

NSW Treasury

Policy and Guidelines: Evaluation

TPG22-22

February 2023



Acknowledgement of Country

The NSW Treasury acknowledges that Aboriginal and Torres Strait Islander peoples are the First Peoples and Traditional Custodians of Australia, and the oldest continuing culture in human history.

We pay respect to Elders past and present and commit to respecting the lands we walk on, and the communities we walk with.

We celebrate the deep and enduring connection of Aboriginal and Torres Strait Islander peoples to Country and acknowledge their continuing custodianship of the land, seas, and sky.

We acknowledge the ongoing stewardship of Aboriginal and Torres Strait Islander peoples, and the important contribution they make to our communities and economies.

We reflect on the continuing impact of government policies and practices and recognise our responsibility to work together with and for Aboriginal and Torres Strait Islander peoples, families, and communities, towards improved economic, social and cultural outcomes.

Artwork: 'Regeneration' by Josie Rose



Contents

Acknowledgement of Country	a
Evaluation Policy and Guidelines	2
1. Evaluation: introduction.....	6
2. Evaluation schedule	9
3. Evaluation design.....	12
4. Evaluation for new proposals.....	15
5. Evaluation practice and reporting	17
Definitions	34
Appendix A. Evaluation TPG: What's new?.....	37
Appendix B. Build evaluation culture and practice	39
Appendix C. Alignment with investment processes.....	41
Appendix D. Types of evaluation and example questions.....	44
Appendix E. Example theory-of-change and logic model: building evidence.....	51
Appendix F. Evaluation workbooks summary.....	53

Key information	
Treasury Policy and Guidelines (TPG) is relevant to?	<input checked="" type="checkbox"/> Government Sector Finance Agencies <input checked="" type="checkbox"/> General Government Sector <input checked="" type="checkbox"/> Public non-financial corporation <input checked="" type="checkbox"/> Public financial corporation <input checked="" type="checkbox"/> State Owned Corporations <input type="checkbox"/> Other
	<input checked="" type="checkbox"/> Executive agencies related to Departments
	<input type="checkbox"/> Subsidiaries of the NSW Government established under the Corporations Act 2001
Date issued	February 2023
Review date	2027
<input checked="" type="checkbox"/> Replaces <input type="checkbox"/> Replaced by	Treasury Circular Program Evaluation TC18-03 NSW Government Program Evaluation Guidelines 2016
Issuing/Publishing entity	NSW Treasury
Related instrument(s)	TPP18-06 Business Case Guidelines TPG23-08 NSW Guide to Cost Benefit Analysis
Document approver	Joann Wilkie, Deputy Secretary, Economic Strategy and Productivity Group, NSW Treasury
Contact	Centre for Economic Evidence cee@treasury.nsw.gov.au
Document contains	
<input checked="" type="checkbox"/> MANDATORY POLICY compliance set out by NSW Treasury.	
<input checked="" type="checkbox"/> RECOMMENDED POLICY reflecting best practice standards.	
<input checked="" type="checkbox"/> GUIDANCE/ADDITIONAL INFORMATION to provide clarity or explain requirements in detail.	

Revision history				
Document version number	Approval Date	Author	Approver	Description

Evaluation Policy and Guidelines

Purpose

These Policy and Guidelines set out mandatory requirements, recommendations and guidance for NSW General Government Sector agencies and other government entities to plan for and conduct the evaluation of policies, projects, regulations, and programs.¹ The Evaluation Policy and Guidelines (Evaluation TPG) is part of an investment framework that informs policy and budget setting in New South Wales. The framework is used to build evidence to support achievement of State Outcomes and improve wellbeing for the people of New South Wales.

Overview

A strong understanding of what works to deliver outcomes and benefits will support improved wellbeing, economic prosperity, and fiscal sustainability for the people of New South Wales. Rigorous and regular assessment of initiatives – including before implementation (ex-ante appraisals) and after implementation (ex-post evaluations) – provides a complete evidence base that can more effectively shape policy and service delivery.

Evaluation (ex-post) is a systematic and transparent process to assess an initiative's appropriateness, effectiveness, efficiency, and social benefits. It is a key tool to improve the design and delivery of ongoing and new initiatives, inform investment decision making and support the delivery of State Outcomes.

Evaluation should be undertaken to verify if initiatives are implemented as planned, and delivering intended social, economic, environmental, and cultural outcomes and benefits. It can be undertaken during the life of an initiative (to inform management of how the initiative is tracking) or at its end (to provide evidence of how well an initiative has worked). Monitoring during the life of an initiative can support performance reporting and evaluation. Best-practice is to plan for monitoring and evaluation at the design stage of an initiative.

All NSW Government agencies are required to regularly examine initiatives to ensure they are achieving intended outcomes and are providing a social benefit to the people of New South Wales. An initiative may be any capital, ICT, recurrent, regulatory or policy action undertaken or resourced by the NSW Government, with the aim to achieve an outcome (or outcomes) for the NSW community.

This Evaluation TPG sets out mandatory policy requirements (compliance set out by NSW Treasury) and provides recommendations (reflecting best practice standards) and guidance to support consistent, high quality and transparent evaluations. Evaluations of government-resourced ongoing and new initiatives should be prioritised, planned,

¹ For the purposes of this policy, 'agency' is all General Government Sector agencies (as listed in NSW Budget Paper No. 1). All General Government Sector agencies and non-commercial public non-financial corporations (as listed in NSW Budget Paper No. 1) are considered within the scope of the policy and are required to adopt the core requirements.

conducted, and reported with reference to the requirements, recommendations and guidance set out in this TPG.

This TPG updates and supersedes the Program Evaluation Guidelines 2016 and TC18-03 Program Evaluation (see *Appendix A. Evaluation TPG: What's new?*). It has been developed, through an extensive consultation process, to support the government's commitment to evidence-informed policy and investment.

This TPG is one of a suite of investment policies and guides – including [Outcome Budgeting \(TPP18-09\)](#), [NSW Government Business Case Guidelines \(TPP18-06\)](#), [NSW Gateway Policy \(TPG22-12\)](#) and [NSW Government Guide to Cost-Benefit Analysis \(TPG23-08\)](#) – that establish the framework to provide evidence to inform policy, prioritise investment, and maximise social benefits.

This TPG, and supporting *evaluation workbooks* and *evaluation resources*, provide guidance and information on planning, implementing, and reporting on process, outcome, and economic and cost-benefit analysis evaluations. They apply to both internally conducted and externally commissioned evaluations. They do not, however, seek to cover the full range of activities that are conducted by agencies to develop, manage, and review initiatives.

Summary of Requirements

Mandatory
<p>1. Evaluation</p> <ul style="list-style-type: none"> a) All NSW Government agencies are required to coordinate monitoring and periodic evaluation of their initiatives, both ongoing and new. b) Initiatives resourced by the NSW Government must be regularly examined to assess their achievement of intended outcomes and net social benefit to the people of New South Wales.
<p>2. Evaluation schedule</p> <ul style="list-style-type: none"> a) Clusters² are required to prioritise and schedule evaluations in line with the recommendations and guidance within this TPG, in response to initiative management needs and in discussion with NSW Treasury. b) The government may require new or ongoing initiatives to be evaluated, considering factors such as the Premier's priorities or Government commitments, size and risk, and external reporting requirements. c) Clusters should prioritise evaluation of new and ongoing initiatives of significant size, government priority, and risk. d) Under the NSW Gateway Policy (TPG22-12) assurance framework, a Gate 6 post-implementation assurance process may be applied, with benefits realisation and evaluation processes subject to independent expert review.

² NSW Government departments, agencies and organisations are arranged into groups called clusters.

Mandatory
<p>3. Evaluation design</p> <p>a) Monitoring and evaluation should be coordinated, as relevant to the initiative and the timing of the evaluation, to examine:</p> <ul style="list-style-type: none"> • implementation • appropriateness, including continued relevance to government and community priorities • efficiency • effectiveness in delivering outcomes • distribution of outcomes and benefits, and effects on equity (where relevant) • net social benefit and value for money.
<p>4. Evaluation for new proposals</p> <p>a) All proposals seeking government resources are required to include evidence to support initiative design and appraisal, and to plan for evaluation.</p> <p>Evaluation evidence:</p> <ul style="list-style-type: none"> • for new initiatives, identify the evidence that has informed the development and appraisal of the submission • for ongoing initiatives, identify when the initiative was last evaluated or subject to other review processes (e.g., audit or assurance), the findings and recommendations, and the agency response. <p>Evaluation planning:</p> <ul style="list-style-type: none"> • identify when the initiative will next be evaluated, or when evaluation timing will be decided • include high-level monitoring and evaluation planning in all submissions seeking NSW Government investment of \$10 million and above (this may be under a monitoring and evaluation framework or under a Benefits Realisation Management Framework), tailored to the size, priority, and risk of the initiative • incorporate resourcing requirements for monitoring and evaluation into the business case funding submission (or set funding aside) or include a plan for when resourcing will be determined and obtained.
<p>5. Evaluation practice and reporting</p> <p>a) Clusters, in consultation with Treasury, are required to nominate initiatives for evaluation reporting. Annual reporting to Treasury will include evaluations completed in the previous financial year, and evaluations due for completion in the upcoming financial year.</p> <p>b) Agencies are mandated to proactively and publicly release the findings of evaluations, unless there is an overriding public interest against disclosure of the information, in line with the <u>Government Information (Public Access) Act 2009</u>.</p>

Document Structure

This Evaluation TPG set out mandatory policy requirements and Treasury recommendations for monitoring and evaluation, with supporting guidance. It is supported by online [evaluation workbooks](#) and [evaluation resources](#).

Table 1: TPG structure

Section		Content
1. Evaluation: introduction		<ul style="list-style-type: none"> What is evaluation? What is a NSW Government 'initiative'? Why is evaluation undertaken? What are principles for evaluation?
2. Evaluation schedule		<ul style="list-style-type: none"> What initiatives should be prioritised for evaluation?
3. Evaluation design		<ul style="list-style-type: none"> What are the types of evaluation? How do these relate to each other? How can evaluation be tailored to initiative size, priority, and risk?
4. Evaluation for new proposals		<ul style="list-style-type: none"> What evaluation evidence should support proposals that seek government resources? What monitoring and evaluation planning should be undertaken at business case development?
5. Evaluation practice and reporting	1. Foundations for evaluation	<ul style="list-style-type: none"> What are the foundations for evaluation? What is a logic model?
	2. Monitoring and evaluation framework	<ul style="list-style-type: none"> What are the steps to build a monitoring and evaluation framework at initiative design and implementation?
	3. Evaluation plan: Design the evaluation	<ul style="list-style-type: none"> What are the steps for designing an evaluation?
	4. Evaluation plan: Manage the evaluation	<ul style="list-style-type: none"> What concerns and relationships should be addressed in managing evaluation design and implementation? What are standards related to cultural appropriateness and ethical conduct?
	5. Evaluation plan: Use the right expertise	<ul style="list-style-type: none"> When and how can independent expertise be utilised in an evaluation?
	6. Evaluation plan: Report and use evaluation findings	<ul style="list-style-type: none"> What steps support planning to report and use evaluation results? What is an example template for an evaluation report?
Definitions		<ul style="list-style-type: none"> What is the meaning of key technical terms in the TPG?
Appendices	Appendix A. Evaluation TPG: What's new?	<ul style="list-style-type: none"> What are the updates to the Program Evaluation Guidelines 2016?
	Appendix B. Build evaluation culture and practice	<ul style="list-style-type: none"> What are ways to build a culture of evaluation and support the practice of evaluation?
	Appendix C. Alignment with investment processes	<ul style="list-style-type: none"> How does evaluation align with other NSW Government processes and guidelines?
	Appendix D. Types of evaluation and example questions	<ul style="list-style-type: none"> What are the types of evaluation? How can evaluation also consider efficiency and equity? What questions may be relevant to different types of evaluation?
	Appendix E. Example theory-of-change and logic model: building evidence	<ul style="list-style-type: none"> What is an example of a theory of change? What is an example of a logic model?
	Appendix F. Evaluation workbooks summary	<ul style="list-style-type: none"> How do the evaluation workbooks support monitoring and evaluation throughout the investment lifecycle?

Evaluation workbooks, provide information to support monitoring and evaluation, including example templates. They are designed to be used flexibly as relevant to the initiative, the context in which it is being evaluated, and the resources allocated to the evaluation.

Evaluation resources, including technical notes, provide detailed technical guidance to the terms and methods used in evaluation, and identify best-practice standards.

1. Evaluation: introduction

Mandatory

- a) All NSW Government agencies are required to coordinate monitoring and periodic evaluation of their initiatives, both ongoing and new.
- b) Initiatives resourced by the NSW Government must be regularly examined to assess their achievement of intended outcomes and net social benefit to the people of New South Wales.

Guidance/Information

What is evaluation?

Evaluation is a systematic and transparent process that can be used to assess the appropriateness, efficiency, effectiveness, and net social benefits of an initiative, after it has been implemented. It can be undertaken at different stages of an initiative's life, including when it is underway or when it is completed.

Evaluation should be undertaken with a view towards informing decision-making. Be clear regarding:

- what is to be evaluated and why
- who will receive and use the evaluation findings
- what decisions the evaluation will inform.

What is a NSW Government initiative?

NSW Government investments are undertaken to deliver outcomes that benefit the people of New South Wales, improve wellbeing, and contribute to State Outcomes (the goals that the government is seeking to achieve for its citizens).

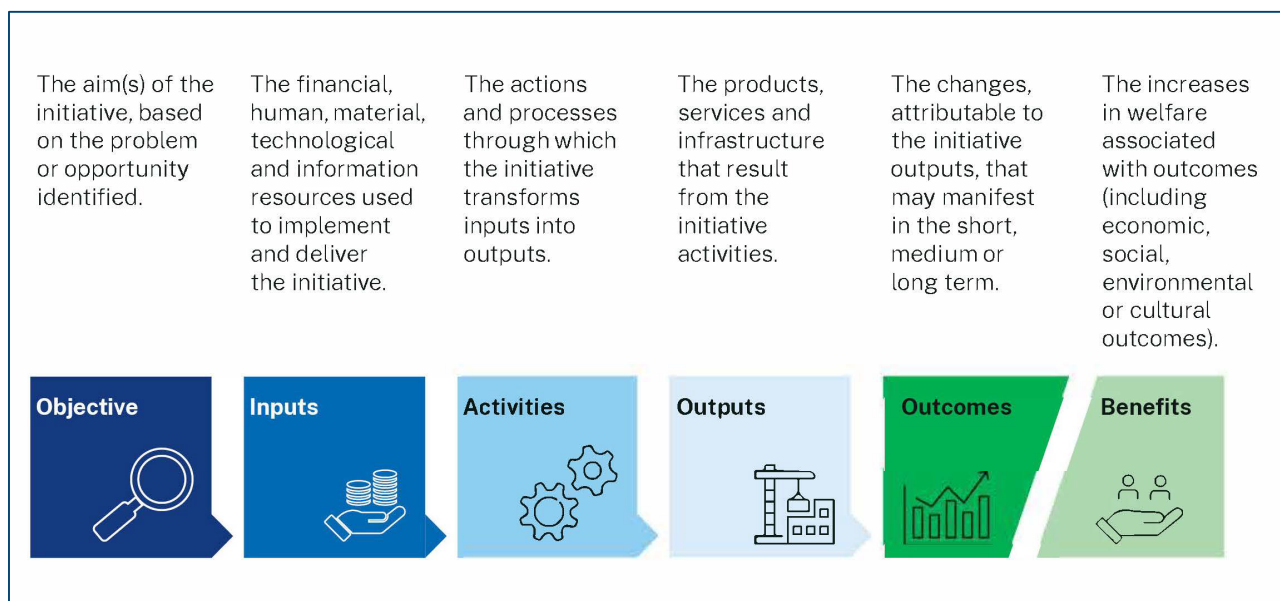
A NSW Government initiative may be any capital, recurrent, ICT, regulatory or policy action undertaken with the aim to achieve a positive outcome for the NSW community. This can be a *program*³, policy, strategy, service, project, or any series of related events. Initiatives may be delivered directly by government agencies or by government funded delivery partners from non-government sectors.

Initiatives can vary in size and structure; from a small initiative at a single location, a series of related events delivered over a period, or whole-of-government reforms with many components delivered by different agencies or governments.

The impacts of an initiative will include outcomes, and associated benefits (see *Figure 1*). In general, outcomes can be understood as the extent of changes that follow from an initiative, and benefits as the improvement in welfare related to that change (benefits are a measure of the value of outcomes to the NSW community – which may be monetary or non-monetary).

³ Under the Outcome Budgeting structure, a 'program' is "the collection of activities, tasks, divisions or function of an agency designed to deliver specific outputs that contribute towards achieving a State Outcome" (TPP18-09). An 'initiative' may sit within a program, contribute to the outcomes of more than one program, or be a program itself.

Figure 1: Initiative logic model



Why evaluate?

Evaluations are a key tool to inform evidence-based decision-making.

