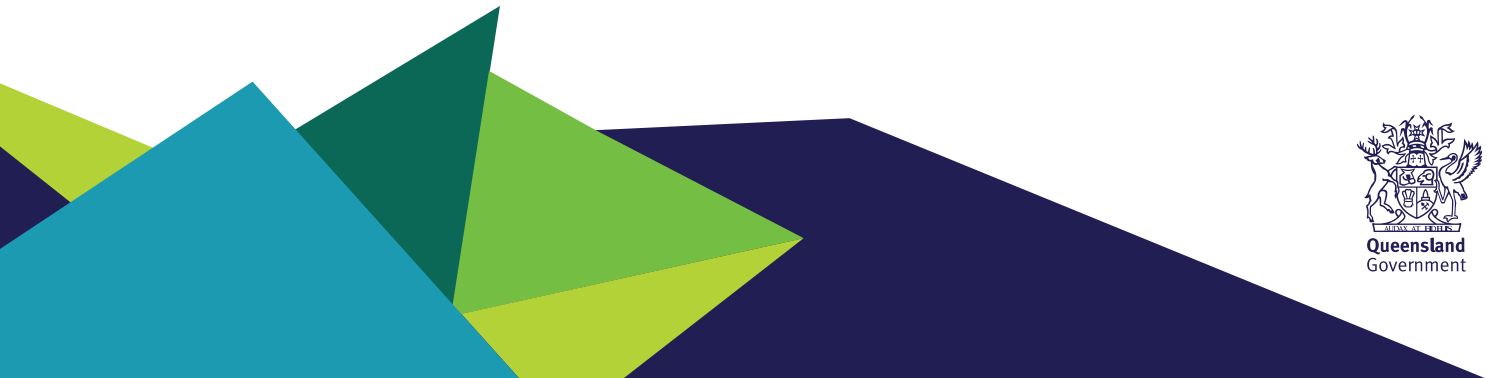
Queensland Government

Program Evaluation Guidelines

Second Edition
2020



Queensland Government

Contact for enquiries

Queensland Treasury is the custodian of this document.
All enquiries should be directed to:

Email: PEG@treasury.qld.gov.au
Phone: 07 3035 6849

Acknowledgements

The second edition of the *Queensland Government Program Evaluation Guidelines (PEG)* was developed by the Queensland Government Statistician's Office, Queensland Treasury. It contains content adapted from similar published works and reflects input from various Queensland Government agencies. All contributions are greatly appreciated.

Attribution

Content from this document should be attributed to:

Queensland Government Statistician's Office, Queensland Treasury, *Queensland Government Program Evaluation Guidelines, Second Edition, 2020*

Copyright

Licence

Contents

Introduction to the Program Evaluation Guidelines

These guidelines outline a broad set of wise-practice principles to assist agencies to plan, commission, manage and conduct consistent, transparent and high-quality evaluation.

Background and context

Queensland Treasury released the first edition of the Queensland Government Program Evaluation Guidelines (PEG) in November 2014, to clarify expectations and provide guidance for the evaluation of public sector programs.

This edition of the PEG contains new and revised content to support the application of contemporary evaluation practice within the context of government priorities.

Why program evaluation matters

At its broadest level, Government business is about responding to issues with robust public policy.

Government policy is a statement of intent, implemented through policy instruments including laws, advocacy, funding and direct action. Developing and implementing programs is one way Government delivers policy outcomes.

Program evaluation assesses a program's effectiveness, efficiency, appropriateness, utility, equity, value for money and/or sustainability. For government-funded programs, evaluation can support learning, performance improvement, decision-making and accountability.

Evaluation plays an essential role in the development, implementation and assessment of government programs by helping stakeholders to understand:

- how to target government investment carefully so that every dollar spent makes a difference
- whether a program is working as intended, for who, how, under what circumstances, why and at what cost
- whether a program is achieving its objectives and constitutes an appropriate policy response
- whether a program is generating any unintended consequences
- whether a new set of activities is required to respond to any opportunities, risks or needs identified.

In the PEG, the term 'program' means a set of discrete, time-based activities that respond to an identified need or policy position. Many evaluation principles that apply to programs can also be applied to systems, policies, strategies, interventions, initiatives, services or projects, including trial or pilot activities.

Using the PEG

The PEG should be used by senior agency officers, program managers, policy officers and evaluators to support evaluations of government-funded programs and align expenditure to government priorities. The PEG provides users with a high-level of understanding of how to evaluate programs and enhance evaluation capability.

The following Queensland Government resources and responsibilities complement the PEG:

Performance Management Framework (PMF)

The PMF is designed to guide public sector agencies in performance management, focussing on planning, measuring and monitoring performance and public reporting.

Project Assessment Framework (PAF)

The PAF is a whole-of-government project assessment process that presents a common approach for assessing projects at critical stages in their life cycle.

