensure the brief includes the information at *Appendix 5*.

7. Governance, roles and responsibilities

We can all contribute to building and fostering an evaluative culture in the department.

Table 1: Roles and responsibilities in building evaluative capability

Employees	Role / Responsibilities
Secretary	**Lead** a culture of evaluation: • Create an environment which recognises the importance of evaluation. • Report to Ministers/Cabinet on the program of works in the Department's rolling evaluation schedule.
Executive	**Drive** a culture of evaluation: • Build evaluation into key business processes, recognising the importance of evaluation and ensuring appropriate resourcing/independence/quality of evaluations in their portfolio. • Define which decisions are to be informed by evaluation. • Ensure relevant evaluation findings are acted upon.
Evaluation experts in the Department	**Build** and **strengthen** a culture of evaluation: • Provide expert evaluation advice and support to other departmental employees, including reviewing of evaluation plans and reports, and assisting with other activities e.g development of program logic, discussions about appropriate use of data, and selection of appropriate indicators. • Conduct, and communicate findings of, complex evaluations. • Coordinate a centralised repository of completed evaluations.
Program managers	**Coordinate** evaluation in your area of responsibility: • Ensure staff are aware of the Evaluation Framework. • Ensure programs can be evaluated, by identifying clear outcomes, and measurable performance indicators from outset. • Promote learning and program improvement through analysis of current evidence base and use of evaluation findings. • Work collaboratively with evaluation experts and share learnings widely.
Employees	**Contribute** to a culture of evaluation by: • Practice evaluative thinking in program design. • Formulate program logic for projects in area of responsibility. • Work collaboratively with evaluation experts. • Document learnings to inform future evaluations.
Evaluation Reference Group *(to be established)*	• provide strategic and subject matter expertise to facilitate the department meeting whole-of-government evaluation requirements • ensure a robust process for prioritising and reporting evaluations • provide advice on the setting and delivery of strategic evaluation priorities • provide advice on the effective operation and delivery of Justice cluster evaluation activity

References

BetterEvaluation (2014) *BetterEvaluation Rainbow Framework*, retrieved April 2017 from http://www.betterevaluation.org/sites/default/files/Rainbow%20Framework.pdf.

Department for Correctional Services (2014) *Research and Evaluation Framework*, Government of South Australia, Adelaide.

Department of Education and Communities (2014) *NSW Department of Education and Communities Evaluation Framework*, NSW Government, Sydney.

Department of Justice and Regulation (2014) *Corrections Victoria Evaluation Framework 2015-2018*, State of Victoria, Melbourne.

Department of Premier and Cabinet (2016) *Criminal Justice Evaluation Framework*, Queensland Government, Brisbane.

Department of Premier and Cabinet (2017) *Evaluation Toolkit*, NSW Government, Sydney.

Department of Premier and Cabinet (2016) *NSW Government Program Evaluation Guidelines*, NSW Government, Sydney.

Department of Premier and Cabinet (2016) *Program Evaluation Circular C2016-01*, NSW Government, Sydney.

Gertler, PJ, Martinez, S, Premand, P, Rawlings, LB and Vermeersch, CMJ. (2016) *Impact Evaluation in Practice (second edition)*, World Bank Group, Washington DC.

Jones, C. (2011) *Intensive judicial supervision and drug court outcomes: Interim findings from a randomised controlled trial*, Sydney: NSW Bureau of Crime Statistics and Research.

Jones, C. & Weatherburn, D. (2004) *Evaluating Police Operations (1): A Process and Outcome Evaluation of Operation Vendas*, NSW Bureau of Crime Statistics and Research, Sydney.

Justice Health and Forensic Mental Health Network (2014) *Research Strategy 2014 – 2017*, NSW Government, Matraville.

Justice Health (2016) *Research and Evaluation Framework 2015 – 2019*, Victorian Government, Melbourne.

NSW Treasury (2017) *NSW Government Guide to Cost-Benefit Analysis*, Sydney.

OECD (2010) *Glossary of Key Terms in Evaluation and Results Based Management*, OECD, Paris.