Understand an initiative in practice

Evaluation examines how an initiative works in practice. It can examine if an initiative is implemented as planned and if it is producing outcomes and benefits as intended. It assesses to what extent, in what situations and why an initiative is working (or not working).

Assess the effectiveness of initiatives in achieving State Outcomes

Evaluation can be used to examine and strengthen the relationship between investing in initiatives and achieving relevant State Outcome Indicator and Program Performance Measure targets. It can support the NSW Government's commitment to ensuring that all government spending is achieving State Outcomes for the people of New South Wales.

Support adaptive management and inform decision-making

Evaluation can be used to improve the design and delivery of an initiative and provide evidence to support continuing or expanding an effective initiative, or ending or changing an ineffective initiative.

Develop an evidence base of 'what works' in New South Wales

Evaluation will contribute to our evidence base of 'what works' in diverse communities and populations across New South Wales. As evaluation evidence is accumulated and shared, it will increase accountability and transparency regarding how initiatives work.

Inform future initiative design and forecasting

Evaluation that identifies the effectiveness of an initiative, and its actual costs and benefits, will provide evidence that can be used to improve the design and appraisal of future initiatives. Evaluation evidence will support longer-term policy design and investment decision-making.

Evaluation principles

Ten principles underpin effective monitoring and evaluation processes.

Table 2: Monitoring and evaluation principles

Principle	Guidance
1. Be clear regarding initiative purpose	Commence monitoring and evaluation planning with a clear awareness of the outcomes and benefits the initiative is intended to achieve, and how these align with State Outcomes. Use logic models.
2. Plan early for monitoring and evaluation	Build monitoring and evaluation planning and resourcing into initiative design and implementation. Develop evaluation planning across the ongoing initiatives of a cluster and budget for evaluation.
3. Design monitoring and evaluation with a view towards informing decision-making	When planning monitoring and evaluation, be clear regarding what needs to be known about how the initiative works in practice, the expected timing of outcomes and benefits, and what information is needed for decision-making and by when. Allow adequate time for evaluation.
4. Use sound evaluation methods	Tailor best-practice evaluation methods to suit the initiative's size, significance and risk, and the information required for decision-making. Consider what is practical and achievable.
5. Ensure proper governance and oversight	Establish governance structures and processes to ensure effective partnerships and oversight of monitoring and evaluation design, implementation, and reporting. Be clear regarding responsibilities.
6. Build stakeholder involvement in decision-making processes	Use collaborative and inclusive approaches to work with relevant stakeholders (internal and external to government) to design and conduct monitoring and evaluation. Use evaluation as an opportunity to build relationships with, and learn from, delivery partners, customers/clients, and the community.
7. Be ethical in design and conduct	Incorporate ethical considerations and requirements into monitoring and evaluation design and conduct.
8. Ensure the right mix of expertise and independence	Coordinate internal and external expertise as needed to ensure objective and robust evaluation. Use experienced evaluators who are independent from the delivery agency and project team, but always include initiative managers in evaluation planning.
9. Be transparent	Publicly release key evaluation information and findings. Provide information in formats that meet stakeholder communication needs.
10. Value ongoing learning	Be open to new learnings. Develop a formal response to consider and integrate evaluation findings. Share lessons learned.

Build a culture of evaluation

Evaluation is most effective when supported by an agency's leaders, adequately resourced, and part of a culture of enquiry and learning within an agency. Evaluation can be supported by building a culture of evaluation, developing evaluation capabilities, and integrating monitoring and evaluation into agency practices (see *Appendix B. Build evaluation culture and practice*).

2. Evaluation schedule

Mandatory

- a) Clusters are required to prioritise and schedule evaluations in line with the recommendations and guidance within this TPG, in response to initiative management needs and in discussion with NSW Treasury.
- b) The government may require new or ongoing initiatives to be evaluated, considering factors such as the Premier's priorities or Government commitments, size and risk, and external reporting requirements.
- c) Clusters should prioritise evaluation of new and ongoing initiatives of significant size, government priority, and risk.
- d) Under the [NSW Gateway Policy \(TPG 22-12\)](#) assurance framework, a Gate 6 post-implementation assurance process may be applied, with benefits realisation and evaluation processes subject to independent expert review.

Recommendations

It is recommended that cluster lead agencies prepare and manage an internal forward evaluation schedule that strategically prioritises evaluation activity across their cluster.

- Schedule evaluations to ensure that evidence will be available to inform investment decision-making (including decisions about continuing, expanding, changing, or ending an initiative).
- Consider the larger portfolio of initiatives and uniqueness of the investment when selecting initiatives for evaluation.
- Plan for evaluation of **new** initiatives as relevant to the key stages of initiative implementation and of outcomes and benefits realisation.
- Plan for regular examination of **ongoing** (recurrent) initiatives. Aim to evaluate ongoing initiatives, within every five years (as practicable), as part of a continuous improvement cycle.
- Consider the role for evaluation, where performance monitoring, including through Outcome Budgeting processes, indicates:
 - progress is not on track to meet targets set by the cluster at the initiative, program, or State Outcome level
 - there is uncertainty regarding the alignment and contribution of an initiative to relevant State Outcomes
 - there is opportunity to examine best-practice and lessons learned.
- Incorporate evaluation planning into relevant cluster and agency business planning processes.

Guidance/Information

Be strategic about evaluation

Decisions need to be made about what will be evaluated. Initiatives should be prioritised for evaluation considering existing evidence and previous evaluation activity, and their:

- size — total cost of initiative (or relevant group of initiatives)
- strategic priority — role in achieving State Outcomes and cluster objectives
- risk — uniqueness and delivery complexity.

Evaluation is good practice regardless of the size of an initiative. Evaluation can be undertaken strategically across the activities of a cluster to inform internal decision-making and to build an evidence-base. A cluster may plan evaluations, for example, to:

- assess alignment of an initiative with government or community priorities
- investigate best-practice and share learnings
- examine how a portfolio of initiatives with a shared objective is performing
- support decision-making regarding an initiative.

Coordinate monitoring and evaluation

Evaluation will complement and be supported by other forms of performance monitoring and measurement. Evaluation is most effective when supported by a continuous and systematic process of collecting and analysing information on the implementation and impacts (outcomes and benefits) of an initiative.

Use monitoring to track the activities and outputs of an initiative and observe its outcomes and benefits. Monitoring processes can be designed to report on performance, inform ongoing improvement, and support evaluation (see *Evaluation design*).

Evaluation will usually investigate a broader range of information on initiative performance and context than is feasible to monitor on an ongoing basis and involves research to address specific questions regarding the performance of the initiative.

Align with agency and cluster business processes

Monitoring and evaluation are one component of a suite of processes that support government decision-making and investment. Be aware of other processes and guidelines relevant to your initiative, and align monitoring and evaluation of an initiative with other business planning and performance reporting processes across the cluster (see *Appendix C. Alignment with investment processes*).

Coordinate monitoring and evaluation with NSW Gateway Policy investor assurance timeframes as relevant to the initiative. Use the Benefits Realisation Management (BRM) framework and this TPG to inform the design of monitoring and evaluation. Where an initiative will be subject to Gateway review, the agency should ensure that BRM and evaluation planning also meet any additional requirements for the review.

Tailor evaluation to suit the initiative

An initiative may be nominated for evaluation by the government as part of funding or other agreements, by a cluster in agreement with Treasury, or through internal cluster decision-making process (nominated by the cluster or an agency within the cluster). The design of the evaluation should be appropriate to the initiative. *Table 3* and *Chapter 3* provide a guide.

Table 3: Initiative scale and evaluation design characteristics

Scale	Initiative characteristics	Evaluation characteristics
A	<p>Size: High total cost of initiative (or relevant group of initiatives) <i>or</i></p> <p>Priority: Priority initiative for NSW government e.g.:</p> <ul style="list-style-type: none"> • Premier's Priorities • election commitment • across-government initiative (multiple delivery partners) • external reporting requirements (for example, Commonwealth-State funding) <i>or</i> <p>Risk: High risk, for example.:</p> <ul style="list-style-type: none"> • Tier 1 or 2 under NSW Gateway Policy • delivery complexity (for example, multiple delivery partners) • risk to achieving State Outcomes. 	<p>Nominated by: Cabinet or Cabinet committee, Cluster</p> <p>Reported to: Cabinet or Cabinet committee. Ministers or senior executives consider and respond to findings.</p> <p>Scope: Coordinated use of monitoring and evaluation to report on outcomes, net social benefits, and value for money.</p> <p>Governance: Executive oversight. Steering committee or Advisory group to include key delivery partners and key stakeholders.</p> <p>Conduct: To support independent and robust evaluation:</p> <ul style="list-style-type: none"> • use evaluators external to the lead agency • for technically complex evaluations, consider a peer review of methods and findings by subject experts external to the evaluation team. <p>Stakeholders: Involve stakeholders impacted by the initiative in confirming and valuing benefits.</p> <p>Resources: Dedicated evaluation budget.</p>
B	<p>Size: Moderate to high total cost of initiative (or relevant group of initiatives) For example, above \$50 million (over four years for ongoing initiatives) <i>or</i></p> <p>Priority: Priority initiative for cluster (including achieving State Outcomes) <i>or</i></p> <p>Risk: Moderate to high risk, for example:</p> <ul style="list-style-type: none"> • Tier 1 or 2 under NSW Gateway Policy • delivery complexity • risk to achieving State Outcomes • limited evidence base • innovative or unique activity (for example, pilot activity). 	<p>Nominated by: Cluster, in discussion with Treasury</p> <p>Reported to: Cabinet committee. Senior executives consider and respond to findings.</p> <p>Scope: Coordinated use of monitoring and evaluation to report on outcomes, net social benefits, and value for money.</p> <p>Governance: Steering committee or Advisory group to include key delivery partners and key stakeholders.</p> <p>Conduct: To support independent and robust evaluation:</p> <ul style="list-style-type: none"> • use evaluators external to the business/delivery team for aspects of the evaluation where independence needs to be demonstrated • for technically complex evaluations, consider peer review of methods and findings by subject experts independent of the evaluation team. <p>Stakeholders: Engage with relevant internal and external stakeholders.</p> <p>Resources: Dedicated evaluation budget within initiative budget.</p>
C	<p>Size: Low to moderate cost of initiative, for example less than \$50 million (over four years for ongoing initiatives) <i>or</i></p> <p>Priority: Opportunity for strategic investigation, for example:</p> <ul style="list-style-type: none"> • opportunity to respond to changing government or community priorities • opportunity to share learnings or demonstrate best-practice • evidence required for upcoming decision-making, including funding <i>or</i> <p>Risk: Low to moderate risk, for example:</p> <ul style="list-style-type: none"> • Tier 3 or 4 under NSW Gateway Policy • Risks to achieving outputs or outcomes. 	<p>Nominated by: Cluster or agency</p> <p>Reported to: Executive leadership. Executive or Initiative managers consider and respond to findings.</p> <p>Scope: Coordinated monitoring and evaluation to provide evidence of outcomes and benefits. Ensure foundations established for economic analysis.</p> <p>Governance: Initiative manager and Advisory group.</p> <p>Conduct: Internal, or mix of internal and external, expertise. Consider peer review by evaluators external to the business/delivery team.</p> <p>Stakeholders: Engage with relevant internal and external stakeholders.</p> <p>Resources: Funded from within a larger program or Cluster evaluation budget.</p>

3. Evaluation design

Mandatory

- a) Monitoring and evaluation should be coordinated, as relevant to the initiative and the timing of the evaluation, to examine:
- implementation
 - appropriateness, including continued relevance to government and community priorities
 - efficiency
 - effectiveness in delivering outcomes
 - distribution of outcomes and benefits, and impacts on equity (where relevant)
 - net social benefit and value for money.

Recommendations

Evaluation should be tailored to the size, strategic significance, and risk of the initiative. Regardless of the size of an initiative, it is good practice to plan for monitoring and evaluation.

Coordinate monitoring and evaluation for:

- initiatives over \$10 million to provide evidence of costs, outcomes, and benefits
- initiatives over \$50 million to provide evidence of outcomes and net social benefits and assess value for money
- all Tier 1 and 2 capital infrastructure initiatives, and a sample of Tier 3 and 4 capital infrastructure initiatives, to provide evidence of outcomes and net social benefits, and assess value for money (tiering is under [Gateway Investor Assurance frameworks](#))
- all pilot initiatives to report on outcomes and benefits, and to consider value for money.

Different types of evaluation will be relevant at different stages in the initiative lifecycle.

- Use process evaluation, as relevant to initiative management needs, to examine implementation and provide information to improve delivery or to support outcome evaluation. Identify implementation costs to inform economic evaluation.
- Use robust outcome evaluation designs to provide evidence of the causal links between the initiative and outcomes, and to support evaluation of effectiveness. Use evidence of outcomes and benefits to inform economic evaluation.
- Use economic evaluation to identify and report the key economic, social environmental and cultural impacts of an initiative (costs and benefits) on social welfare. Use cost-benefit analysis (CBA) to assess the net social benefits and value for money of an initiative. Use supplementary methods to assess value for money where it is not feasible to quantify and monetise benefits.

Guidance/Information

Types of evaluation

The most comprehensive and effective plan for evaluation of an initiative will include elements of process, outcome, and economic evaluation over its life. Economic evaluation can complement different stages of evaluation. Examination of distribution and effects on equity (including achievement of equity goals where an initiative has been undertaken to promote equity) can also complement different types of evaluation, where relevant.

Process evaluation

Process evaluation examines initiative implementation and delivery. It focuses on the inputs, activities, and outputs of an initiative. It can be designed to examine the contexts in which the initiative is operating, identify who the initiative is reaching and their experiences, identify issues with delivery, review appropriateness (to needs, community and conditions), assess efficiency, and provide information for process improvements. Process evaluation can include questions about distribution and equity of access to the outputs of an initiative.

Process evaluation can be undertaken early to check if an initiative is being implemented as intended, or to support outcome evaluation by helping distinguish implementation issues from design issues.

Outcome evaluation

Outcome evaluation examines if and how an initiative is leading to intended changes. It measures the extent of change and the degree to which the initiative has contributed to the observed changes. It can also be designed to capture unintended outcomes. Outcome evaluation can include questions about distribution and equity of outcomes.

There is a range of outcome evaluation research designs that can be used, from experimental to quasi-experimental to non-experimental designs. The strongest methods for examining causal attribution will use a control group (as in randomised experimental design) or a comparison group (as in quasi-experimental design) to compare change with and without the initiative, and support conclusions about the extent to which a change is caused by the initiative.

Outcome evaluation can be undertaken when the initiative is expected to achieve outcomes (across a sufficient sample size). It will provide an important foundation for economic evaluation, including cost-benefit analysis (CBA).

Economic evaluation

Economic evaluation identifies and measures the impacts of an initiative relative to its costs, to provide an assessment of value for money or net social benefit. Economic evaluation will build on evidence collected from process and outcome evaluations, and from initiative monitoring (including under Benefits Realisation Management).

Cost-benefit analysis (CBA)

CBA is a comprehensive examination of an initiative's incremental impacts on welfare (including economic, social, environmental, or cultural impacts) and assessment of whether the investment delivered a net social benefit to the people of New South Wales. CBA is NSW Treasury's preferred method for economic evaluation (ex-ante and ex-post) as it allows for the objective comparison of net social benefits across different initiatives

and portfolios. It is recommended for large, complex, or risky initiatives.

CBA should (as relevant) include analysis of distribution and equity of costs and benefits.

CBA evaluation of an implemented initiative can be undertaken when an initiative is due to realise benefits or at its completion (these are both known as ex-post CBAs).

Where a CBA appraisal (ex-ante) was undertaken (when designing the initiative and business case), ex-post CBA can be used to test the accuracy of cost and benefit forecasts presented in the business case, and inform, where relevant, a business case variation report and a lessons learned report.

Cost-effectiveness analysis (CEA)

Cost-effectiveness analysis (CEA) (including Cost Utility Analysis (CUA)) compares costs against a single unit of an outcome and is a secondary option to assess value for money where it is not feasible to quantify or monetise benefits.

Other supplementary methods to assess value for money can be considered when needed to complement evaluation at different stages of the initiative life. For example, cost-efficiency analysis, examining costs against outputs, can complement process evaluation. Cost-consequence analysis can complement an outcome evaluation and provide an important foundation for CBA.

Benefits Realisation Management (BRM)

Benefits Realisation Management (BRM) — as a standard approach to understand, plan, manage and report, and evaluate benefits realisation — can be coordinated with evaluation to provide evidence of initiative outcomes and net benefits.

Monitoring (under a BRM benefits register) will support reporting of net social benefits where an ex-ante CBA appraisal comprehensively forecast costs and benefits, key costs and outcomes and benefits have been monitored (and are within the sensitivity testing forecasts), and there is evidence that reported outcomes and benefits are attributable to the initiative.

However, where monitoring has not provided clear evidence of outcomes and benefits, there have been substantial changes to the initiative in implementation, or there is now evidence that the assumptions of the ex-ante CBA were not sufficiently accurate — ex-post CBA would be needed to provide information of net benefits (rather than relying on BRM).