Financial Accountability Act 2009

The *Financial Accountability Act 2009* was enacted to govern public sector financial administration in Queensland. It requires that "accountable officers and statutory bodies achieve reasonable value for money by ensuring the operations of the department or statutory body are carried out efficiently, effectively and economically".

Queensland Audit Office (QAO)

In addition to these frameworks, the QAO's role includes enhancing accountability and improving public services by conducting performance audits on the efficiency and effectiveness of public services.

Further resources and *common evaluation terms* can be found at the back of this document.

Planning for evaluation

Evidence-informed program decision-making is strengthened by well-planned, timely evaluations. Good evaluation planning considers factors such as the policy context, program lifecycle, key questions, scope, authorising environment, resources, timeframes, stakeholder expectations and any changing conditions.

Policy context

The priorities of government drive policy decisions. Evaluation can support policy-makers in understanding whether their key objectives are being met. Evaluators of government programs should familiarise themselves with current government priorities and consider how these might influence their evaluative choices.

Program lifecycle

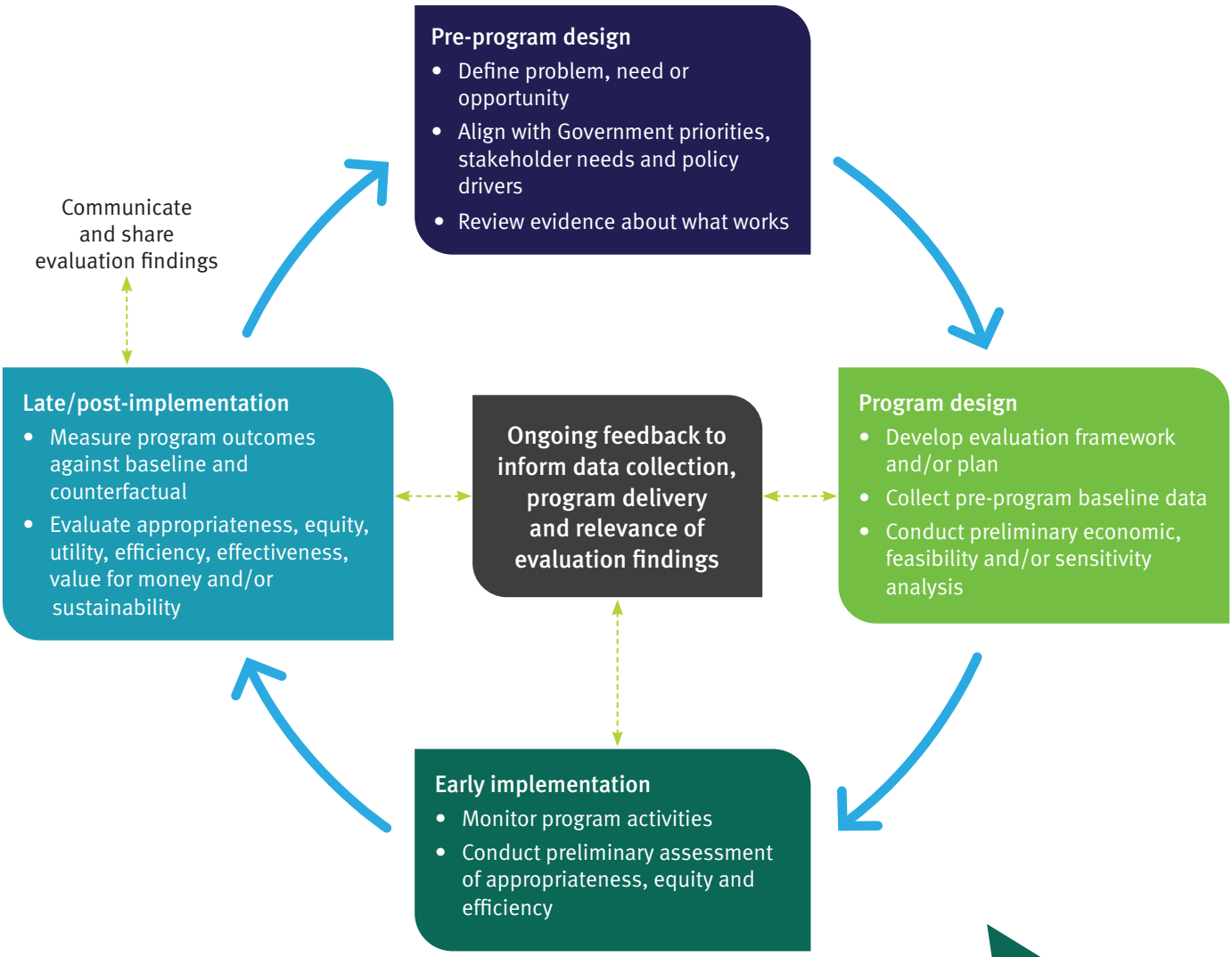
Evaluation is essential for new programs, trials and pilots and should be integrated throughout all phases of a program's lifecycle where possible (Figure 1).

However, when evaluative thinking starts in the early phases of the program's lifecycle, the evaluation will provide answers to a broader range of questions.