Queensland Treasury and Trade (2014) *Program Evaluation Guidelines*, Queensland Government, Brisbane.

United Nations Evaluation Group (2010) UNEG Quality Checklist for Evaluation Reports, United Nations,

Weatherburn, D. (2009) *Policy and Program Evaluation: recommendations for criminal justice policy analysts and advisors*, NSW Bureau of Crime Statistics and Research, Sydney.

Webber, A. (2012) *Youth Justice Conferences versus Children's Court: A comparison of cost-effectiveness*, Sydney: NSW Bureau of Crime Statistics and Research.

Appendix 1: Glossary

Term	Definition
Audit (performance)	Performance audit reports concern the efficiency, effectiveness, economy (and in certain circumstances, compliance aspects) of a particular government activity.[22]
Baseline	Information collected before or at the start of a program that provides a basis for planning and/or assessing subsequent program progress and outcomes.
Cost-benefit analysis	An appraisal and evaluation technique that estimates the costs and benefits of a project or program in monetary terms.[23] A comprehensive Guide to Cost-Benefit Analysis is available from NSW Treasury.
Cost-effectiveness analysis	A technique for comparing the costs of alternative proposals to find the minimum cost solution which achieves the given objective.[24] This does not assess the net impact on social welfare, and should be used only as a supplementary approach to cost-benefit analysis.
Economic evaluation	Involves the identification, measurement, and valuation of a program's economic costs and benefits.
Effectiveness	The extent to which a program achieves its objectives.
Efficiency	The extent to which a program is delivered with the lowest possible use of resources, to the areas of greatest need, and continues to improve over time by finding better or lower cost ways to deliver outcomes
Evaluation	A rigorous, systematic and objective process to assess the effectiveness, efficiency, appropriateness and sustainability of programs.
Experimental Design	Considered the strongest methodology for demonstrating a causal relationship between program activities and outcomes. It measures changes in the desired outcome for participants in an intervention group and those in a control group who do not differ in any systematic way (e.g. randomised controlled trials). Results are independent of selection processes and any associated bias
Expert opinion	The views of a person generally considered to be very knowledgeable in a particular field.
Findings	Factual statements about a program which are based on empirical evidence. Findings include statements and representations of the data, but not interpretations, judgments or conclusions about what the findings mean or imply
Inputs	The financial, human and material resources used for a program.
Key Performance Indicators (KPIs)	A tool to measure our performance in achieving our outcomes. They are simple, clearly defined indicators which allow us to assess current progress against our program objectives.
Monitoring	A process to periodically report against planned targets (Key Performance Indicators). Monitoring is typically focused on outputs

[22] Audit Office of NSW, 2017, *Performance Audit Reports* (viewed 24 July 2017) http://www.audit.nsw.gov.au/publications/performance-audit-reports
[23] NSW Treasury, March 2017, *NSW Government Guide to Cost-Benefit Analysis*, Sydney.
[24] Ibid.

	rather than outcomes and is used to inform managers about the progress of a program and to detect problems that may be able to be addressed through corrective actions
Outcome	A result or effect that is caused by or attributable to the program.
Outcome evaluation	Outcome evaluation seeks to verify a causal link between pre-defined program activities and outcomes. It identifies the overall positive or negative outcome, and ideally for whom and under what conditions the program is most effective. It is preferable that it also considers any unintended consequences for participants or stakeholders. Outcome evaluation should occur when the program has been running long enough to produce reliable results
Output	The products, goods, and services which are produced by the program
Pre and post studies	A pre and post study involves observations about participants before and after the program. It measures whether a change in outcome measures has occurred since a program has commenced, but, unless a control group is involved, it cannot attribute the change to the program.
Process evaluation	Process evaluation investigates how a program is delivered, and may consider alternative delivery processes. It can help to distinguish implementation issues from ineffective programs. Process evaluations draw on data from a wide range of sources (e.g. document review, observation, surveys, qualitative inquiry and analysis of administrative program data) depending on the nature of the program being evaluated.
Program	A set of activities managed together over a sustained period of time that aims to deliver outcomes for a client or client group.
Program logic	The program logic is a diagram that illustrates the logical linkage between the identified need or issue that a program is seeking to address; its intended activities and processes; their outputs; and the intended program outcomes. A template is provided overleaf.
Program review	Typically quicker, more operational assessments of ‘how we are going’, often to inform continuous improvement. Reviews generally take place after implementation has started and may be useful when there is insufficient information to conduct an evaluation.
Research	Research is a systematic process of inquiry and discovery in order to produce knowledge and understanding. It is based on the process of hypothesis generation and testing. Hypotheses are typically developed on the basis of the existing body of knowledge.
Stakeholders	Those groups who affect and/or could be affected by an organisation’s activities, products or services and associated performance. They are likely to have an investment in evaluation and its findings.
Value for money	Value for money is achieved when the maximum benefit is obtained from the program provided within the resources available.