Evaluation Plan

An evaluation plan may include different types of evaluation to be undertaken at one point in time, or different evaluations scheduled over time, as relevant to the implementation and impact stages of the initiative, and the need for evaluation findings to inform decision-making (see 4. *Evaluation for new proposals*, Figure 2).

Appendix D. Types of evaluation and example questions includes more detail on the types of evaluation, and example questions related to the different types of evaluation (these lists are not prescriptive or exhaustive). Evaluation questions should be tailored to the initiative and the information needed for decision-making.

4. Evaluation for new proposals

Mandatory

- a) All proposals seeking government resources are required to include details on prior evaluation and to plan for future evaluation.

Evaluation evidence:

- for new initiatives, identify the evidence that has informed the development and appraisal of the submission
- for ongoing initiatives, identify when the initiative was last evaluated or subject to other review processes (for example, audit and assurance), the findings and recommendations, and the agency response.

Evaluation planning:

- identify when the initiative will next be evaluated, or when evaluation timing will be decided
- include high-level monitoring and evaluation planning in all submissions seeking NSW Government investment of \$10 million and above (this may be under a monitoring and evaluation framework or under a Benefits Realisation Management Framework) tailored to the size, priority, and risk of the initiative
- incorporate resourcing requirements for monitoring and evaluation into the business case funding submission (or set funding aside) or include a plan for when resourcing will be determined and obtained.

Recommendations

The monitoring and evaluation framework should incorporate:

- a logic model that identifies intended activities/deliverables and key impacts (for example, inputs, outputs, outcomes, and benefits)
- a forward evaluation plan that identifies the evaluation types and timeframes, and scopes the evaluation purpose, key questions, and design and methods
- a monitoring and data collection plan (or plan to develop a monitoring and data collection framework) to track:
 - implementation (inputs, activities, and outputs)
 - impacts (outcomes and benefits)
 - information relevant to assumptions and risks, or reference to where this information will be captured.

The level of resources required for evaluation should be determined based on the risk, size, and uniqueness of the initiative.

Guidance/Information

Prepare for evaluation

At the business case stage, under a monitoring and evaluation framework, forward planning for evaluation may be high level. More detailed individual evaluation plans can be developed prior to implementing an evaluation.

The logic model for the initiative can be used as a foundation from which to consider what will be evaluated, as well as when and how it will be evaluated (see *Figure 2* for example of types and timing of evaluation over the life of an initiative and *5.II. Monitoring and evaluation framework* for more detailed guidance).

The framework should be designed in consultation with the evaluators within your cluster, or other experts, and involve relevant stakeholders.

Monitoring and evaluation will be most effective when they are considered at the design stage of an initiative and incorporated into implementation planning.

Early planning for monitoring and evaluation can enhance the design of an initiative, by asking questions about intended outputs, outcomes, and benefits and how these will be measured (including ensuring that required data is collected).

Early planning can also increase the range of evaluation methods that can be used, and support the use of more robust designs. For example, experimental and quasi-experimental evaluation designs require early planning and the collection of baseline data. Where modelling is used for ex-ante CBA appraisal of an initiative, consider how the model could be designed to allow it to be updated for use in ex-post CBA.

Early planning will provide opportunity to include monitoring and evaluation in relevant processes such as communications planning, stakeholder engagement planning and risk management for the initiative.

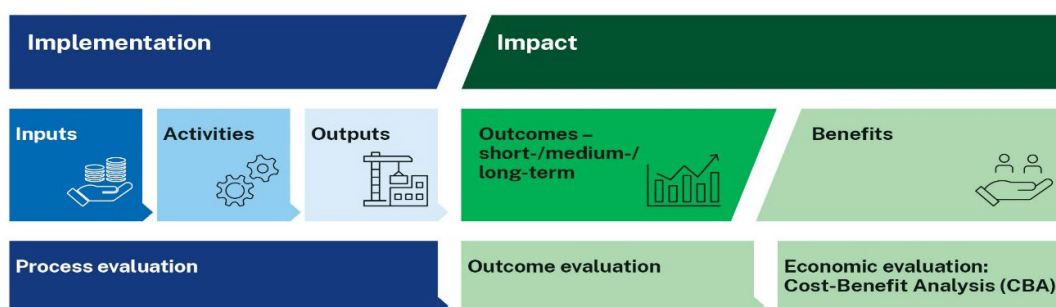
When a monitoring and evaluation framework is established during initiative design, resourcing requirements can be included in the business case funding submission.

The evaluation approach, and the time and funding required for evaluation, should be appropriate to the size, risk, and uniqueness of the initiative.

It is best-practice to estimate the resources required based on planned monitoring and evaluation activities, with reference to costs of comparable evaluations where available.

Where the costing is not based on identified activities, sufficient funds should be allocated to monitoring and evaluation. For example, it may be appropriate to allocate one to five per cent of the initiative budget for monitoring and evaluation for recurrent or regulatory initiatives, up to an appropriate capped amount. For capital initiatives, the percentage may be lower. For pilot initiatives, the percentage may be higher.

Figure 2: Types of evaluation relevant at different stages of initiative implementation



5. Evaluation practice and reporting

Mandatory

- a) Clusters, in consultation with Treasury, are required to nominate initiatives for evaluation reporting. Annual reporting to Treasury will include evaluations completed in the previous financial year, and evaluations due for completion in the upcoming financial year.
- b) Agencies are mandated to proactively and publicly release the findings of evaluations, unless there is an overriding public interest against disclosure of the information, in line with the Government Information (Public Access) Act 2009.

Recommendations

Evaluations, including those reported to Cabinet or Cabinet Committees, should include:

- a logic model
- clear evaluation purpose and key evaluation questions
- a robust design and methods
- clear presentation and assessment of results
- a statement on how results will inform future decisions.

Reporting of CBA results should include:

- description of major costs and benefits
- assessment of net social benefit for New South Wales
 - net present value
 - qualitative assessment of likely net social benefit
- benefit-cost ratio, or supporting assessment of value for money
- comparison of ex-ante appraisal and ex-post CBA results, to draw lessons learned and strengthen future ex-ante appraisals.

Guidance/Information

Evaluation should be conducted and reported in line with the principles and standards in this TPG.

Build monitoring and evaluation on clear documentation of the reason for the initiative and how the initiative is intended to lead to change, using a logic model (see 5.I. *Foundations for evaluation*).

Use a monitoring and evaluation framework to provide high-level planning for data collection and evaluation over the life of an initiative (see 5.II. *Monitoring and evaluation framework*).

Design evaluation to provide robust evidence to inform decision-making (see 5.III. *Evaluation plan: Design the evaluation*).

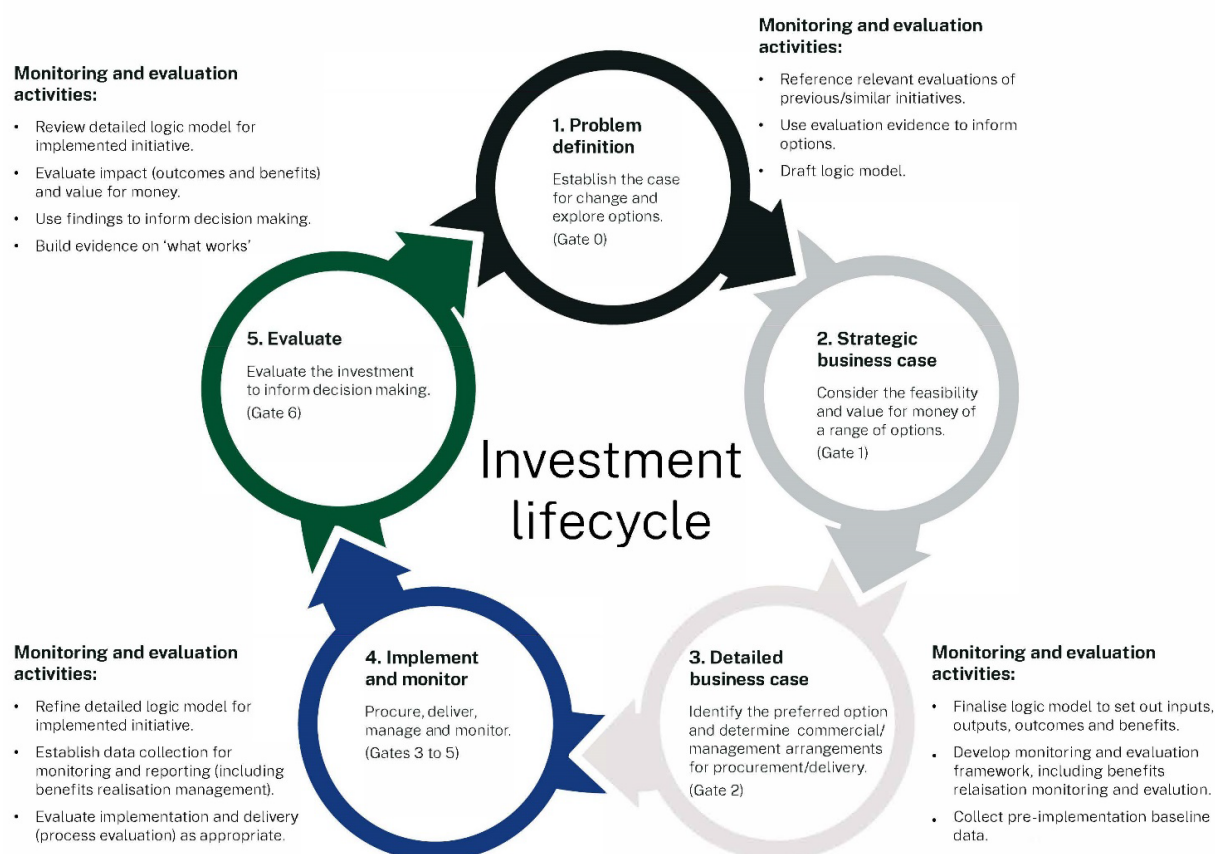
Ensure that the right people are involved in designing and planning the evaluation, including community stakeholders, and that evaluation processes and reporting meet cultural appropriateness and ethical conduct standards (see 5.IV. *Evaluation plan*):

Manage the evaluation).

Coordinate expertise internal and external (to the business and delivery teams) as required to ensure independent and robust evaluation (see 5.V. *Evaluation plan: Use the right expertise*).

Be clear regarding how monitoring and evaluation planning, practice and reporting will fit with the investment life-cycle of the initiative (from initiative conception and design, through to completion), including how monitoring information and evaluation evidence will be used to inform decision-making at different stages of the initiative (see *Figure 3*).

Figure 3: Investment lifecycle



The *evaluation workbooks* provide more detailed guidance on monitoring and evaluation activities (see *Appendix F. Evaluation workbooks summary*).

5.1. Foundations for evaluation

Key points

Monitoring and evaluation should be built on clear understanding of:

- the reason for the initiative
- how the initiative is intended to lead to change

Use a logic model to clearly set out how the initiative is intended to lead to change.

Evaluation foundations

Monitoring and evaluation will be supported by clear documentation of:

- the **reason for the initiative**, including the reason for government intervention, the objective of the initiative, the context in which the initiative operates (including links to other initiatives), and the counterfactual scenario (i.e., what would happen without the initiative)
- a **theory of change** — a summary narrative, based on evidence, logic, or theory, that describes how and why the initiative will achieve intended outcomes and benefits
- a **logic model** — a diagram that illustrates how an initiative is intended to work, by setting out inputs, activities, and outputs, and linking these with outcomes and benefits.

It is also important to identify the assumptions regarding causal links (for example, how outputs are intended to lead to outcomes, and how outcomes will produce benefits), other activities that support achieving links, and any risks to these links.

Before commencing an evaluation, revisit the reason for the initiative, theory of change and logic model presented in the business case and other supporting materials (including benefits realisation planning). Identify where these may need to be updated to fit new contexts or the initiative as implemented.

The logic model can be used to set out the expected timing of an initiative's implementation and impacts. It can also provide a foundation for scoping evaluation questions and prioritising what should be monitored.

Where an ongoing initiative does not have a current logic model, develop this as a first step in an evaluation. It is good practice for all initiatives to have current logic models as well as monitoring and evaluation frameworks. The process of developing a logic model can be used to build a shared understanding of the initiative.

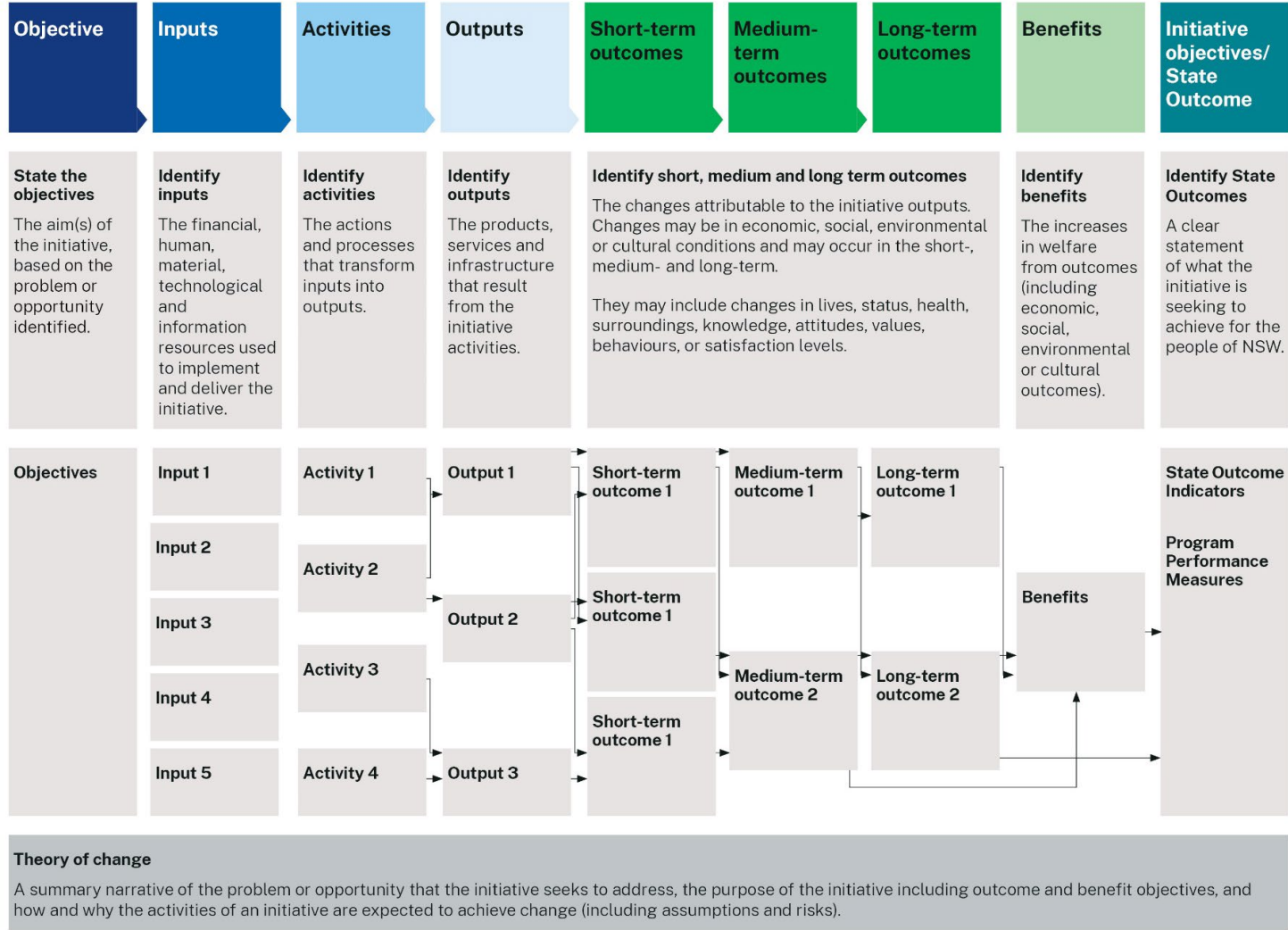
The detail and structure of a logic model will vary, depending on initiative size and complexity (*Figure 4* presents an example logic model):

- A simple logic model may include inputs, outputs, and outcomes.
- A more detailed model can identify benefits (the improvements in social welfare) that are expected to follow from outcomes.
- The logic model should identify how the initiative aligns with State Outcomes or other broader objectives.

Large or complex initiatives may have an overarching objective and include a number of inter-related sub-initiatives with individual objectives. An overarching summary logic model can be developed for the larger initiative, with detailed logic models for individual sub-initiatives (see *Workbook VIII. Complex initiatives*).

Evaluation practice and reporting

Figure 4: Example logic model



Appendix E. Example theory-of-change and logic model: building evidence presents an example theory-of-change and logic model relevant to an initiative to build an evidence bank.

5.II. Monitoring and evaluation framework

Key points

A monitoring and evaluation framework plans for collecting data and for evaluation over the life of an initiative. Use the framework to:

- scope evaluation purpose, key questions, and design and methods
- plan for evaluation at key stages and decision-points
- guide data collection for monitoring, performance reporting and evaluation
- align initiative monitoring and evaluation with broader cluster processes.