For example:

To answer questions about a program's effectiveness, it is important to collect primary baseline data (through methods such as surveys and interviews) prior to program implementation. Along with any available secondary data, this can establish pre-program conditions against which change can be measured.

Figure 1: Integrating evaluation throughout a program's lifecycle



Program context and purpose

Before considering evaluation designs and methods, it is important to understand what the program is trying to achieve in terms of measurable objectives and how the program is designed to contribute to the change desired.

Often, this includes identifying likely external, contextual and confounding factors as well as the assumptions that link program inputs, outputs and outcomes. Program logic and theory of change is often used to identify, understand and visually represent this.

This information forms a key component of the evaluation plan, along with the evaluation questions, scope, design, methodology, authorising environment, resources, timeframes and stakeholder expectations.

Evaluation objectives, questions and scope

The purpose or objectives of an evaluation can often be translated into evaluation questions or criteria. For example, an evaluation may seek to understand whether a program worked. This can be expressed in the form of evaluation questions such as:

- To what extent was the program effective in achieving intended outcomes?
- To what extent can outcomes be uniquely attributed to the program (as opposed to counter explanations or the counter factual)?
- Do outcomes represent value for money?
- How equitably and efficiently were benefits distributed among stakeholders?

Other common evaluation questions include:

- Was the program implemented as intended and if not, why not?
- How efficiently was the program implemented?
- How accessible and useful was the program?
- How appropriate was the program design and implementation?
- Before and/or since program implementation, did economic analysis indicate a net benefit?
- Were there any unintended consequences or outcomes (if so, how were they managed)?

For programs with multiple delivery strategies, each strategy may need to be evaluated individually as well as collectively. Examining individual elements of a multi-faceted program can help to answer:

- which program initiatives are providing the greatest impact
- which elements of program delivery are most effective in generating desired outcomes
- is greater impact achieved when specific strategies are combined into a package of initiatives
- in what contexts are mechanisms of change triggered to achieve desired outcomes?

KEY CONSIDERATION

Measuring the extent to which a program was effective usually requires an understanding of its purpose, implementation and any influencing conditions.

These questions will be particularly relevant for the evaluation of whole-of-government programs, where multiple agencies are working towards shared understanding of collective objectives.

To answer any evaluative question, an evaluation's objectives need to be clearly defined and the measures of performance should be specific, measurable, accessible, relevant, robust and timely.

Performance measures can be derived from existing secondary sources such as operational administrative databases or official statistics. However, when these sources are not available, accessible, fit-for-purpose or of sufficient data quality, then surveys or other primary data collections may be required.

Tools such as evaluability assessments, literature reviews and needs assessments can assist evaluators to determine whether the right conditions, data and people are present for an evaluation to proceed, while also providing greater clarity about the program and its target groups.

Evaluation designs and methods

A quality evaluation has a sound, fit-for-purpose design that clearly articulates the research methods needed to gather evidence to achieve its objectives.

Evaluation design should consider what information is required and available, and how it will be collected. It will also need to address any information limitations or deficiencies.

For example:

Suppose the purpose of the evaluation was to understand whether outcomes from a program implemented in one location could be generalised to other locations. The design would need to consider place-based, program participant and contextual factors.

If the evaluation purpose was to understand a program's effectiveness, the design would use data collected prior to, during and after program implementation to compare recipients of the program with a similarly-matched group not receiving the program. It would seek to understand how the program contributed to outcomes, who was affected, how and under what conditions.

Appropriate evaluation designs can control for many factors that may compromise the ability to demonstrate a relationship exists between a program's actions and observed outcomes (see *Threats to validity* in Common evaluation terms at the end of this document).

KEY CONSIDERATION

Wherever possible, an evaluation should collect evidence using a range of quantitative and qualitative methods. Through a process of triangulation, the results can:

- reduce potential bias associated with using only one source of evidence
- enable discovery of new information about the program, its effects and other factors related to the theory of change.

Ethical principles

Protecting evaluation participant privacy and upholding ethical standards of conduct are essential elements of research involving human participants. Evaluations often present a number of moral, ethical and political concerns that need to be appropriately managed.

A key principle for ethical evaluation is ensuring that no physical, psychological or reputational harm arises from evaluation activities. Well-designed evaluations have governance and risk management systems in place to prevent, monitor and manage the unintended consequences that could arise.

All evaluation activities, irrespective of the specific stakeholders involved, should exercise sensitivity and respect towards different beliefs, perspectives and cultures.

Stakeholder engagement and communication

Evaluation plans work best when negotiated with those involved in implementing the program, including key stakeholders, government and non-government agencies.

Developing a stakeholder engagement and communication plan can add clarity about:

- who the key stakeholders for the evaluation are, for example program sponsors, implementers, users and beneficiaries
- the extent and type of engagement required, for example whether stakeholders will provide input in co-designing the program and/or its evaluation
- how ethical, culturally appropriate and accessible engagement will occur
- roles and responsibilities, for example provision of administrative data, culturally relevant advice or access to stakeholders for consultation
- what will be communicated to stakeholders, when, how and by who
- expectations in relation to the timing and detail of evaluation findings to be released.