Appendix 2: Program Logic

A program logic is a diagram used to show how a program is expected to bring about changes to achieve expected outcomes, with regard to its specific problems, goals, and design. It illustrates what the program is trying to achieve. Below is a template for a program logic model.[25]

Program objective: A specific statement describing what the program aims to do e.g. a behaviour change program for domestic violence offenders and their families, delivered in the community, which aims to reduce the risk of reoffending.

		Inputs	Outputs		Outcomes – Impacts		
			Activities	Participation	Short-term	Medium-term	Long-term
Situation Context for the program e.g. why is the program needed? Consider creating a problem statement. e.g. The reoffending rate for DV is higher than other crimes.	**Priorities** Consider state & departmental priorities that are driving a reform agenda e.g. Premier's priorities; DJ Corporate Plan discuss DV reoffending	**(What we invest)** consider resources, including staff, money, time, materials, equipment, partners, volunteers e.g. $5 million; 1 program manager; 2 new part-time trainers over 3 years.	**What we do** Specify the action being taken, such as delivery of services; development of products; consultations undertaken e.g. delivery of program; information provided to families	**Who we reach** who is involved? Consider clients, participants, decision-makers and stakeholders e.g. offenders, service providers, families of offenders.	**Learning**: changes in awareness; knowledge; attitude; skills e.g. within 6 months, increased participation rates.	**Action**: changes in behaviour, practice, decision-making or policies e.g. within 12 months, improved program completion rates; offenders access support services when needed.	**Conditions**: changes to social; economic environments e.g. within 5 years, rates of DV reoffending are down; family wellbeing and safety is enhanced.
					Specify each of these timeframes for your project. Outcomes should be linked to program objectives.		
		Assumptions: what beliefs do you have about the program/its operation e.g. partner agencies will be involved; inmates will participate in the program			**External factors:** what is outside your control but may impact your program? e.g. funding, political will, other programs available to the target audience		
Evaluation							

[25] Adapted from Taylor-Powell, E (2011) Logic models: A framework for program planning and evaluation. University of Wisconsin. For further guidance about program logic, see also https://aifs.gov.au/cfca/how-develop-program-logic-planning-and-evaluation

Appendix 3: Evaluation plan

The evaluation plan is a document that sets out what is being evaluated and why, how the evaluation will be conducted, and how the findings will be used. Evaluation plans that are developed in consultation with stakeholders can help ensure a common understanding of the purpose of an evaluation and its process.

The specific content and format of an evaluation plan will vary according to the program being evaluated. At a minimum, plans should include the following information:

Subject of the evaluation This should include a very brief overview of the program, including its history, objectives and key stakeholders. Add your program logic diagram if one has been developed.
Strategic priority? How is the program aligned to the Premier's priorities, corporate priorities and/or other plans? What is the scale of the evaluation (tier 1-4)?
Purpose of the evaluation Why is the evaluation being conducted? Consider the decisions that will be made as a result of the findings. Are you trying to: - assess the program on its merits and against its claims - guide your decisions about whether to stop it, or continue it, and if so in what form - enable ongoing improvements and adjustments - identify factors that need to be taken into account as the program is expanded - identify emerging needs, gaps or priorities or - collect baseline data for use in future evaluation?
Primary audience Who will receive and use this information? e.g. Cabinet, Cabinet Committee, Ministers, Executive, policy-makers, program staff, stakeholders, agencies running similar programs, or the public.
Key evaluation questions What are the key questions the evaluation seeks to answer?
Resourcing, budget and timeframe Who will commission/conduct and manage the evaluation? Will it be conducted internally or externally? What are the human and material resources to be allocated to the evaluation? Who will be doing what? Do you have a quarantined evaluation budget? What is the expected timeline? What are the important milestones? Try to ensure the findings are available before funding runs out, so the evaluation can inform future investment decisions.
Data and methodology What are your indicators? Describe how the data will answer your evaluation questions. Where will the data come from – existing datasets, surveys, document analysis, stakeholder interviews? Who will be the data custodian? What alternative methods were decided against, and why?