Plan for monitoring and evaluation

A monitoring and evaluation framework will plan for and guide monitoring, evaluation, reporting and improvement over the life of an initiative.

When planning monitoring and evaluation, be clear regarding what information will be required by who and by when. Ensure that the monitoring and evaluation framework meets the information needs of initiative managers and stakeholders.

Monitoring is a means to track and report on performance, support ongoing improvement, and inform evaluation. Forward planning for monitoring can:

- provide a focus on performance and outcomes
- build feedback processes that support adaptive management
- anticipate the data required for evaluation, including ensuring baseline data are collected.

Evaluation can be scheduled to investigate if the initiative is being implemented as intended, and if it is achieving expected outcomes and benefits. Forward evaluation planning should:

- scope key evaluation questions, design, and methods
- plan for evaluation considering the lifecycle of the initiative and the information required for decision-making
- identify 'owners' responsible for evaluation.

Ensure that monitoring and evaluation, and different types of evaluation, are coordinated over the life of the initiative. For example, consider where process evaluation should identify the monetary and non-monetary implementation costs of an initiative to inform economic evaluation. Outcome evaluation provides an important foundation for CBA; consider how the outcome evaluation could be designed to provide evidence of outcomes and benefits that will support an economic evaluation.

Consider also how monitoring and evaluation of the initiative will align with broader program or cluster level monitoring and evaluation frameworks.

Use the logic model

Use the **logic model** as the foundation for the monitoring and evaluation framework. To support planning for evaluation:

- identify the intended timeframes for initiative implementation and for realising outcomes and benefits (and consider when these can be evaluated)
- determine the purpose of evaluation at different stages of the initiative's life, scope the key evaluation questions to be addressed at these stages and choose evaluation methods that will address these questions (see *Figure 5*), and

- consider:
 - the key activities, outputs, outcomes and benefits that should be tracked
 - what other information should be monitored to address key evaluation questions.

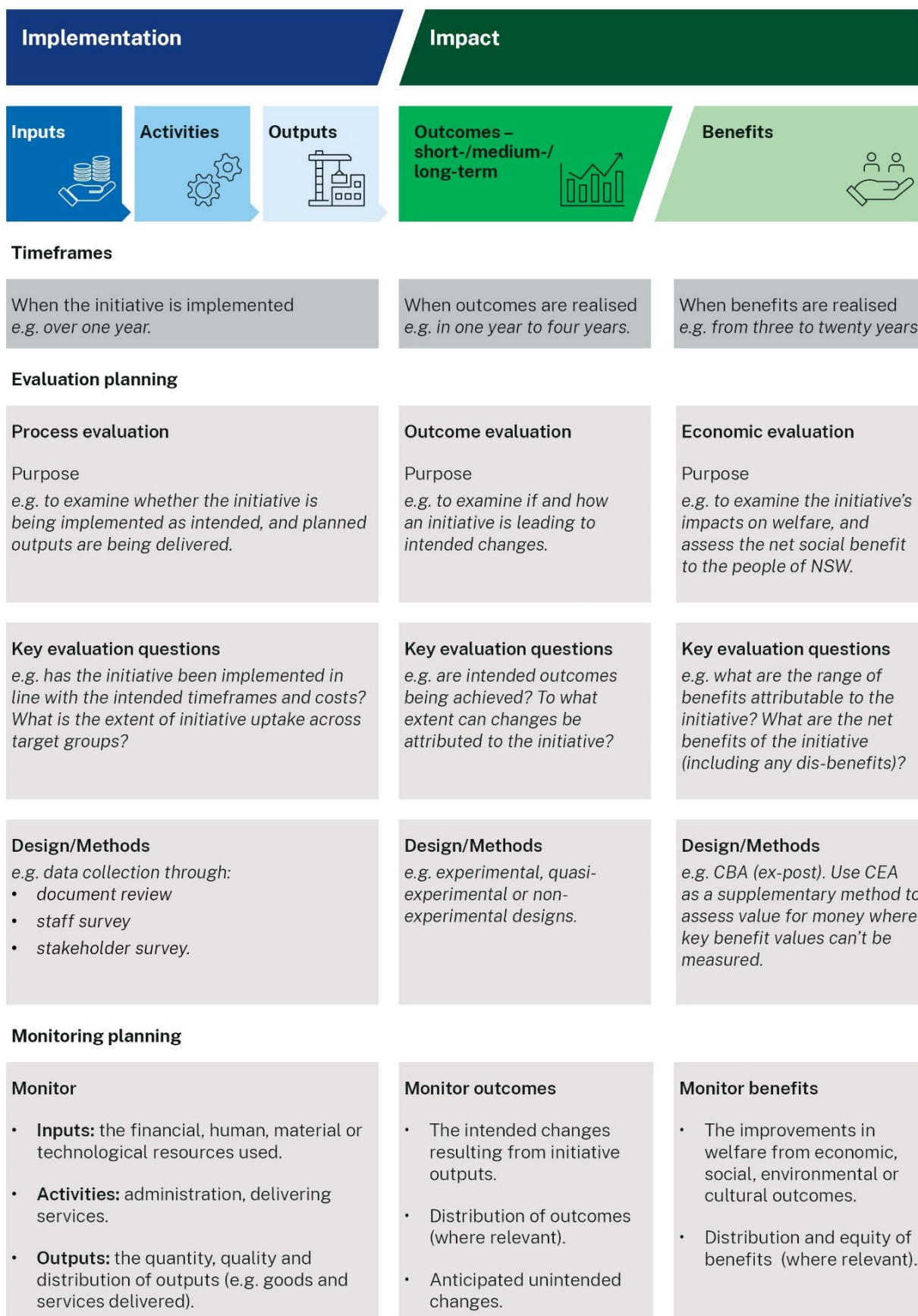
Include the monitoring and evaluation framework in the initiative business case

See *Table 4* for an example checklist for monitoring and evaluation planning relevant to a business case submission.

Table 4: Guidance - example business case checklist for monitoring and evaluation planning

Include:	Guidance:
Logic model	<ul style="list-style-type: none"> • Present the objective(s) of the initiative • Identify the initiative inputs, outputs, outcomes and benefits • Identify timeframes for implementation, and for realisation of outcomes and benefits
Forward evaluation plan and schedule	<ul style="list-style-type: none"> • Plan for evaluation (process, outcome and economic) considering: <ul style="list-style-type: none"> ○ the implementation timeframes and expected timing of impacts ○ the information required for decision-making • Scope evaluation purpose, key evaluation questions and methods (to frame monitoring) • Identify 'owners' responsible for evaluation
Monitoring and data collection framework	<ul style="list-style-type: none"> • Identify key measures/indicators to monitor for: <ul style="list-style-type: none"> ○ implementation (inputs, activities and outputs) ○ impacts (key outcomes and benefits) ○ assumptions and risks • Plan for data collection and storage (as part of the implementation plan)
Cost monitoring and evaluation requirements	<ul style="list-style-type: none"> • Identify key activities and the resources required for: <ul style="list-style-type: none"> ○ ongoing monitoring ○ scheduled evaluations • Include costings in the funding request (or plan for when evaluation costs will be determined, and resourcing will be obtained).

Figure 5: High level monitoring and evaluation planning



5.III. Evaluation plan: Design the evaluation

Key points

Evaluations should provide robust evidence to inform decision-making. When designing the evaluation, undertake to:

- understand the context of the evaluation
- design your evaluation in response to a clear evaluation purpose and key questions
- ensure evaluations are as rigorous as possible, aim to produce valid and reliable findings, and reach sound conclusions
- ensure that the selected methods can be applied within the available time and resources
- plan to present and assess evaluation results and conclusions.

Involve the right people in designing and managing the evaluation (see 5.IV. *Evaluation plan: Manage the evaluation*). Consider also who else may bring value to the evaluation design and implementation (see 5.V. *Evaluation plan: Use the right expertise*).

Establish context for the evaluation

Evaluation should commence with a clear understanding of why evaluation is being undertaken (at this point in time), what evaluation has already been undertaken, what wider evidence is available on the performance of this or similar initiatives, and who will receive and use the evaluation findings.

Design the evaluation

Evaluation purpose

Each evaluation should be undertaken with a view to informing decision-making. Reference the logic model to consider what aspects of the initiative the evaluation will examine, and where the initiative will inform decision-making regarding the initiative. Define the evaluation purpose in terms of what evidence is needed and what decisions it is likely to inform (the purpose of the evaluation will be aligned to, but distinct from, the initiative objectives). Ask questions such as:

- why does the evaluation matter?
- to whom will the findings be important and why?
- when are the findings needed?
- how will the findings be used?

Evaluation questions

Evaluation questions should be designed to address the evaluation purpose and provide information that will meet the needs of decision-makers and stakeholders.

Focus the evaluation on a small number of higher order evaluation questions (three to seven), to ensure that the evaluation is focused on key concerns, and that the scope is manageable.

In refining evaluation questions, consider:

- the initial questions in the monitoring and evaluation framework
- the current information needs of decision-makers and managers
- the data collected from monitoring, and any trends, issues or gaps for further examination
- who is impacted by the initiative and current stakeholder concerns.

The evaluation may also need to address additional questions according to government priorities or partnership agreements.

Evaluation design and methods

All evaluations should aim to be rigorous, produce valid and reliable findings, and reach sound conclusions.

The design used for evaluation will determine the methods to be used, the data collection that will be required, and the strength of conclusions that can be made regarding the attribution of the initiative to outcomes and benefits.

Different types of evaluation will use different evaluation designs and methods. For example, an outcome evaluation using an experimental design may be supported by case-study research. It is likely that more than one method will be required, with evidence from different data sources used to answer the evaluation questions. The evaluation resources include more detailed technical information on different evaluation designs and methods.

The evaluation should be fit for purpose. This means the evaluation design should be tailored to the scale of the initiative and the significance of the evaluation. The design should use methods that will provide credible and comprehensive evidence, while ensuring that the selected methods can be applied within required timeframes and allocated resources.

Where required data are unavailable and there is no suitable substitute, work closely with evaluation experts to determine to what extent evaluation will be feasible. In such cases, use the evaluation to inform recommendations for a forward data collection strategy.

Reflect on evaluation design, methods and evidence

The evaluation design, and the methods used to gather evidence, will determine the validity of evaluation conclusions.

At key stages of the evaluation, and as a final assessment process undertaken as part of the evaluation, consider the:

- strengths of the evaluation design and method, as implemented
- robustness of evidence gathered
- validity of evaluation conclusions.

This assessment should be included in the evaluation report (see 5.VI. *Evaluation plan: Report and use evaluation findings*).

5.IV. Evaluation plan: Manage the evaluation

Key points

Develop a project plan as you design the evaluation. The plan should:

- ensure that the right people are involved in designing and planning the evaluation (including community stakeholders) and establish clear roles and responsibilities
- consider cultural appropriateness of data collection and evaluation and activities
- uphold ethical conduct requirements
- partner and collaborate with First Nations people and communities in evaluating initiatives delivered to or that impact upon them.

The project plan can be designed to draw upon a mix of internal and external expertise (see 5.V. *Evaluation plan: Use the right expertise*).

Establish a project plan

Evaluation project management planning should be undertaken at the same time as designing the evaluation. Include the activities required for evaluation design, delivery, analysis and reporting. Also plan for governance, stakeholder consultations and collaboration, ethics requirements, peer review, and procurement (as needed). Establish an owner for the plan.

The project plan will set the timetable for the evaluation activities and should establish clear milestones and deliverables. Identify when findings are needed to inform decision-making. Allow adequate time for collaborative processes, approval or authorisation processes, and data collection (consult with data custodians early to ensure data will be available and fit for purpose). Build flexibility into the evaluation timetable to allow for unforeseen events or adjustment of methods.

The project plan should also identify key risks that the evaluation may face (e.g., for stakeholder engagement, commissioning, access to data, ethical conduct and timelines) and establish strategies for ongoing monitoring and management of risks.

Involve the right people

The evaluation manager should ensure that the right people are included in the evaluation design and conduct. Identify key stakeholders who will hold an interest in the evaluation. An evaluation's relevance and recommendations will be stronger when designed with the people who are involved with initiative delivery and with the people who are the target of the initiative or otherwise experience its impacts (this can include customers or clients who are intended to benefit, or anyone who otherwise experiences positive or negative impacts, from the initiative).

Evaluation governance: Ensure that stakeholders are involved in evaluation governance. Always involve the initiative manager and delivery partners in evaluation design and management planning. For a large or complex evaluation, governance can include a steering committee to provide direction or contribute to decision-making at key stages of the evaluation.

Stakeholder collaboration: Identify when and how stakeholders will collaborate in evaluation processes (e.g., developing evaluation questions, designing the evaluation, valuing benefits and testing the practicality of recommendations). Stakeholder representatives can participate in advisory groups or be more directly involved through inclusive processes (see *Table 5*). Allow adequate time for meaningful collaboration. Be transparent about the scope of the evaluation and what it can achieve.

Project teams: Establish clear roles and responsibilities within project teams. Be clear regarding how decision-making processes work.

Meet relevant standards

Ensure that processes for working with people and communities meet relevant standards related to cultural appropriateness and ethical conduct.

Cultural appropriateness in evaluation

Evaluation activities should recognise, respect and be responsive to the culture and values of participants, communities and the cultural setting in which the initiative operates. Cultural appropriateness in an evaluation involves recognising and acknowledging differences in knowledge, language, norms, beliefs, practices, family and community structure and dynamics, and priorities.

Consider cultural appropriateness at the inception of an initiative and in all evaluation activities, including evaluation planning, design, management, data collection (including consultations), analysis and reporting findings. Use up-to-date guides and standards from within the cluster or representative organisations.

Ethical conduct in evaluation

Ethical conduct should be a key consideration in all monitoring and evaluation processes. Identify what ethical issues need to be addressed, and where ethics clearance will be required. In cases where formal ethical review is not required, ethical practices should still be followed.

All activities, such as data collection, consultation and reporting, need to protect the privacy, dignity and confidentiality of participants and consider any culturally sensitive information. Ethics review and approval are required, and evaluations should be overseen by a Human Research Ethics Committee (HREC), where there is a potential risk of harm to people participating in the evaluation, as informants or as evaluators. Types of harm can range from loss of privacy or benefits to participants, damage to vulnerable groups, or physical, social or mental harm. Allocate adequate resources and time for ethical conduct approval processes.

Several organisations have developed ethical guidelines and standards to support researchers and evaluators, including the [National Health and Medical Research Council \(NHMRC\)](#) and [Australian Evaluation Society \(AES\)](#) (see also *Resources. Useful references*). Refer to the ethics policies or guidelines of your cluster.

Work with First Nations people and communities

Put First Nations people and communities at the centre of the evaluation of initiatives that are delivered to or that impact upon them. Where consistent with the preferences of the community, collaborate with community representatives in evaluation including in identifying priorities, co-design, governance, delivery, data management and critical self-reflection on processes and findings. Allow adequate time for genuine and respectful communication and collaboration, and for building evaluation capabilities where appropriate.

Recognise and respect cultural and intellectual knowledge and property. The way First Nations people conceive and recognise their cultural and intellectual contributions to evaluation may differ to non-indigenous approaches. Ensure issues of data sovereignty and governance are considered in consultation with key stakeholders. Discuss agreements related to the contribution of expertise and cultural and intellectual property rights upfront.

Follow guidance relevant to First Nations people and communities. Refer to the [NSW OCHRE Plan \(2013\)](#), [Commonwealth Productivity Commission \(2020\) Indigenous Evaluation Strategy](#), and the [Australian Evaluation Society \(2021\) First Nations Cultural Safety Framework](#). Evaluators should adhere to a number of ethical principles identified by peak research bodies in Australia, including the [Australian Institute of Aboriginal and Torres Strait Islander Studies \(AIATSIS\) Code of Ethics for Aboriginal and Torres Strait Islander Research](#), and the [NHMRC Ethical conduct in research with Aboriginal and Torres Strait Islander Peoples and communities](#).