Authorising environment

Recognising and reporting public value using appropriate authority structures brings legitimacy, transparency and accountability to government activities.

Having a robust authorising environment that provides sound oversight for evaluation design, implementation and reporting activities is important. This should be built into an evaluation strategy and be factored into the design of new programs. Those providing the oversight should have a strong understanding of evaluation concepts, and be committed to driving evidence-based policy and evaluation-informed decision-making.

Clear governance structures can ensure:

- stakeholders understand their evaluation roles and responsibilities including who is responsible for leadership, decision-making and undertaking evaluation activities
- performance can be optimised through appropriate planning, monitoring, risk management, innovation and learning
- accountability is defined and transparent
- an authorising environment exists for formal data requests and ethics approvals
- internal and external perceptions of legitimacy exist around evaluation activities.

When existing program governance arrangements are unavailable or unsuitable, it may be necessary to build new processes to support planned evaluation activities.

KEY CONSIDERATION

Where the program involves multiple delivery partners, consideration should be given to formalising roles and responsibilities through Memoranda of Understanding, agreements, contracts or committees.

Resources and budget

The scope and level of resourcing will have a bearing on the quality and usefulness of the evaluation. As a general guide, the resources and budget devoted to this should be informed by the program's scale, complexity, risks, resourcing, budget and intended outcomes.

Programs that have significant budgets, are complex, large-scale, of strategic significance or high risk are generally prioritised by government agencies and will typically have a larger budget for evaluation.

Comparatively, government typically exercises more discretion about the level of resourcing required to evaluate programs with lower spending and lower impact. However, there are cases where a high level of resources and budget are required to evaluate a program with seemingly low total expenditure. This can include where the findings from the evaluation:

- will be used to inform decisions about whether to roll out the program to a wider area and/or client group such as with a pilot or trial
- are to be generalised or used as evidence of another program's effectiveness.

Timeframes and reporting

The time needed to evaluate will vary depending on the program's complexity, whether any approval processes are required, and the strength of evidence needed for the evaluation.

Realistic timeframes will consider factors such as:

- appropriate design and planning phases
- stakeholder engagement
- authorising support and approval
- data collection, analysis and reporting.

Short-to-medium term monitoring and evaluation can report on implementation efficiencies, outputs delivered, knowledge and skills acquired, and any initial attitudinal shifts. However, behavioural and population level changes may take years or even generations to be realised. Measuring these changes typically involves collecting baseline data and progressively collecting follow up information as the mechanisms for change start taking effect.

For evaluations that extend over a longer timeframe, it is important to have pre-determined reporting points built into the program and evaluation plans.

These points provide an opportunity to share monitoring information on how the evaluation is tracking and see whether any interim findings can be used to improve the program or the evaluation, such as to improve program delivery or data collection.

In planning for reporting, evaluators should consider the evaluation's timeframe, the expectations of the audience, the technical expertise available for report writing, and how each report will be used. Given decision-makers may use reported information to inform program resourcing decisions, evaluators should ensure data reported is of high-quality and relevant to the decision-makers' needs.

A schedule for reporting and sharing findings can be included as part of the stakeholder engagement and communication strategy, as a dedicated release plan of the findings, and/or within the evaluation's framework or plan.

Managing expectations

Evaluation sponsors should be made aware of any limitations of the evaluation's design during the planning stage, as well as any issues that may impact data collection and interpretation throughout the evaluation.

This will help with the initial and ongoing management of expectations about:

- what the evaluation can achieve
- when and how results can be produced, including in what format they will be delivered
- what can be reasonably stated in reporting.

When observed outcomes are due to both the program and other external factors, the relationship between the program and outcomes is based on association, correlation or contribution, but not causation. When there is no relationship between a program and an observed outcome, it should not be described as a program outcome. When there are many factors that contribute to an outcome, an evaluator should describe the extent to which the program contributed to observed outcomes, how it contributed, for who and under what conditions.

Ongoing learning

A key feature of evaluation is the ability to obtain, review and communicate feedback about the program for ongoing learning and continuous improvement. This can help to:

- refine the theory of change and data collection strategies
- identify emerging risks, needs and opportunities
- maintain and strengthen relationships among program and evaluation stakeholders
- enhance the relevance of evaluation findings for future program or evaluation design.

Managing uncertainty and change

Programs implemented in uncertain, contested, emergent or dynamic contexts can create challenges for evaluators attempting to follow tightly structured plans. These contexts can lead to unplanned, but necessary changes to a program's design, implementation, stakeholders, authorising environment and consequently, the evaluation design and methods.

In some cases, an evaluation of effectiveness may need to be reset with a new baseline. In other cases, an adaptive, agile evaluation framework of change built around evaluation questions can identify emerging pathways of change and associated enablers and mechanisms.