Implementation State who is going to do what, and when. What are their responsibilities (e.g. conducting surveys, overseeing data collection) and what is the timeline for their key steps/milestones?
Risks and mitigation What risks to the evaluation exist, and how are they being addressed? Potential risks can include inability to recruit participants, low response rates or inconclusive findings.
Dissemination What are the plans to disseminate and/or publish the findings? How will the results be used? Where possible, evaluation findings should be publicly released to improve transparency, contribute to the evidence base and reduce duplication.
Privacy and ethical considerations What privacy/ethical issues have you identified and how will they be addressed? Does the evaluation need ethics clearance from an approved body?
Consultation Who has been consulted on this evaluation plan?
Expert evaluation review Have internal evaluation experts, such as BOCSAR, been consulted? Please attach their advice/comments.

Appendix 4: Program scale & evaluation

Use this table to determine the scale of your program (Tier 4, 3, 2 or 1) and whether or not it requires evaluation.[26]

	Scale	Program	Evaluation
Evaluation expected	4	• Strategic priority for government • Accountable to Cabinet or Cabinet committee; or multiple Ministers • Substantial investment • High-risk or controversial • Multiple agencies involved in delivery • Other factors: lack of evidence base; external reporting requirements; innovative trial	Formal evaluation mandatory. The report must be provided to Cabinet or a Cabinet Committee. Funding should be quarantined for an independent evaluation and dedicated staff. The evaluation should be led by evaluation experts from inside the Department, a consultant evaluator, or the Centre for Program Evaluation (Treasury).
	3	• Strategic priority for DJ • Accountable to Minister(s) • Significant investment (relative to cluster/agency) • Moderate to high-risk • Other factors: lack of evidence base; internal reporting requirements	Evaluation expected. The report must be provided to responsible Minister(s). Designated evaluation budget and staffing; the evaluation should be led by evaluation experts from inside the Department or a consultant evaluator.
Evaluation at Secretary's discretion	2	• Named in DJ Strategic Plan or identified as an emerging priority • Accountable to DJ Executive • Moderate department investment • Low to moderate risk • Other factors: not recently reviewed	Budget for evaluation and staff appropriate to agreed methodology. Staff should be external to the program to ensure independence.
	1	• Low priority • Accountable to Directors • Limited department investment • Low risk • Other factors: similar to other evaluated programs found to be successful	Budget for a review and a staff member (external to the program) to conduct it.

[26] Adapted from the NSW Government Program Evaluation Guidelines

Appendix 5: Evaluation Report - briefing

Once the evaluation report is finalised, you will need to brief the Secretary and Minister(s) about the findings and their implications. At a minimum, your brief should include the following information:

Subject of the evaluation This should include a very brief overview of the program, including its history, objectives and key stakeholders.
Key findings What are the key conclusions and findings arising from the evaluation? How are these findings being used?
Resourcing, budget and timeframe Did the evaluation stay within its anticipated budget, resourcing and timeframes? What is the public release date?
Expert evaluation review Have internal evaluation experts reviewed the final report? Please attach their advice/comments.
Audience and dissemination Who will receive and use this information?
Communication to DPC and NSW Treasury DPC and Treasury are to receive copies of all completed evaluation reports. Please provide an electronic copy of your report to evaluation@justice.nsw.gov.au, where staff in the Office of the Secretary will coordinate distribution to these agencies.

THIS PAGE INTENTIONALLY LEFT BLANK