Table 5: Stakeholder interest and potential roles in evaluation

Stakeholder	Interest in evaluation	Potential roles in evaluation
Executive leadership	Senior executives are responsible for ensuring the delivery of outcomes, identifying the decisions that will be informed by the evaluation, communicating findings to Ministers, and ensuring that findings are acted upon.	<ul style="list-style-type: none"> • Endorse and advocate evaluation activity • Establish expectations of the evaluation • Approve funding • Approve public release of reports • Use the evaluation to inform decisions regarding the initiative
Initiative manager	The initiative manager holds responsibility for delivery of outcomes and for implementing change.	<ul style="list-style-type: none"> • Develop logic model (with key stakeholders) at initiative start • Scope focus of evaluation, evaluation questions, and potential risks • Appoint evaluation manager (including commissioning where used) • Ensure feedback to executive and initiative stakeholders • Be transparent regarding response to recommendations • Participate in the evaluation steering committee
Cluster evaluation unit/ evaluators	Cluster evaluation units or internal evaluators may be well positioned to manage the overall evaluation process and governance arrangements and provide an independent (from the project and delivery team) perspective. Engage technical experts, such as from the economics unit within a cluster, to contribute to components of evaluation.	<ul style="list-style-type: none"> • Provide advice on: <ul style="list-style-type: none"> ◦ evaluation standards and design ◦ performance reporting requirements across the cluster ◦ other appraisal/evaluation activities across the cluster ◦ ethical standards and required approval processes • Develop evaluation design (with initiative manager and team) • Undertake evaluation or assist in commissioning (e.g., assist with draft tender terms of reference (ToR) and project management) • Participate in evaluation steering committee or an advisory group • Provide peer review
Delivery team	People involved in delivering the initiative will have important background knowledge to the initiative and practical experience in considering what is and is not working. Involving the delivery team can focus the evaluation on important concerns, ensure that evaluation design and stakeholder engagement strategies are feasible, and support retention of learnings within the agency.	<ul style="list-style-type: none"> • Provide input to evaluation design, (e.g., develop or review logic model, identify knowledge gaps and scope evaluation questions) • Provide data from administrative systems • Coordinate collaboration with stakeholders (e.g., workshops) • Undertake data collection • Participate in aspects of evaluation (e.g., process evaluation) • Act as an audience to test preliminary evaluation findings and the feasibility of recommendations • Implement endorsed recommendations
Subject matter experts	Academic and other professionals can contribute evaluation or subject matter expertise and will have an interest in evaluation design and findings.	<ul style="list-style-type: none"> • Provide evaluation or subject matter expertise • Participate in an evaluation advisory group • Provide peer review
Customers/ clients and community	Initiative recipients, and people who otherwise experience the impacts of an initiative, will have an interest in evaluation design, findings and recommendations. Collaborate with stakeholders to ensure that the evaluation concerns and design are appropriate to the community. Identify where and how inclusive and participatory engagement can be established.	<p>Key points for involvement/co-design may be:</p> <ul style="list-style-type: none"> • developing the logic model • identifying outcomes and benefits that matter • developing and testing evaluation questions • selecting measures of outputs, outcomes and benefits • designing the evaluation • identifying risks • providing responses to evaluation questions • testing outcomes assumptions • testing benefits assumptions and estimates • designing recommendations in response to evaluation findings.

5.V. Evaluation plan: Use the right expertise

Key points

The evaluation workplan can be designed to draw upon a mix of internal and external expertise. External evaluators can be commissioned for any stage of the evaluation project plan. Provide a clear Request for Quotation, to ensure the tender is effectively designed and costed.

Independent evaluators (external to the business area and delivery teams) can be engaged to deliver the evaluation or aspects of the evaluation. This may be an evaluation unit or other teams with evaluation expertise within the cluster, external consultancy teams or academic teams.

Independent evaluators can be brought in at any appropriate point in the evaluation process, for example, to assist with refining the initiative logic model, designing monitoring and data collection, planning the evaluation, designing evaluation questions, selecting evaluation methods, identifying risks, managing data collection, or coordinating evaluation across larger scale or complex initiatives.

Well managed partnerships can bring flexibility to the evaluation, provide an external perspective, assist in meeting a significant workload or tight deadlines and build evaluation knowledge within the agency. The allocation of roles can vary. Consider the:

- required technical skills or subject matter expertise
- credibility of the evaluator in the eyes of the intended audience
- need to establish the independence and objectivity of the evaluation
- time and resources available to support evaluation.

Cultural expertise may be required to ensure culturally appropriate design and implementation. Where an initiative that affects First Nations people and communities is being evaluated, identify opportunities for First Nations businesses to design and conduct the evaluation, or for First Nations evaluators to contribute at key points (see also 5.IV. *Evaluation plan: Manage the evaluation*).

Where specialist expertise is required, ensure it is brought in early enough to contribute to decision-making on evaluation design, planning and other key components of the evaluation. For example, where an experimental design is to be used, early planning will be required to establish comparison/control groups, collect baseline data and establish data linkages (as required).

The engagement can be designed to support learnings within the agency. This might include, for example, debriefing by the evaluator about the evaluation process at key points or involving initiative staff in key data collection activities (such as fieldwork). An internal evaluation coordinator can be used to facilitate staff involvement.

Peer review processes can provide quality assessment for internally and externally conducted evaluations. Peer review may be undertaken by internal or external experts, depending on the profile of the initiative and the complexity and risk of the evaluation. It can be used before finalising key stages of the evaluation, such as the design, data analysis or ex-post CBA benefits framework, and in relation to aspects of the evaluation such as cultural appropriateness.

When undertaking **procurement**, comply with NSW Government and your own agency guidelines and policies. Consult the [NSW Government Procurement Policy Framework](#). Contact your cluster's internal procurement team for specific guidance.

In a procurement Request for Quotation (RfQ), include all the information a consultant will need to effectively design and cost their proposal, including information on the initiative, context for the evaluation, evaluation requirements, what data are available and what expertise is required. *Table 6* provides information that may be useful to include when completing the [Buy NSW Request for Quotation – Part A – specification template](#). Consider where the commissioning agency should own the intellectual property rights for data collected and models developed.

Table 6: Example information for consultancy brief

Include:	For example:	Provide information regarding:
Background	Initiative summary	<ul style="list-style-type: none"> The objectives of the initiative Key stakeholders of the initiative How the initiative links with other initiatives to deliver outcomes and benefits
	Evaluation purpose	<ul style="list-style-type: none"> Why the evaluation is being undertaken What decisions the evaluation will inform The target audiences for the evaluation
	Initiative appraisal/ evaluation history	<ul style="list-style-type: none"> The (pre-implementation) appraisal undertaken when designing the initiative Previous evaluation of the initiative that has been undertaken The information or reports that will be made available to the consultant
Scope of work: Evaluation requirements	Project description	<ul style="list-style-type: none"> The types of evaluation that will be undertaken (e.g., process, outcome and CBA) What is in-scope for the evaluation (and what may be out-of-scope) Expectations related to communications and collaboration with stakeholders
	Issues/risks	Challenges or risks that may impact the evaluation, for example: <ul style="list-style-type: none"> monitoring history and availability of data delivery complexity (e.g., multiple delivery partners) ethical conduct approval timelines potential challenges to stakeholder communications and collaboration
	Key Deliverables and Dates	Deliverables: <ul style="list-style-type: none"> key deliverables (e.g., evaluation plan, data sets, evaluation report(s) and presentations) how findings will be shared with stakeholders (e.g., presentations and publications) Dates: <ul style="list-style-type: none"> when the evaluation findings are needed to inform decision making key milestones Be clear about who will retain the intellectual property produced during the evaluation (including to data and modelling)
	Key stakeholders	<ul style="list-style-type: none"> The people and groups who will have an interest in evaluation findings The people who will be involved in the evaluation (and their roles and responsibilities) The expectations related to stakeholder engagement, for example: <ul style="list-style-type: none"> steering committee advisory groups collaborative stakeholder processes Any expectations regarding involvement of the delivery team or other agency staff in evaluation processes
	Agency's obligations	<ul style="list-style-type: none"> The data and information that will be provided to the consultant The support that will be provided by agency staff

Evaluation practice and reporting

Include:	For example:	Provide information regarding:
Initial assumptions OR Identify evaluation tasks	Logic model	<ul style="list-style-type: none"> The logic model for the initiative <i>OR</i> <i>Assist with developing or refining a logic model</i>
	Counterfactual	<ul style="list-style-type: none"> What would be expected to happen in the absence of the initiative <i>OR</i> <i>Determine a realistic counterfactual</i>
	Key evaluation questions (KEQs)	<ul style="list-style-type: none"> The evaluation questions <i>OR</i> <i>Develop KEQs</i>
	Evaluation methods	<ul style="list-style-type: none"> The methods considered suitable and feasible to answer the KEQs <i>OR</i> <i>Propose and develop evaluation methods</i>
	Outcomes and benefits	<ul style="list-style-type: none"> The measurable outcomes and benefits <i>OR</i> <i>Scope or measure outcomes and benefits</i>
	Stakeholders	<ul style="list-style-type: none"> The key stakeholders who should be included in the evaluation <i>OR</i> <i>Undertake stakeholder mapping or develop a stakeholder management plan</i>
	Data	<ul style="list-style-type: none"> Other data sources that are suitable and available <i>OR</i> <i>What data the consultant is expected to collect</i>
	Links to other initiatives	<ul style="list-style-type: none"> The broader context in which the initiative operates <i>OR</i> <i>Identify other activities (State, national or local) that interact with, complement or replicate the initiative</i>

See *Workbook V. Evaluation plan: Use the right expertise* for detailed guidance regarding:

- designing evaluation to use external expertise
- key criteria to consider in choosing a consultant to conduct an evaluation.

5.VI. Evaluation plan: Report and use evaluation findings

Key points

Plan to report and use evaluation findings:

- a process for reporting and responding to the evaluation findings should be planned for in advance of the evaluation.
- evaluation evidence should be used to inform the design and appraisal of future initiatives.
- lessons learned from evaluation will be of interest to a range of audiences.

Plan to report the evaluation

Plan to report evaluation findings and conclusions, including summary information on the evaluation design and management processes. Include an assessment of the strengths of the evaluation method(s) and evidence, and of the validity of evaluation conclusions. The report may also include recommendations for further actions.

Consider who the evaluation findings will be reported to, and how this information will be used. The evaluation is likely to be of interest to a range of stakeholders. Different reports can be developed to suit the information and communication needs of different audiences. Consider differences in language and literacy levels, as well as access to communication channels.

In presenting the evaluation report, the agency may separately provide comment on the evaluation findings and recommendations and publicise the next steps for the agency.

Plan for evaluation to be used in decision-making

A formal process should be established for considering the evaluation findings and developing a response. If early findings can be used to inform initiative improvements, consider sharing these findings in less formal ways.

Develop a plan for responding to the evaluation results and recommendations. Assign response 'owners' and establish mechanisms to track progress against key indicators.

Evaluation also presents the opportunity to improve the assumptions used in the initiative design, logic model and appraisal. Where evaluation is undertaken as part of Benefits Realisation Management, it will provide evidence to update the benefits register and benefits report and, where relevant, inform a business case variation report and lessons learned report. A discussion of ex-post CBA results, relative to the ex-ante CBA forecasts, should be provided where there are significant differences that warrant investigation. Use learnings to inform the design and appraisal of future initiatives.

Share learnings

Aim to share learnings from evaluation widely, across the agency and cluster, and with other organisations.

A key principle in this TPG is that evaluation processes should be transparent, and findings should be publicly released. Agencies are mandated to proactively and publicly release the findings of evaluations unless there is an overriding public interest against disclosure of the information (see Requirement 5.b) in this TPG).

Best-practice involves publication on the commissioning agency's website of the full report and methods used. Agencies may also choose to release a plain English executive summary and statement of findings on their website.

To support the use of evidence in ongoing decision-making, build a central repository of learnings and evidence of 'what works'.

Sharing findings can:

- enhance accountability and transparency
- improve awareness of the outcomes and benefits delivered by initiatives
- contribute to building an evidence base to inform decision-making
- reduce duplication and overlap of evaluation efforts across NSW Government
- support discussion of standards of evaluation and use of evidence.

Example evaluation report

Figure 6 presents an example template for a comprehensive evaluation report. Not all evaluation reports would necessarily contain each element listed. The focus of the report will be *Evaluation deliverables*, with content defined by the design of the evaluation and the information produced. see *Workbook VII. Example evaluation report template* for more details.

Figure 6: Example evaluation report

1. **Executive summary**— initiative description, evaluation purpose, design and methods, findings and recommendations (one to three pages, designed as a stand-alone summary)
2. **Background**
 - Reason for the initiative
 - Initiative characteristics
 - Initiative logic (theory-of-change/logic model)
3. **Evaluation context**
 - Reason for the evaluation (at this point in time)
 - History of evaluation
 - What is already 'known' (data and literature review)
 - Evaluation stakeholders
4. **Evaluation purpose and scope**
 - Purpose of the evaluation
 - Key evaluation questions
 - Evaluation design and methods
5. **Evaluation deliverables**
 - **Evaluation results**
The evidence that the evaluation provides to address the evaluation questions
 - **Evaluation conclusions**
For example, assessment of the degree to which the initiative is implemented as intended, appropriate, efficient, effective or providing net benefits and value for money
 - **Assessment of evaluation methods and evidence**
The strengths and weakness of evaluation design and methods, and how these have impacted results and the ability of the evaluation to answer the evaluation questions
 - **Recommendations**
Suggested responses to evaluation findings (if relevant to the evaluation design)

Appendix: Evaluation workplan information

Separate report: Agency response (where appropriate)

Separate report: Business Case variation and lessons learned report (where relevant)

Definitions

Term	Definition
Activities	The actions and processes of an initiative that transforms inputs into outputs.
Appropriateness	In evaluation, appropriateness may relate to the extent to which an initiative addresses an identified need, aligns with government priorities, or is suitable for a community or context (including cultural appropriateness and ethical considerations).
Attribution	The extent to which a change is caused by, or attributable to, an initiative (relative to other external factors/causes).
Benefit	An increase in welfare associated with an initiative's outcomes (including economic, social, environmental or cultural outcomes). Benefits need to be first be understood as changes in condition, i.e. outcomes. In CBA, benefits are a measure of the value of the outcomes of an initiative to the NSW community – they may be monetary or non-monetary (methods exist to monetise non-market benefits).
Benefits Realisation Management (BRM)	A process of identifying, organising, managing and measuring benefits so that intended benefits are achieved. It is a continuous process running through the whole life of the initiative. See Benefits Realisation Management Framework .
Benefit-Cost Ratio (BCR)	The ratio of the present value of total benefits to the present value of total costs.
Benefits Register	A register encompassing all quantitative and qualitative benefits. It is a key component of BRM and can also be called a Benefit Tracking Tool.
Comparison group	A group of participants in the study, randomly selected from a sample, who are not exposed to the initiative being evaluated.
Control group	A group of participants in the study, selected from a sample using techniques other than randomisation, who are not exposed to the initiative being evaluated.
Cost-Benefit Analysis (CBA)	An evidence-based method for systematically organising and presenting the impacts of initiatives on society's welfare, including economic, social, cultural or environmental impacts. A CBA supports assessment of the net benefits relative to a situation without the proposed initiative, and comparison of net benefits across different initiatives, by estimating costs and benefits in monetary terms. See TPG23-08 NSW Government Guide to Cost-Benefit Analysis .
CBA ex-ante	A CBA appraisal undertaken prior to the implementation of an initiative, to compare options and to assess the net social benefits of a proposed initiative.
CBA ex-post	A CBA evaluation that is undertaken when an initiative is underway (interim or 'in media res' ex-post CBA) or completed (final ex-post CBA) to assess the net social benefits of an implemented initiative.
Cost Consequence Analysis (CCA)	A form of economic evaluation where the costs of an initiative are measured and presented alongside multiple outcomes attributed to the initiative. Where accepted in a discipline (for example, in health economics) CCA can contribute to developing evaluation findings. CCA is not a preferred method of economic evaluation in this TPG because it does not produce a singular cost-to-impact ratio (e.g., a cost-effectiveness ratio or a benefit-cost ratio). CCA can provide a foundation for a CBA by capturing a range of outcomes that can be quantified and valued in monetary terms as benefits or described qualitatively.
Cost Effectiveness Analysis (CEA)	A form of economic evaluation that estimates costs in monetary terms against a single unit of an outcome.
Costs	Costs are the resources used to implement and deliver the initiative. Costs include the direct and indirect costs (monetary and non-monetary) of implementing an initiative. In CBA, costs also include decreases in welfare resulting from an initiative's outcomes (including economic, social, cultural or environmental). Under the BRM framework, negative impacts on welfare are also described as dis-benefits.
Counterfactual	The 'without' scenario or base case, is what would be expected to happen had the initiative not been implemented – usually being the next most likely scenario to have happened – such as a 'business as usual'/'no policy change' scenario or 'next most likely' option to have been implemented.