For example:

Families and Schools Together (FAST) is an early intervention and prevention program designed to strengthen family functioning and build protective factors for children. FAST is an international program that uses established evaluation processes, but it has been adapted for the Northern Territory by FAST NT. When FAST NT first prepared to evaluate their program, it became clear that many of the evaluation tools used to assess outcomes (e.g., validated psychometric tools) would not work in the remote Indigenous communities where they were rolling out the program because the language and concepts that the tools used were not always meaningful for participants. Using these tools, therefore, risked collecting meaningless data.

To overcome these issues, FAST NT began work with an external consultant to develop new evaluation processes that would translate their pre-packaged measurement tools into something meaningful to the Indigenous community.

Part of the challenge was developing an evaluation process that built a robust evidence base using measurement tools that were meaningful to local participants while also satisfying reporting requirements.

The new process needed to be:

- **culturally relevant**
- **meaningful to participants**
- **able to measure change**
- **easily administrated**
- **consistent with measures included in the international tool**
- **able to meet the analytic requirements of reporting frameworks.**

Part of the task was replacing the psychometric survey tool (which was not producing reliable results) with a narrative inquiry tool that used pictures and symbols that were meaningful to local participants but could also be translated into definitive outcomes.

Adapted from the Australian Institute of Family Studies, *Evaluating the outcomes of programs for Indigenous families and communities - Adapting evaluation methods and measures, 2017.*

Commissioning and managing evaluation

The skills, expertise and commitment of those conducting an evaluation are critical to achieving its objectives. Determining who should manage an evaluation will depend on a variety of factors such as its priority, resourcing requirements, capabilities and the need for independence.

Choosing an evaluator

Ideally, an evaluator should be chosen during program development to ensure that program and evaluation planning can occur simultaneously. When deciding whether to commission an evaluation externally through an evaluation service provider or use internal resources such as an agency evaluation or research unit, there are several factors to consider (Table 1).

Table 1 – Benefits and challenges in using internal and external evaluators

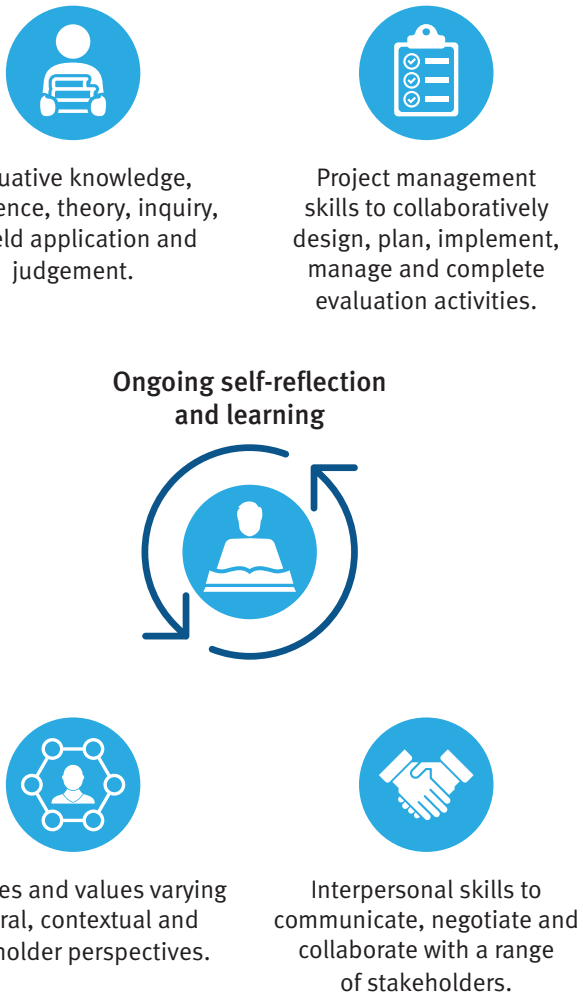
Internal Evaluator	External Evaluator
Benefits	
• Has detailed knowledge about program design and implementation. • Can help to build evaluation capacity, knowledge and skills within the agency and across government. • Can be less costly. • May be able to achieve better employee and stakeholder 'buy-in' for the evaluation.	• Perceived independent perspectives and insights. • Offers specialist technical or professional skills and expertise. • Provides their reputation and experience. • May deliver more efficient evaluation activities.
Challenges	
• May have less evaluation experience. • May draw resources away from program delivery. • May inhibit the candour of stakeholders who are consulted as part of the evaluation. • May reduce the actual or perceived validity of the evaluation's results.	• Can be more costly. • May not have an adequate understanding of the program, its implementation context or target groups. • May have difficulty engaging employees and stakeholders in the evaluation process.

Where an agency does not have the time, skills or resources necessary to carry out larger-scale or complex evaluations without assistance, it might be appropriate to commission external evaluators for selected activities, such as data collection or analysis, while the agency still designs, manages and reports the evaluation.