Definitions

Term	Definition
Cultural appropriateness	The extent to which evaluation activities recognise, respect and are sensitive and responsive to the culture and values of the participants and communities in which the initiative operates.
Dis-benefit	A decline in welfare resulting from an outcome, perceived as a negative by one or more stakeholders.
Economic Evaluation	An evaluation that measures value for money of an initiative. It identifies and measures an initiative's costs relative to its impacts. CBA is NSW Treasury's preferred method of economic evaluation as it supports assessing the net impacts on an initiative on social welfare (relative to a situation without the proposed action), and comparison of net benefits across different initiatives (see Cost-benefit analysis). Other forms of economic evaluation may compare an alternative action to achieve the same output (see Cost-efficiency analysis) or outcome (see cost-effectiveness analysis).
Effectiveness	The extent to which an initiative achieves its objectives (usually related to outcomes).
Efficiency	The extent to which an initiative delivers outputs relative to cost. It can also be understood as the extent to which an initiative delivers optimal outcomes relative to cost.
Equity goals	Goal(s) to reduce identified disparities between sub-groups within a population. Disparities may be based on characteristics such as disability status, income, geography/region, ethnicity, religion, gender or age.
Evaluation	A systematic and transparent process that can be used to assess the appropriateness, efficiency, effectiveness or net social benefits of an initiative.
Evaluation design	The overarching design that determines the methods to be used, the data collection required and the level of attribution which can be validly applied to an initiative.
Evaluation methods	The tools (e.g. surveys) and procedures (e.g. recruitment) used to collect and analyse data.
Evaluation plan	A plan that outlines how an evaluation will be designed and implemented.
Evidence	Information or a body of facts, quantitative or qualitative, can be used to assess the validity of a proposition or inform a decision. To be considered robust, evidence should be relevant to the context, credible, accurate and complete.
Experimental Design	Experimental designs, or randomised control trials (RCTs), randomly assign participants to either an intervention group, or to a control group (a group not exposed to the initiative) to test for causal attribution of an initiative to outcomes.
Gateway Assurance System	The NSW Government project assurance process that mandates independent peer reviews by experts at defined critical decision points (or gates) in the life of an initiatives (capital, ICT and recurrent).
Initiative	Any capital, ICT or recurrent, regulatory or policy action undertaken with the aim to achieve outcomes for the NSW community. This can be a program, policy, strategy, service, project or any series of related events.
Inputs	The financial, human, material, technological and information resources used to implement and deliver the initiative.
Logic Model	A diagram that illustrates how an initiative is intended to work, by systematically setting out inputs, activities and outputs, and linking these with impacts (including outcomes and benefits). May also be described as a program logic.
Monitoring	A continuous and systematic process of collecting and analysing information about the implementation and impacts of an initiative.
Monitoring and Evaluation Framework	A framework that plans for, and guides, monitoring and evaluation over the life of an initiative.
Net Present Value (NPV)	The difference between the present value of benefits and present value of costs.

Definitions

Term	Definition
Net social benefit	In economic evaluation, net social benefit is the change in welfare that is derived from an initiative. This can be measured as the difference between total benefits and total costs attributable to the initiative (includes social, economic, environmental and cultural impacts). In ex-post CBA, relevant decision criteria used to assess net social benefits are the net present value (NPV) and the benefit-cost ratio (BCR). Where key benefits cannot be quantified, qualitative assessment of likely net social benefits should be reported.
Non-Experimental Design	An evaluation design that does not involve the allocation of an intervention and control or comparison group. These designs are also referred to as descriptive or observational studies.
Outcome	Changes that are attributable to the initiative outputs. Changes may be in economic, social, environmental or cultural conditions and occur in the short, medium or long term. They may include changes in lives, status, health, surroundings, knowledge, attitudes, values, behaviours or satisfaction levels.
Outcome evaluation	An evaluation that examines if, to what extent and how an initiative is leading to intended outcomes.
Output	The products, services and infrastructure that result from the initiative activities.
Peer review	A process where the quality of the work is evaluated by a reviewer, external to the business area and delivery team, with subject matter expertise.
Process evaluation	An evaluation that examines initiative implementation and delivery. It focuses on the inputs, activities and outputs of an initiative.
Program Performance Measure (PPM)	Under Outcome Budgeting, a PPM is a quantitative or qualitative measure of Program performance that is used to demonstrate change and that details the extent to which Program results are being or have been achieved.
Quasi-Experimental Design	An evaluation design that does not use randomisation to assign participants to an intervention group or to a comparison group (a group not exposed to the initiative), to test for causal attribution. It is also known as non-randomised experimental design.
Stakeholders	Members of the NSW community involved with or impacted by an initiative. This includes customers or clients who are intended to benefit from, and anyone who otherwise experiences the impacts of, the initiative (who lives, works, visits or invests in New South Wales, including businesses).
State Outcome	The primary purpose for which Budget funding is being expended, which clearly explains to the public the goal that a subnational government is seeking to achieve for its people. NSW State Outcomes are accompanied by Outcome Indicators. See TPP18-09 Outcome Budgeting.
State Outcome indicator	Under Outcome Budgeting, a State Outcome indicator is a measure of effectiveness that can reasonably demonstrate the performance of the government in achieving a State Outcome.
Value for money	Value for money is achieved when the maximum benefit is obtained from the available resources. Value for money is supported by: <ul style="list-style-type: none"> • maximising output from the use of available inputs (or minimising inputs for a pre-determined level of output), • effectiveness in achieving the intended outcomes, and • maximising benefits and achieving a net social benefit (i.e., benefits exceed costs).
Welfare	Welfare is synonymous with wellbeing of an individual, group, community or the entire society (in this case, the community of NSW). Changes in welfare may be related to social, economic, environmental and cultural outcomes. In a CBA, social welfare refers to the aggregate change in benefits and costs across all NSW residents.

Appendix A. Evaluation TPG: What's new?

This TPG presents NSW Government evaluation requirements. It includes mandatory requirements, recommendations and guidance in the new NSW Treasury Policy and Guidelines single instrument format. It applies to all:

- NSW Government agencies, being all General Government Sector agencies (as listed in NSW Budget Paper No. 1), and other government entities.
- individual capital, recurrent, regulatory or policy actions, undertaken or resourced by the NSW Government, with the aim to achieve an outcome for the NSW community.

This TPG combines, updates and supersedes Treasury Circular Program Evaluation TC18-03 and the NSW Program Evaluation Guidelines 2016. It retains the general principles and technical approaches to evaluation as established in the Program Evaluation Guidelines 2016. It has been updated to reflect developments in NSW Government policies and in evaluation best practice. This includes aligning and integrating guidance with best-practice established in updated Treasury investment policies and guidelines, including:

- [Government Sector Finance Act 2018 No 55](#)
- [NSW Government Guide to Better Regulation \(TPP19-01\)](#)
- [Outcome Budgeting \(TPP18-09\)](#)
- [NSW Government Business Case Guidelines \(TPP18-06\)](#)
- [Benefits Realisation Management Framework](#)
- [NSW Government Guide to Cost-Benefit Analysis \(TPG23-08\)](#)
- [NSW Gateway Policy \(TPG22-12\)](#).

Key developments relate to:

Outcome Budgeting

The NSW Government commitment to Outcome Budgeting establishes a reform process to focus investment decision-making to achieve State Outcomes and manage resources to promote increased accountability, transparency and better value for taxpayer dollars ([Outcome Budgeting TPP18-09](#)).

Consistent with the move to Outcome Budgeting, this TPG recommends that evaluation of initiatives be undertaken:

- with attention to the role of the initiative in achieving any of the Outcome Indicator and Program Performance Measure targets set to achieve relevant State Outcomes
- where performance monitoring, including through the Outcome Budgeting processes, indicates issues with achieving targets, uncertainty regarding the alignment and contribution of an initiative to achieving relevant State Outcomes, or opportunity to examine best-practice and lessons learned.

Monitoring and evaluation frameworks (for Business Cases)

Under TPP18-06, a business case should include planning for Benefits Realisation Management (BRM) and for evaluation as key components of implementation planning.

This TPG introduces guidance on high-level monitoring and evaluation planning at initiative design and implementation. This is relevant to business case development under TPP18-06. This guidance is also relevant to establishing monitoring and evaluation planning across the suite of ongoing initiatives within a cluster.

Benefits Realisation Management (BRM) framework

The BRM framework presents a process for identifying, organising, managing and measuring benefits over the life of an initiative.

This TPG introduces guidance on coordinating BRM with initiative monitoring and evaluation processes. It identifies the role of BRM in establishing monitoring of an initiative's impacts (outcomes and benefits) and supporting the reporting of the net social benefits of an initiative. It provides guidance to support evaluation planning under BRM.

Cost-Benefit Analysis (ex-post)

The NSW Government Guide to Cost-Benefit Analysis (TPG23-08) presents CBA as a method for systematically organising and presenting information on all the impacts of an initiative, including social, economic and environmental impacts (the term 'cost-benefit analysis' replaces what had previously been referred to as 'economic appraisal', to clarify the scope of the analysis). TPG23-08 states that evaluation of large, complex or risky initiatives should be as rigorous as possible, applying full ex-post cost-benefit analysis and empirical measurement of outcomes wherever feasible.

This TPG introduces guidance on undertaking economic evaluation aligned to the implementation phase of the initiative and in coordination with process and outcomes evaluations. It also provides enhanced guidance on undertaking CBA as an ex-post evaluation (during or at the end of the implementation of an initiative). This TPG identifies the role for ex-post CBA in informing business case variation and lessons learned reporting.

NSW Gateway Policy

Assurance processes, under the NSW Gateway Policy (TPG 22-12), provide a risk-based framework for reviews by independent experts at up to seven defined critical decision points (or gates) in the life of an initiative (capital, ICT and major recurrent).

This TPG identifies where Gateway reviews may examine BRM and evaluation processes over the life of an initiative. It also supports the use of Gateway tiering systems to inform assessment of initiative risk when prioritising initiatives for evaluation and designing evaluation.

Other additional guidance

Distribution and equity: this TPG identifies questions that can be asked regarding distribution and equity in different types of evaluation.

Stakeholder relationships: this TPG includes enhanced guidance on stakeholder engagement and collaboration in evaluation, consistent with developments in evaluation practices that increase the participatory and co-design role of the community.

The enhanced guidance on stakeholder collaboration is consistent with NSW Government's commitment to a different way of working with, and in support of, First Nations communities, by building strong working partnerships that have at their heart respect for local First Nations culture, leadership and decision making (NSW Government OCHRE Plan). Treasury is developing further guidance for incorporating First Nations cultural values into initiative planning, design, appraisal and evaluation. This will support the objectives of improving outcomes for First Nations communities and improving policy design and public investment decisions.

Appendix B. Build evaluation culture and practice

Key points

The practice of evaluation will be supported within an agency by:

- building a culture of evaluation
- integrating monitoring and evaluation into agency practices.

Monitoring and evaluation are most effective when adequately resourced, supported by leaders, and part of a culture of enquiry and learning within the organisation.

To build a **culture of evaluation**, an agency should aim to:

- be results-oriented
- provide a safe environment of learning and growth, with monitoring and evaluation positively recognised as an opportunity to learn, continuously improve and adapt (including a culture of learning from failure and transparency with evaluation findings)
- ensure understanding and appreciation of roles and responsibilities related to monitoring and evaluation and how this will support decision-making
- continuously work with the community to ensure that initiatives are appropriate and relevant to their needs.

To build **evaluation practices**, an agency should aim to:

- ensure collaboration between initiative designers and evaluators early in its life
- integrate monitoring and evaluation into business-as-usual practices
- develop data collection systems and facilitate access to data sets
- ensure evaluation occurs, and findings are communicated, before key decisions are made
- use findings from monitoring and evaluation to improve design and delivery of initiatives
- share learnings across the agency and with other organisations
- seek ongoing improvement in the quality of monitoring and evaluation.

Table 7 presents a range of activities that can be undertaken across a cluster to promote monitoring and evaluation and to embed these into business practices.

Table 7: Build evaluation culture and practice - example activities

Cohort	Build a culture of evaluation	Embed monitoring and evaluation
Executive Leadership	<ul style="list-style-type: none"> Promote the role of evaluation in evidence-based decision making. Identify an evaluation lead in the executive team. Formalise and promote the use of monitoring and evaluation in key business processes. Support open and accountable monitoring and evaluation. Be supportive of innovation, allow for failure, promote learning and growth. Support staff (and stakeholder) training in monitoring and evaluation. 	<ul style="list-style-type: none"> Identify where evaluation can inform decision-making. Develop a forward evaluation framework, across the activities of a cluster, aligned to agency and cluster objectives (including State Outcomes). Build monitoring and evaluation into current initiatives. Formalise monitoring and evaluation planning as part of initiative design and planning. Plan resourcing for monitoring and evaluation. Communicate evaluation findings to executive and ministers. Develop formal responses to evaluation findings. Share evaluation findings within the cluster. Track, respond and act upon evaluation findings. Publicly release evaluation findings.
Evaluation experts (including cluster evaluation units)	<ul style="list-style-type: none"> Communicate resourcing and evaluation reporting requirements to leadership. Promote the importance of evidence, and building an evidence base, to support decision-making. Maintain communications regarding evaluation with NSW Treasury. Share learnings for evaluation conduct (e.g., design, procurement, community collaboration) and findings across the cluster. Establish mechanisms such as communities of practice to share learnings and build capabilities. Provide or facilitate evaluation training and skills development. Provide guidance on ethical conduct. 	<ul style="list-style-type: none"> Assist executive leadership with developing a strategic evaluation framework and plan across the cluster. Collaborate with initiative managers and teams to establish monitoring and evaluation frameworks for new and ongoing initiatives. Use collaboration to build evaluation capabilities across the cluster. Provide guidance on procurement, consistent with cluster procurement processes. Assist with in-house evaluations, including monitoring and evaluation design, and providing independent perspectives in evaluation. Enable flexible evaluation teams – include service delivery teams both to enhance evaluation design and learnings, and to build evaluative capabilities in teams. Provide peer review and quality assurance.
Initiative manager and Service delivery team	<ul style="list-style-type: none"> Ensure sufficient time and resources are allowed to conduct evaluation. Understand when to use evaluation tools – consult Evaluation TPG and agency guidelines and seek advice from evaluators. Provide opportunities for teams to share informal evaluative thinking – what is going well, what isn't going well, what can be improved. Engage in evaluation training and skills development. Build ongoing relationships with community stakeholders. Engage early with stakeholders to better understand initiative impacts and best evaluation approaches. Establish mechanisms for ongoing performance feedback. Be open to new learnings. 	<ul style="list-style-type: none"> Build monitoring and evaluation frameworks for new and ongoing initiatives – bring evaluation expertise into the development of business cases and monitoring and evaluation planning at initiative design and implementation. Establish logic models for new and ongoing initiatives. Identify key evaluation questions for new and ongoing initiatives. Identify data and information available for initiative monitoring (and any gaps). Establish additional ongoing data collection, as needed, to measure outcomes and benefits realisation. Align evaluation projects with state outcomes and other performance reporting. Present and communicate evaluation findings 'upwards' and across the agency.

Appendix C. Alignment with investment processes

Monitoring and evaluation are relevant to a number of investment guidelines and processes (see *Table 8* and *Figure 7*).

Table 8: Alignment with NSW Guidelines and Policies

Guidelines	Process	Monitoring and evaluation
Government Sector Finance Act 2018 No 55	In 2013, Treasury committed to a reform program known as Financial Management Transformation (FMT), which aimed to improve financial governance across the NSW public sector. The Government Sector Finance Act 2018 introduces a single framework for public sector financial management.	The Evaluation TPG sets out standards for evaluation evidence and transparency that are aligned with the Act objects to promote and support sound financial management, efficient and effective government, stewardship of government resources and related money, and the keeping and sharing of performance information in the government sector for the purposes of decisions about resource allocation.
Outcome Budgeting (TPP18-09)	<p>Outcome budgeting introduces State Outcomes as clear statements of what the NSW Government is seeking to achieve for the people of New South Wales.</p> <p>Each State Outcome is accompanied by Outcome Indicators. Every <i>program</i> (collection of activities, tasks, divisions or functions of an agency that contribute towards achieving a State Outcome) is accompanied by Program Performance Measures (PPM).</p> <p>Outcome Assessment is a comprehensive and in-depth process to assess performance of State Outcomes and <i>programs</i>.</p>	<p>Evaluation can be conducted at a whole of <i>program</i> level (including all sub-initiatives within the <i>program</i>), for individual initiatives, or groups of initiatives. Initiative evaluations can provide evidence that can contribute to <i>program</i> level review of performance against State Outcomes and of constituent <i>programs</i>.</p> <p>The impact of an initiative on State Outcomes will vary according to the size and purpose of an initiative. Large initiatives may have a direct impact on a State Outcome. Smaller initiatives may contribute towards or combine with other initiatives to produce State Outcomes. Performance information for individual initiatives will likely be more comprehensive and focused on the detail of initiative activities and their specific outcomes but will be aligned to the Outcome Indicators and Program Performance Measures for a State Outcome and <i>program</i> respectively.</p> <p>This Evaluation TPG recommends that evaluation be considered where performance monitoring identifies that Outcome Indicators and Program Performance Measures are not being met, where an initiative's current contribution to achieving State Outcomes is not clear, or where there is opportunity to examine best-practice or lessons learned.</p>
NSW Government Business Case Guidelines (TPP18-06)	<p>All General Government agencies and Government businesses are required to provide Treasury with business cases for critical investment, policy or regulatory decisions, based on the size and risk profile of an initiative.</p> <p>A detailed business case will include a case for change and a logic model, identify key benefits, and include CBA. Economic appraisal should follow TPG23-08 (as below).</p>	<p>A detailed business case will include implementation planning that establishes monitoring, and schedules evaluation. The logic model should set out the outputs, outcomes and benefits, and the causal links and assumptions, that will be tested in evaluation. Forward planning for evaluation will consider what should be evaluated, how and when. Resourcing for monitoring and evaluation can be incorporated into the funding request, or otherwise quarantined.</p> <p>In turn, post-implementation evaluation provides an evidence base for future business cases and CBAs. Funding submissions will be strengthened by demonstrated evidence of effectiveness, realisation of benefits, and improvements to processes.</p>

Appendix C. Alignment with investment processes

Guidelines	Process	Monitoring and evaluation
<p><u>NSW Government Guide to Better Regulation (TPP19-01)</u></p>	<p>The Guide to Better regulation assists agencies in developing regulation which is required, reasonable and responsive to the needs of New South Wales.</p> <p>All new and amending regulatory proposals need to demonstrate application of the seven Better Regulation principles, to support fully informed decision making.</p>	<p>The Better Regulation principles incorporate monitoring and review.</p> <p>Planning for monitoring of regulation compliance and impacts must be undertaken (Principle 3).</p> <p>Regulations should be periodically reviewed to ensure their continued efficiency and effectiveness (Principle 7). Reviews should generally be conducted after five years.</p>
<p><u>NSW Government Guide to Cost-Benefit Analysis (TPG23-08)</u></p>	<p>CBA is an appraisal and evaluation technique that sets out the economic, social, cultural and environmental impacts (costs and benefits) of an initiative.</p> <p>CBA is required as part of a business case for any new or significantly amending initiative, to be completed for capital expenditure of \$10 million or more, recurrent expenditure on a case-by-case basis, and for initiatives considered high-risk.</p> <p>CBA is recommended for post-implementation evaluation.</p>	<p>When conducted as an ex-ante appraisal (when scoping a potential initiative), a CBA is used to inform decision-making. An ex-ante CBA will identify, and quantify as feasible, the anticipated benefits (improvements to welfare) of the initiative. These identified benefits will be the foundation for benefits monitoring, and for review in evaluation.</p> <p>When completed after an initiative has been implemented (ex-post), a CBA can be used to assess whether benefit estimates are realised and compare the actual costs and benefits of an initiative with the ex-ante estimates.</p> <p>In turn, ex-post CBA can yield a rigorous evidence base as a reference for future business cases and CBAs, both for ongoing implementation of the initiative and other initiative proposals.</p>
<p><u>Benefits Realisation Management (BRM) framework</u></p>	<p>BRM is a comprehensive process of identifying, planning, managing and evaluating the intended benefits of an initiative throughout its life. Benefits are reviewed and updated over time, when benefits are realised, or where new benefits or dis-benefit emerge.</p>	<p>The BRM framework provides guidance on monitoring outcomes and benefits to ensure benefits from investments are realised. Effective benefits monitoring will track performance and inform evaluation of an initiative.</p> <p>The final phase of the Four Phases of BRM (Evaluate) investigates the realisation of outcomes and benefits. Deliverables include an updated benefits register, an updated benefits report, a business case variation report and a lessons learned report.</p>
<p><u>NSW Gateway Policy (TPG22-12)</u></p>	<p>Gateway is a project assurance process that mandates reviews by independent experts at up to seven defined critical decision points (or gates) in the life of an initiative (capital, ICT and recurrent).</p> <p>The Gateway Policy uses a risk-based approach to determine when, and at what decision points, Gateway reviews should be undertaken. Gate 0 to Gate 5 are applicable during initiative development and delivery. Gate 6 (post implementation review) is held when the outputs and outcomes can be assessed.</p> <p>The Gateway Review process aligns with current NSW Government policy and strategies. Reviews are not a replacement for a delivery agency's internal governance. It is intended that delivery agencies use existing project documentation and not create or customise documents for the review.</p>	<p>Planning for BRM and evaluation should commence as early as possible (from Gate 0).</p> <p>Gate 2 includes review of the Detailed Business Case. Key documents required include the benefits register, the logic model and evaluation plan.</p> <p>Initiatives with a Tier 1 high profile/high risk rating will trigger a Gate 6 post-implementation independent assurance review (the review process is optional for Tiers 2 and 3). This does not imply that evaluation is only required for initiatives identified for Gateway review.</p> <p>The Gate 6 review, depending on its timing, may review benefits realisation and/or evaluative practice.</p> <p>Where an initiative will be subject to a Gate 6 review, the agency should ensure that BRM and evaluation planning also meet any unique requirements for the review.</p> <p>This Evaluation TPG identifies that Gateway assessments can be used to assess the risk level of an initiative when identifying priorities for evaluation.</p>

Figure 7: Alignment with NSW guidelines and policies over the investment lifecycle

	Problem definition	Strategic business case	Detailed business case	Implement and monitor	Evaluate
Activities	Establish the case for change and explore options.	Consider the feasibility and value for money of a range of options.	Identify the preferred option and set up commercial/ management arrangements for procurement and delivery.	Procure, deliver, manage and monitor the project.	Evaluate the investment to inform decision making.
Monitoring and Evaluation activities	<ul style="list-style-type: none"> Reference relevant evaluations of previous/similar initiatives. Use evaluation evidence to inform solutions. 	<ul style="list-style-type: none"> Refine case for change, objectives, high level logic model and theory of change. 	<ul style="list-style-type: none"> Develop detailed logic models for preferred options Develop benefits realisation management plan Develop monitoring and evaluation framework 	<ul style="list-style-type: none"> Refine detailed logic model for implemented initiative Monitoring and data collection (include benefits register) Process evaluation (as relevant) 	<ul style="list-style-type: none"> Refine detailed logic model for implemented initiative Process, outcome and economic evaluation including CBA Reporting findings to inform decision making
Alignment with other NSW policies/guidelines	Outcome Budgeting <ul style="list-style-type: none"> Business Case Guidelines: Stage 0 Guide to Better Regulation: Principles 1-2 	<ul style="list-style-type: none"> Business Case Guidelines: Stage 1 Guide to Better Regulation: Principles 3-6 	<ul style="list-style-type: none"> Business Case Guidelines: Stage 2 Guide to Better Regulation: Principles 3-6 	<ul style="list-style-type: none"> Outcome Budgeting Guide to Better Regulation: Principle 7 	<ul style="list-style-type: none"> Outcome Budgeting Guide to Better Regulation: Principle 7
	Guide to Better Regulation: Principles 1-2	<ul style="list-style-type: none"> Guide to CBA: Steps 3-7 Benefits Realisation Management Framework: Phase 1 and 2 NSW Gateway Policy: Gate 1 	<ul style="list-style-type: none"> Guide to CBA: Step 8 Benefits Realisation Management Framework: Phase 2 NSW Gateway Policy: Gate 2 	<ul style="list-style-type: none"> Benefits Realisation Management Framework: Phase 3 NSW Gateway Policy: Gates 3-5 NSW Gateway Policy: Gate 0 	<ul style="list-style-type: none"> Guide to CBA: Step 9 Benefits Realisation Management Framework: Phase 4 NSW Gateway Policy: Gate 6

Appendix D. Types of evaluation and example questions

Process evaluation

What is process evaluation?

Process evaluation examines initiative implementation and delivery. It focuses on the inputs, activities and outputs of an initiative. It can be designed to examine the context in which the initiative is operating, identify issues with delivery, review the efficiency of the initiative, and provide information for process improvements.

What does process evaluation involve?

A process evaluation will examine whether an initiative is being implemented, and intended outputs are being delivered, as intended. It is an opportunity to check implementation against the design and logic model of the initiative, and to check outputs against targets and wider standards. Where implementation issues are identified, further analysis can explore the initiative's current operating conditions and identify any barriers. Where the initiative has not been implemented as designed, the evaluation can be used to improve understanding of how the initiative is working in practice.

A process evaluation will also consider whether the initiative is achieving its intended reach and responding to customer and other stakeholder needs. This can include examining the pattern of uptake or engagement with services, including by different sub-groups or in different contexts, and considering the appropriateness of delivery methods and services.

A process evaluation may also usefully identify the costs of implementation, and how, where and why these have varied from initial projections. Identifying both monetary and non-monetary implementation costs will provide important foundations for economic evaluation.

How can process evaluation be designed?

The design of a process evaluation will depend upon the characteristics of the initiative, the questions that are relevant to ask regarding its implementation (see example questions in *Table 9*), and the extent to which monitoring information has been collected.

Where monitoring has been established, data from a suite of measures of inputs, activities and outputs can provide information for evaluation. Additional data will be collected, such as customer experience and satisfaction information, as defined by the evaluation questions and design (see also *Resources. Technical notes: Evidence in evaluation (data types and collection)*).

The process evaluation may incorporate analysis of cost-efficiency, to examine the efficiency with which inputs are used to produce outputs (see *Economic evaluation and cost-benefit analysis (CBA)*).

Process evaluation can also consider distribution and equity of outputs, as relevant to the initiative and the purpose of the evaluation.

When can process evaluation be undertaken?

Process evaluation can be undertaken at any stage in the implementation of an initiative, where there are questions regarding implementation of the initiative.

When commenced early (e.g. within 6 to 12 months of an initiative commencing), process evaluation can check if the initiative is being implemented as intended and inform changes to design or improvements to delivery.

Process evaluation can support outcome evaluation. When an initiative hasn't met its intended outcomes, process evaluation can help distinguish implementation issues from questions on the effectiveness of initiative design.

Table 9: Process evaluation - example key evaluation questions

Implementation

- Has the initiative been implemented as designed? Where and why have adjustments been made?
- Has the initiative been implemented within intended timeframes and costs? What adjustments have been made to the schedule or budget?
- Where the initiative is delivered under partnership agreements, have these delivery relationships operated as intended?
- Are there challenges or barriers impacting implementation? How can these be mitigated?
- Have known risks been managed? Have additional risks been identified?
- What is the degree of customer or stakeholder satisfaction with the initiative delivery?

Distribution

- Is the initiative reaching the target populations? To what extent?
- Are service delivery systems reaching marginalised groups? To what extent?
- What is the rate of uptake across target groups?
- Does the reach and uptake differ for different sub-groups or contexts?
- To what extent have the needs of recipients and other key stakeholders been met?

Analysis

- What is known regarding the quantity and quality of outputs?
- How does implementation compare to established standards?
- How do the delivered outputs compare to quantity and quality targets and standards?
- Are there opportunities to improve efficiency in delivery?
- Were contextual factors (political, social, economic, cultural) adequately considered in the design and implementation of the intervention?

Where implementation issues were identified, additional questions may be:

- Does the governance structure effectively support implementation?
- Is delivery aligned with the objectives?
- What were the key factors contributing to timing or cost variance? What is recommended for future timing or costing assumptions?

When undertaken with outcome evaluation, additional questions may be:

- In what ways has the design supported (or hindered) achieving outcomes?
- Is the design appropriate to achieve outcomes?

See also: *Table 11 (analysis of cost-efficiency)*

Outcome evaluation

What is outcome evaluation?

Outcome evaluation examines if and how an initiative is leading to intended outcomes. It seeks to quantify the extent of any change and the degree to which the initiative has contributed to them. It can include assessing if the initiative is the best option to achieve desired outcomes.

What does outcome evaluation involve?

An outcome evaluation gathers evidence on the extent of change resulting from an initiative (outcomes) and examines the initiative's role in producing this change (causal attribution). Outcomes may be realised in the short, medium or longer term. They may relate to status (such as levels of education, health or employment) or conditions (such as environmental). They may include changes in people's knowledge, attitudes, values, behaviours, or satisfaction levels. The outcomes examined may include benefits (increases in social welfare associated with outcomes).

Further, a comprehensive outcome evaluation examines where outcomes have been achieved and by whom. It may investigate when and where the initiative is more effective, how outcomes may be different for different groups, the likelihood of outcomes continuing, and what contexts and factors influence the achievement of outcomes.

The outcome evaluation is also an opportunity to identify any additional positive or negative outcomes and benefits from the initiative that were not foreseen in initiative design or ex-ante appraisal.

Where the initiative directly contributes to achieving a State Outcome, the outcome evaluation can be designed to assess the contribution of the initiative to achieving the State Outcome. Where the initiative is one of a number of interdependent activities that contribute to achieving a State Outcome, the evaluation may focus on the outcomes that are directly attributable to the initiative.

How can outcome evaluation be designed?

There is a range of outcome evaluation research designs that can be used, from experimental to quasi-experimental to non-experimental designs (see example questions in *Table 10 and Resources. Technical Notes: Outcome evaluation designs* for further information).

An outcome evaluation needs to be based on strong evidence, including baseline measurements. A robust evaluation will be able to rule out alternative explanations for any changes in outcomes. The strongest methods for examining causal attribution will use a control group (as in randomised experimental design) or a comparison group (as in quasi-experimental design). Experimental and quasi-experimental evaluations need forward planning, and usually ethics approval.

An outcome evaluation that demonstrates the initiative caused the observed outcomes is an important foundation for economic evaluation. It may be useful to consider if the initiative is the most cost-effective method to deliver on outcomes, particularly if this question was not asked in the ex-ante appraisal. Where the benefits associated with outcomes can be valued, CBA can be used to assess net social benefits (see *Economic evaluation and cost-benefit analysis (CBA)*).

Outcome evaluation can be designed to consider distribution and equity of outcomes, and specifically ask to what extent the needs of specified groups have been met.

When can outcome evaluation be undertaken?

Outcome evaluations can be undertaken when an initiative is expected to achieve outcomes. The timing of the evaluation will need to balance the requirements for performance information with the timeframes over which outcomes are expected to be achieved. Depending upon the characteristics of the initiative, outcome evaluation may be undertaken from one to four years after the initiative is established and stable (with sufficient sample size). Where intended outcomes are expected to take a long time to be realised, evaluate short to medium term outcomes as an indicator of the likelihood

achieving the expected final outcomes. For ongoing initiatives, outcome evaluation can be scheduled at a nominated timeframe, such as every five years.

Table 10: Outcome evaluation - example key evaluation questions

Outcomes realisation

- What are the actual changes (outcomes) delivered by the initiative?
- Are intended short, intermediate or longer-term outcomes delivered by the initiative?
- When are outcomes being realised?
- What is the key effect size (scale of change) and range? How has this been identified?
- What is the counterfactual (what change in key indicators would have happened in the absence of the initiative)? How has the counterfactual been estimated (is there a control or comparison group)?
- Are there any unintended outcomes for participants or stakeholders?

Distribution

- Are outcomes being achieved for targeted groups?
- What is the distribution of outcomes among different groups?
- Does the initiative produce different outcomes in different contexts or stakeholder groups? What situations separate outcomes for groups?
- Who does the initiative work well for? In what circumstances?
- Who has been subject to unintended outcomes?

Analysis

- To what extent can changes be attributed to the initiative (was the initiative the cause of observed outcomes)? What other factors (may have) contributed to the outcomes?
- Under what conditions has the initiative been most effective? What factors support success?
- How does the initiative compare with other activities that produce the same or similar outcomes?
- How has the initiative contributed to the achievement of state outcomes? Can a share of observed changes in the relevant state outcomes be reasonably attributed to the initiative?
- Does the initiative contribute to change as expected?
- Where the initiative has equity objectives:
 - does the intervention achieve results for stated equity-focused objectives?
 - to what extent does the initiative change identified inequities?
 - what is the effect on different populations and on the structural drivers of inequity?
 - has there been a change in the factors leading to inequity?
 - are inequities between groups likely to increase, remain stable or decrease after completion of the initiative?

See also: Table 11 (analysis of cost-effectiveness)

Economic evaluation and cost-benefit analysis (CBA)

What is economic evaluation?

Economic evaluation identifies and measures the consequences of an initiative in relation to its costs, to provide an assessment of value for money and support comparison to an alternative course of action. Economic evaluation can be undertaken for an initiative's outputs (analysis of cost-efficiency), outcomes (analysis of cost-effectiveness) or benefits (cost-benefit analysis), as relevant to the stage of initiative implementation. CBA is Treasury's preferred method for economic evaluation as it presents the full range and depth of the impacts of an initiative, and uniquely allows for net social benefits of initiatives to be assessed and compared across programs and portfolios. Economic evaluation should be undertaken with reference to the [NSW Government Guide to Cost-Benefit Analysis \(TPG23-08\)](#). See *Table 11* for examples of evaluation questions for cost-efficiency and cost-effectiveness analysis (CEA) (building upon process evaluation and outcome evaluation respectively) and *Table 12* for examples of evaluation questions relevant to ex-post CBA.

Cost-efficiency analysis – analyse outputs delivered against costs

To complement process evaluation, cost-efficiency analysis can be undertaken to examine the level of outputs (goods and services) delivered relative to costs. The evaluation may examine costs per unit of output, compared to established standards or other activities that produce similar (quality of) outputs (this approach only measures the efficiency of output delivery, not the effectiveness of outputs in achieving outcomes, or the relationship between the quality of outputs and the benefits realised). The evaluation may identify opportunities to reduce delivery costs.

Cost-effectiveness analysis – analyse outcomes against costs

To complement outcome evaluation, cost-effectiveness analysis (CEA) can be undertaken to examine the level of outcomes delivered relative to costs, and to assess if the initiative is the most cost-effective method to deliver intended outcomes compared to other initiatives producing comparable outcomes. CEA may be relevant when an evaluation is timed to examine short- or medium-term outcomes (where end outcomes and their benefits are yet to be realised).