Evaluator competencies

Evaluators provide their clients with skilled, professional services, founded on a range of underlying competencies (see Figure 2).

Figure 2: Evaluator competencies



Adapted from the *AES Evaluators' Professional Learning Competency Framework, 2013.*

Commissioning an evaluation

Where a decision has been made to commission an external evaluator, agencies should first refer to relevant government procurement policies and procedures. For example, the Queensland Procurement Policy 2017 sets out the requirements for procurement of goods and services.

In preparing a tender document to commission an evaluation, agencies should be clear about the program, its policy context, the evaluation's scope, objectives, timeframes, methodological preferences and any other expectations.

KEY CONSIDERATION

Commissioning agencies are likely to be more successful in selecting a suitable provider if they are familiar with the program and policy evidence base and understand what is achievable within available time, resources and budget.

Once an appropriate evaluator has been selected, relevant components from the tender document can be included in a formal contract along with the:

- roles and responsibilities of all involved
- specifications of the arrangement, including costs, agreed outputs and delivery dates
- governance procedures, including an approach for identifying and managing risks and opportunities
- agreed dispute resolution arrangements.

Managing an evaluation

An evaluation process can be managed like any other time-bound project, following good project management principles, practice and tools. Critical to the success of evaluation is open and continuous communication, preferably from the program development phase, between the evaluation's sponsor, program managers, the evaluator and other key stakeholders.

Effective evaluation managers will facilitate regular meetings and program monitoring activities and update and share key documents with any program or evaluation developments, including the program, evaluation, risk management, stakeholder engagement and communication plans.

This will help to ensure that:

- any risks or changes in the policy, program or stakeholder context are detected and managed early
- key evaluation documents, questions, models and frameworks can be revised as more is learnt about the program and its implementation context.

Communicating and sharing evaluation findings

Evidence-informed decision-making should be at the forefront of any program evaluation. While this typically relates to decisions by policy-makers about a program's design, implementation and opportunities for improvement, understanding the information needs of other evaluation stakeholder groups (who may be affected by its findings) is equally important when reporting results.

Reporting results

For many evaluations, an evaluation report will be used to communicate key findings with stakeholders and decision-makers. Reporting would typically include:

- a brief executive summary
- introduction to the evaluation (i.e. purpose and objectives) and the program's policy context
- an overview of the program evaluated, including an updated program logic model
- the evaluation's key questions, scope and target populations, clearly identifying what is and isn't being evaluated
- the evaluation's design, including any theories, methodologies or assumptions, and key information about the policy or program evidence base (where applicable)
- the data collected, and any ethical or cultural considerations managed as part of this process
- results, their interpretation, limitations, opportunities for improvement and any recommendations.

In general, evaluation reports will be objective, balanced and fit-for-purpose. They will consider how any reported information will be used and shared by others.

It is essential that evaluators understand not only the relevance of the evaluation's findings, but the expectations and information needs of its target audience. Knowing what information to include, and how to include it, is critical to effectively communicating the evaluation's results.

In some cases, a static, text-heavy report may not be the most appropriate means to communicate results. Evaluators should consider the use of infographics, summary reports, data visualisation, interactive tables and figures and other presentation types, if these would benefit the reader.

Sharing results

Agencies are encouraged to share the results of their evaluations in the public domain (where permission has been obtained from the evaluation's sponsor). Doing this:

- assists in building the evidence base available to decision-makers
- enhances accountability and transparency of program operations
- provides explanations of the findings to the public, and of the program and evaluation expenditure choices made
- allows for outcome comparison and knowledge sharing across similar programs
- reduces the potential for any duplication and overlap between agencies
- encourages discussion about opportunities to adjust activities in response to evaluation findings.

Communicating changes being implemented as a result of evaluation findings is important, particularly with those who have participated in the evaluation. Providing feedback to participants (where disclosure is allowed) recognises and values their contribution.

KEY CONSIDERATION

Program evaluation can identify opportunities for improvement in its design or operations. This should not be viewed as a problem of evaluation, but rather a chance to make changes to achieve better outputs and outcomes.

Enhancing evaluation capability

Working with agencies to enhance evaluation capability is an important part of developing a culture of evaluation across government. This culture can help to ensure that sound processes for evaluation design, use and implementation are used to develop and assess government programs.

Building capability in evaluation

Designing, implementing and reporting program evaluation requires an understanding of the key elements of the evaluation process. Fostering internal capability in high-quality evaluation can help agencies to:

- articulate the role of evaluation in delivering public value for Queenslanders
- explain the strengths and limitations of evaluations and their role in evidence-informed decision-making
- highlight the alignment, or lack there-of, between evaluation, policy and program design
- enhance skills, knowledge and experience in undertaking evaluation activities
- facilitate sharing knowledge and evaluation across the agency.

This improves an agency's ability to assess the effectiveness of organisational activities and to use evidence to improve performance.

Meaningfully building evaluation capability within an agency, and across government, takes time and investment. For evaluation activities to become embedded in agency practice, the right policies, procedures, people, equipment and culture need to be in place to ensure ongoing monitoring and feedback on agency processes.