The method used to evaluate cost-effectiveness will depend on the type of outcomes, available data and the evidence required for decision-making. A CEA can compare costs per unit of outcome between options that deliver the same outcome. A 'cost-utility analysis' can be undertaken where a common outcome measure can be used. These approaches can be used to compare different ways to achieve a common outcome, but not to compare options with different outcomes. A 'cost-consequence analysis' (CCA) can be used to identify multiple outcomes (this will list a range of outcomes but not produce a cost-outcome ratio) and provide a foundation for CBA.

Table 11: Economic Evaluation - example key evaluation questions

Costs

- What are the initiative's actual implementation costs, including overheads?
- How do actual implementation costs compare with projected or planned costs? What are the reasons for deviations between forecast and actual costs?

Process evaluation: analysis of cost-efficiency

- To what extent does the initiative demonstrate cost efficiency? How do delivery costs compare with best-practice or other initiatives producing similar (quality of) outputs?

Outcome evaluation: analysis of cost-effectiveness

- To what extent does the initiative demonstrate cost effectiveness? How do delivery costs compare with best-practice or other initiatives producing comparable outcomes?

Cost-benefit analysis – analyse benefits against costs

What is cost-benefit analysis (CBA)?

CBA is a comprehensive examination of an initiative's impacts on welfare (including economic, social, environmental and cultural impacts) and assessment of whether the investment delivered a net social benefit to the people of New South Wales.

What does ex-post CBA evaluation involve?

Ex-post CBA is used to present all the impacts (positive and negative, intended and unintended, direct and indirect) that result from changes attributable to the initiative (including in social, economic, environmental and cultural conditions) and assess the extent to which an investment provides a net social benefit to the NSW community.

Where an ex-ante CBA supported implementation of the initiative, ex-post CBA can be used to test the accuracy of cost and benefit forecasts presented in the business case. It can inform updates to the benefits register and benefits report (as a conclusive step in Benefits Realisation Management (BRM)), and where relevant, a business case variation report and a lessons learned report.

How can ex-post CBA be designed?

CBA will identify and seek to quantify the key benefits and costs (including dis-benefits) from an initiative. Benefits are increases in welfare that result from an initiative's outcomes (economic, social, cultural or environmental).

Benefits can be understood qualitatively and quantitatively. Some benefits will be readily measured and monetised. Benefits should be described qualitatively where they are challenging to quantify or value. It is best-practice to involve stakeholders in processes of identifying and valuing benefits or dis-benefits.

The CBA will consider the marginal change in benefits and costs due to the initiative, compared to a scenario without the initiative (the counterfactual), thereby excluding any change that is due to other factors. An ex-post CBA may:

- build on evidence from an outcome evaluation, where the attribution of the initiative to change has been identified, to then identify or estimate the benefits associated with outcomes
- use outcome and benefits monitoring information, including that collected under a BRM framework, and account for changes that would likely have occurred in the absence of the initiative (this may be the approach best suited to evaluating investments, such as significant infrastructure project or a community wide intervention where establishing outcome evaluation with a control group may not be feasible, ethical or appropriate).

Estimating the net present value (NPV) and benefit-cost ratio (BCR) of an initiative, as part of ex-post CBA, provides an assessment of the net social benefit and value for money (comparison of benefits against costs) of an initiative, and supports comparison of the value of the investment against alternative investments. Where it is not feasible to quantify or monetise the relevant key benefits, the likely net benefit to the NSW community can be discussed qualitatively, and CEA used as a secondary option to assess value for money.

Where equity considerations are relevant, the CBA should be designed to consider the distributional impact of an initiative.

When can ex-post CBA be undertaken?

CBA evaluation can be undertaken when an initiative can reasonably be expected to be producing benefits. Benefits may be realised over an extended period, with different benefits realised in different time periods. Evaluation can be undertaken at key stages of benefits realisation (including at initiative completion and long after the initiative has been implemented) to investigate whether benefits are sustained.

Table 12: Ex-post CBA - example key evaluation questions

Costs

- What are the actual implementation costs, including overheads? What are relevant unit costs?
- What are the indirect costs (dis-benefits/reductions in social welfare) associated with the initiative? Are there any material costs that cannot be quantified?
- What are the total costs of the project (including any dis-benefits)?
- How do actual costs compare to forecast (*ex-ante*) costs?

Benefits

- What are the range of benefits attributable to the initiative (including future benefits)?
- Are there any unintended benefits?
- How have benefits been quantified and monetised? Where benefits are non-market, how can these be valued? Are there any material benefits that cannot be quantified?

Distribution

- What is the distribution of benefits (and costs) among different groups?
- Does the initiative produce different benefits in different contexts or for different groups?
- Who, and in what circumstances, does the initiative work well for?

Assess net benefits

- What is the NPV of the initiative to the NSW community? Where material costs or benefits cannot be valued and what is the likely net benefit for NSW (based on evidence)?
- What is the BCR of the initiative to the NSW community?
- To what extent has the initiative delivered value for money?

Ex-ante review

- How accurate were ex-ante assumptions and quantity forecasts?
- How accurate were the price parameters used for ex-ante valuation?
- Did the ex-post BCR and NPV fit within the calculated error margin of the ex-ante CBA?
- For CBA undertaken during the initiative, do the BCR and NPV forecasts need to be revised?
- How can future costing and benefit forecasts for similar projects be improved?

Benefits realisation analysis – refer also to BRM framework

- How do actual outcomes and benefits compare with intended outcomes and benefits?
- Have stakeholders realised the benefits they expected?
- Are benefits sustainable? Will benefits continue beyond the initiative or other government support? What other conditions are required to realise those benefits?
- Are any benefits no longer valid or achievable?
- What endogenous factors (e.g., assumed causal links or value parameters) may have resulted in realised benefits differing from anticipated benefits?
- What external factors may have impacted on the realisation of benefits?
- What lessons have been learned regarding benefits realisation and management?
- Are initiative improvements or corrective actions required?
- Is more in-depth review required? What aspects of the initiative require further investigation?

See also: *NSW Government Guide to CBA (TPG23-08)*. *Technical notes: Ex-post CBA* provides further notes on ex-post CBA to be read in conjunction with TPG23-08.

Appendix E. Example theory-of-change and logic model: building evidence

This appendix presents an example theory-of-change (*Box 1*) and logic model (*Figure 8*) relevant to an initiative to build a bank of evaluation evidence that will support understanding of ‘what works’ and inform investment decision-making.

Box 1: Example theory-of-change

The problem that the initiative seeks to address: The initiative is responding to an identified gap in information on the effectiveness and value for money of government investments.

The purpose of the initiative: The initiative is to build a bank of evidence of ‘what works’ in NSW, to support evidence-informed decision-making, and enhance the effectiveness and net social benefits of government spending.

How and why the activities of an initiative are expected to achieve change:

How

To support evidence development, the agency will provide guidance on requirements and standards for appraisal and evaluation, collate evidence, and assess the quality of evidence to identify examples of best-practice.

These activities will support increased quantity and quality of appraisal and evaluation, which will provide evidence to inform decision-making for individual initiatives.

The increase in evidence will provide a broader evidence base to improve understanding of outcomes and benefits from NSW Government investments and support shared learning (and reduce duplicative analysis).

This will support evidence-informed resource allocation decisions across government, improve effectiveness from government investments in achieving outcomes, and increase the transparency of the outcomes and benefits delivered.

The benefits will include improved wellbeing for the NSW community resulting from increased effectiveness and efficiency of government initiatives and increased net social benefit from government investments.

This will contribute to achieving the State Outcome ‘A sustainable fiscal environment enabling delivery of outcomes’, and support the transition towards Outcomes Budgeting.

Why

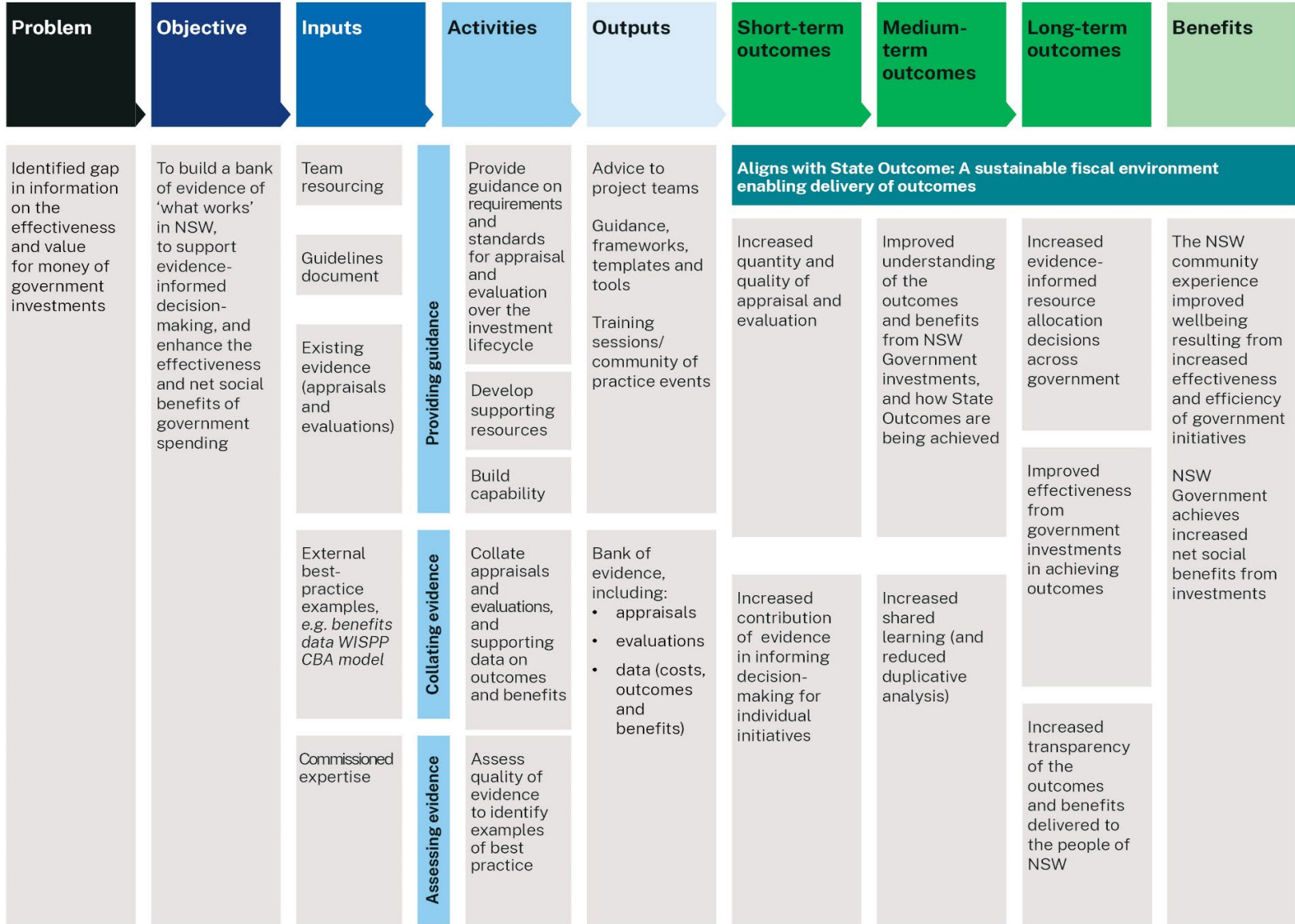
A number of NSW audit reports have documented deficiencies in the understanding of whether government investments achieve expected outcomes and value for money. Without market prices, measuring public sector productivity and deciding public sector resource allocation are often complex. In response, a need to enhance evaluation methodologies and analytical approaches to measure program outcomes, including economic evaluations, has been identified. Improved evidence will support decisions on how government funding can be allocated to generate the greatest welfare improvements for the people of NSW.

The success of Washington State in improving outcomes and reducing costs demonstrates that building evidence on initiative outcomes and benefits can enhance government allocation of resources. The Washington State Institute for Public Policy (WSIPP) produced impartial and high-quality research, and the legislature used the results of their work to change how money was spent. In 2012, the Pew Charitable Trust published a case study examining the impact of their work in juvenile justice programming on spending and public safety outcomes. The report found that through the creation of the benefit-cost model and increased use of rigorous program evaluations, the state experienced:

- greater improvements in crime and arrest rates compared to the national average
- incarceration rates lower than the national average
- an estimated \$1.3 billion dollars in savings per two-year budget cycle.

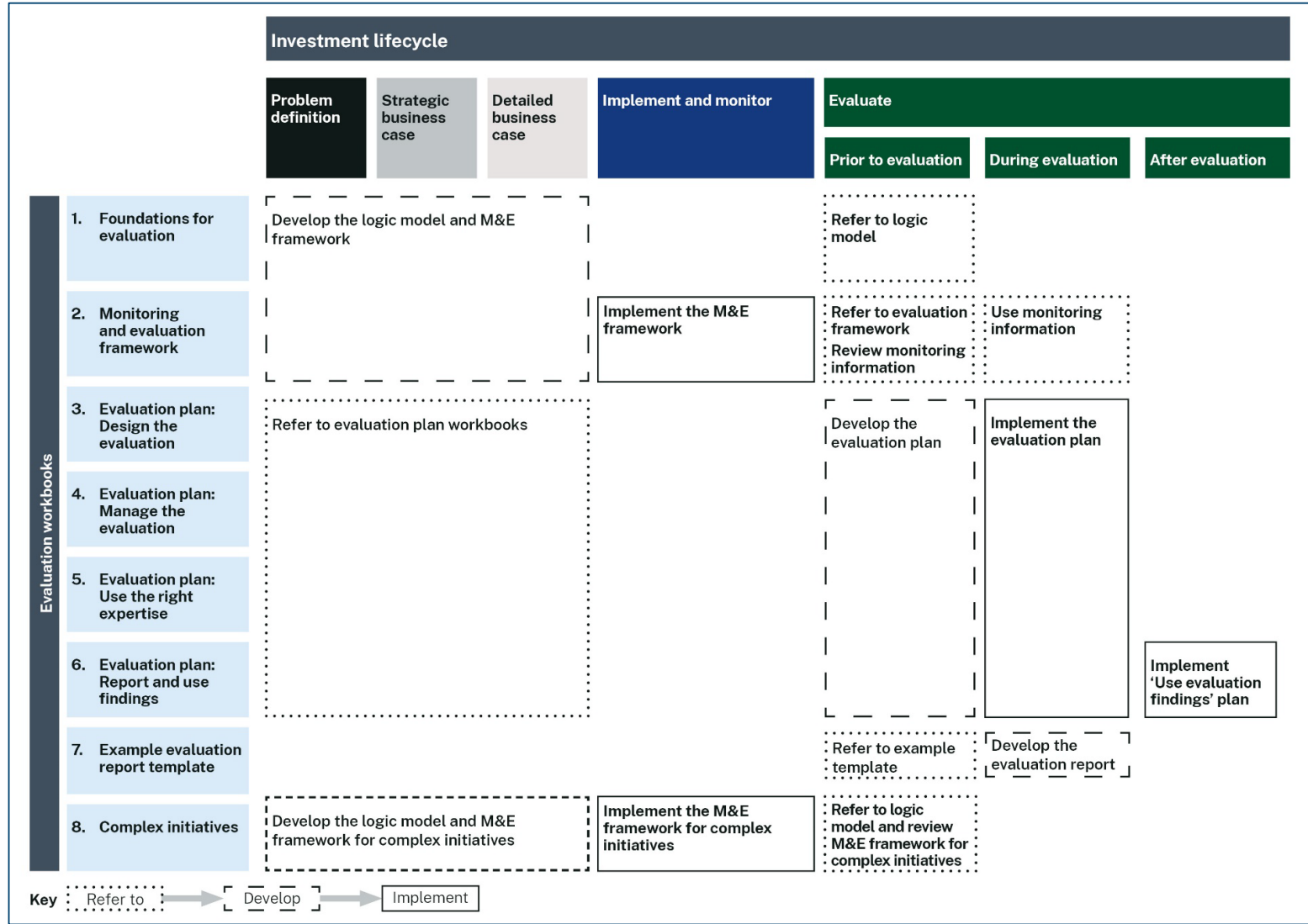
Appendix E. Example theory-of-change and logic model: building evidence

Figure 8: Example logic model



Appendix F. Evaluation workbooks summary

Figure 9: Using evaluation workbooks over the investment lifecycle



52 Martin Place
Sydney NSW 2000

GPO Box 5469
Sydney NSW 2001

W: treasury.nsw.gov.au

This publication is protected by copyright. With the exception of (a) any coat of arms, logo, trade mark or other branding; (b) any third party intellectual property; and (c) personal information such as photographs of people, this publication is licensed under the Creative Commons Attribution 3.0 Australia Licence.

The licence terms are available at the Creative Commons website at:
creativecommons.org/licenses/by/3.0/au/legalcode

NSW Treasury requires that it be attributed as creator of the licensed material in the following manner:
© State of New South Wales (NSW Treasury), (2022).