For capability building to be successful, agencies need to recognise the benefits of adopting evaluation as a necessary and ongoing function.

Creating a culture of evaluation

The Queensland Government encourages agencies to develop a culture of evaluation within a broader context of evidence-informed policy development.

To build a culture of evaluation, Queensland Government agencies should:

- adapt their capability building processes to suit their nature, size and structure
- allocate adequate resources (staff, time, commitment, equipment and effort) to evaluation skills and knowledge development
- ensure their program, group, office and departmental leadership teams play an important role in nurturing and fostering this culture, through explicit and ongoing support
- access quality evaluation information resources, including communities of practice, and professional, academic and non-government bodies
- obtain buy-in from all staff, especially those selected to enhance their evaluation skills, knowledge and capabilities.

Creating a culture of ongoing learning in evaluation methods and techniques can help individuals and agencies stay relevant, enhance their ability to meet requirements for current and future evaluations, and foster continuous improvement of evaluation processes.

Further resources

Many resources about how to approach and conduct evaluations have been developed to guide those working in the evaluation space, a selection of which are included here.

ACT Government (2017). *Evaluation Policy and Guidelines*. ACT Chief Minister's Department, Canberra.

Australian and New Zealand School of Government, Evidence and Evaluation Hub.

Australian Evaluation Society (2006). *Guidelines for the ethical conduct of evaluations*.

Australian Government, Department of Industry, Innovation and Science (2017). *Evaluation Strategy 2017-2021*. Office of the Chief Economist, Canberra.

Australian Institute of Aboriginal and Torres Strait Islander Studies (2011). *Guidelines for Ethical Research in Australian Indigenous Studies*.

Australian Institute of Family Studies website. https://aifs.gov.au/cfca/publications/evaluating-outcomes-programs-indigenous-families-and-communities

Australian National Audit Office (Australian Government) (2014). *Better Practice Guide: Public Sector Governance*.

Better Evaluation website, www.betterevaluation.org

Centre for Epidemiology and Evidence. *Commissioning Evaluation Services:* A Guide. Evidence and Evaluation Guidance Series, Population & Public Health Division. Sydney: NSW Ministry of Health, 2017.

Community Tool Box website, https://ctb.ku.edu/en/evaluating-community-programs-and-initiatives.

Cook, T.D. and Campbell, D.T. (1979). *Quasi-Experimentation: Design & Analysis Issues for Field Settings*. Chicago, Rand McNally.

Denzin, Norman K. (2012) *Triangulation 2.0*, in Journal of Mixed Methods Research Volume 6, Issue 2, April 2012, Pages 80-88.

Government of Western Australia (2015). *Evaluation Guide*. Program Evaluation Unit, Department of Treasury, Perth.

HM Treasury (2011). *The Magenta Book: Guidance for Evaluation*. HM Treasury, London.

Kennedy, Patrick. (2009) *How to combine multiple research options: Practical Triangulation*.

McNeish, D. and Scott, S., *Project Evaluation – A Practical Guide*.

NHMRC (2007). *Australian Code for the Responsible Conduct of Research*

NHMRC (2014). *Ethical Considerations in Quality Assurance and Evaluation Activities*.

NSW Government (2016). *NSW Government Program Evaluation Guidelines*. NSW Government, Sydney.

Department of Housing and Public Works (2017). (Queensland Procurement Policy 2017.) Queensland Government, Brisbane.

Department of the Premier and Cabinet (2017). *Queensland Government Performance Management Framework (PMF)*. Queensland Government, Brisbane.

Department of the Premier and Cabinet (2015). *Criminal Justice Evaluation Framework (CJEF)*. Queensland Government, Brisbane.

Queensland Treasury (2015) *Project Assessment Framework – Policy Overview*. Queensland Treasury, Brisbane.

Shadish, W.R., Cook, T.D, and Leviton, L.C. (1991). *Foundations of program evaluation: Theories of practice*. Newbury Park, CA: Sage.

The Families Commission (2017). *Making sense of evaluation: a handbook for everyone,* Social Policy Evaluation and Research Unit (Superu), Wellington.

U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, *Program Evaluation*.

Common evaluation terms

Term	Brief description
Appropriateness	The extent to which a program is suitable for achieving stated objectives in a given context.
Baseline data	Data collected before a program is implemented to compare with data collected once that program is implemented.
Confounding factors	Any non-program factor that may affect a program's outcomes. Such factors may include individual demographics, contextual information and/or measurement decisions.
Contribution analysis	Analysis of evidence to determine the extent to which program activities have contributed to observed outcomes.
Counterfactual	An assessment of what would have happened had the program not been implemented. This is often achieved by comparing the experiences of individuals, groups or communities who did not participate in the program (comparison group) with the experiences of those who did.
Effectiveness	The extent to which a program is responsible for achieving its intended outcomes. Assessments of effectiveness should consider alternative explanations for observable outcomes and what may have occurred in the absence of the program (i.e. the counterfactual).
Efficiency	The extent to which a program is delivered at the lowest possible cost (technical or productive efficiency), to the areas of greatest need (allocative efficiency) and/or continues to improve over time by finding better or lower cost ways to deliver outcomes (dynamic efficiency).
Equity	The extent to which a program meets the needs of participants in relation to their individual needs. It can be distinguished from equality where participants are treated equally.
Evaluability assessment	A pre-evaluation or early evaluation assessment of the extent to which a program can be evaluated in a reliable and credible fashion. This assessment typically looks at available budget, data and resources as well as the evaluation's likely design, questions, methods and timeframe.
Evaluation framework	Outlines the parameters and approach for what will be evaluated and often covers multiple programs. Frameworks typically include the program's objective, context and logic as well as the evaluation's scope, objectives, principles, guiding questions, monitoring and measurement plan, engagement and communication strategy and governance arrangements.
Evaluation plan	Outlines how an evaluation will be conducted in more operational detail and is usually applied to a specific program. It typically includes the program's objective, context and logic as well as the evaluation's scope, objectives, principles, guiding questions, design, methodology, data collection, analysis and reporting.
Experimental design	A type of evaluation design that involves the random assignment of equally eligible subjects into an intervention group (program participants) or control group (non-program participants). Data is collected both before and after program implementation and results between the two groups compared. A common example of experimental design is a randomised-control trial.
Inputs	The financial, human, organisational or physical resources needed to deliver a program.
Monitoring	The regular collection, analysis and use of program and system level data in relation to their management, performance, productivity and achievements.
Non-experimental design	A type of evaluation design that involves the non-random sampling of subjects from a pre-existing intervention group (program participants). It does not include a comparison group of non-program participants. Common examples of non-experimental designs include interpretation, observation or interaction studies, time-series or retrospective studies, and theory-based evaluation approaches.
Objectives	Clear, measurable statements of what the program or evaluation desires to achieve.

Outcome	The intended or unintended, positive or negative results that may be directly or indirectly influenced by a program. Outcomes are often described in terms of timeframes (short, medium or long term) and according to the type or level of change occurring (e.g. behavioural change at an individual level or social change at a societal level).
Outputs	The quantifiable products or services delivered as a result of a program's processes or activities. Outputs capture what the program produces, rather than what difference the program made.
Policy	A statement of government intent in relation to an issue. The development and implementation of programs is one way that government can act in response to a policy decision.
Program	A set of discrete, time-based activities that aim to achieve an outcome in response to an identified need or policy position. In this guideline, the principles applied to a program may equally be applied to a system, policy, strategy, initiative, intervention, service or project.
Program evaluation	Processes undertaken to assess a program's effectiveness, efficiency, appropriateness, utility, equity, value for money and/or sustainability to support accountability, learning, and performance improvement. An evaluation should provide credible, useful, evidence-based information that enables the timely incorporation of its findings, lessons and recommendations to be incorporated into the decision-making of organisations and stakeholders in a timely manner.
Program logic model	A program logic model (or logic model) is a visual tool to show, in simple terms, how a program operates to produce change. It usually includes a program's theory of change and the assumptions which underlie the conversion of the program's inputs and activities to outputs and outcomes.
Quasi-experimental design	A type of evaluation design that involves the non-random selection of subjects (based on their characteristics) into a pre-existing intervention group (program participants) and a comparison group (non-program participants). Common examples of quasi-experimental designs include difference-in-difference, regression discontinuity and pre/post testing with matched samples.
Realist evaluation	Evaluation which seeks to understand what program elements are effective, for whom, why, under what circumstances and at what cost, recognising that there can be multiple pathways to any given outcome. This approach seeks to identify the underlying mechanisms triggered by its implementation context to produce an outcome.
Stakeholders	Program stakeholders are individuals, groups or organisations who can affect, or be affected by, a program over time. Those with a direct or indirect interest in a program may include program managers, sponsors, participants, service providers and sometimes the wider society.
Theory of change	An explanation of how activities are understood to contribute to a series of outcomes. It is also referred to as program theory and can be represented in a program logic model.
Threats to validity	Factors that can compromise the ability of an evaluation to demonstrate that a) a relationship exists between a program's actions and the observed outcomes (internal validity) or b) the observable outcomes can be applied to other program participants (external validity). Threats to internal validity often relate to the program participants and their characteristics. Threats to external validity often relate to the context within which programs are implemented. Well-designed evaluations can manage the risks associated with such threats to validity.
Utility	The extent to which a program met the information needs or was useful to program participants. Utility should be viewed both in terms of demand and supply. It can include measurement of use, functionality, strengths, weaknesses and learning.
Value for money	Value for money is achieved when the maximum benefit is obtained from the program provided within the resources made available. Usually involves an assessment of input costs, output and outcome quality and benefits. Common methods for measuring value for money include cost effectiveness analysis, cost benefit analysis and social return on investment.