

Queensland Government  
**Program Evaluation Guidelines**

November 2014

## Document details

<b>Security Classification</b>	PUBLIC
<b>Date of Review of Security Classification</b>	November 2014
<b>Authority</b>	Under Treasurer, Queensland Treasury and Trade
<b>Author</b>	Economics Division, Queensland Treasury and Trade
<b>Version</b>	1.0

## Contact for enquiries and proposed changes

The Economics Division in Queensland Treasury and Trade is the custodian of this document. All inquiries regarding this document should be directed in the first instance to:

Email: [PEG@treasury.qld.gov.au](mailto:PEG@treasury.qld.gov.au) Tel: 07 303 51457.

## Acknowledgements

The *Queensland Government Program Evaluation Guidelines* were developed by the Economics Division, Queensland Treasury and Trade, with assistance from the Queensland Government Statistician's Office.

Feedback was also received from staff from within the Department of the Premier and Cabinet and from various agencies. All feedback was greatly appreciated.

## Copyright

*Queensland Government Program Evaluation Guidelines*

Copyright © The State of Queensland (Queensland Treasury and Trade) 2014

## Licence



This document is licensed under a Creative Commons Attribution 3.0 Australia licence. You are free to copy, communicate and adapt the work, as long as you attribute the authors. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/3.0/au>.

## Contents

EXPLANATION OF KEY TERMS .....	1
OVERVIEW .....	2
INTRODUCTION .....	3
Why evaluate?.....	3
How the guidelines are to be used with other Queensland Government frameworks .....	5
Structure of the guide.....	6
1 THE EVALUATION PROCESS .....	7
1.1 Planning the evaluation.....	8
1.2 Conducting the evaluation .....	14
1.3 Using the results.....	14
2 PRACTICAL CONSIDERATIONS.....	16
2.1 Clarifying roles and responsibilities .....	16
2.2 Evaluation timeframes .....	17
2.3 Refining the evaluation process .....	18
2.4 Scale of the evaluation .....	18
3 WHERE TO GO FOR FURTHER INFORMATION .....	20
ATTACHMENT A: EVALUATION APPROACHES .....	21
A.1 Evaluating design, implementation and delivery .....	21
A.2 Evaluating effectiveness .....	22
A.3 Evaluating efficiency .....	25
A.4 Further resources.....	29
ATTACHMENT B: COLLECTING EVALUATION DATA .....	30
B.1 Types of data.....	30
B.2 Collection methods .....	31
B.3 Selecting the right method .....	33
B.4 Further resources.....	35

## EXPLANATION OF KEY TERMS

**Effectiveness:** The extent to which a program is responsible for a particular outcome or outcomes. To ascertain effectiveness requires consideration of other potential influences on the outcomes of interest and the counterfactual (what would have happened in absence of the program).

**Efficiency:** The extent to which a program is delivered at the lowest possible cost, to the areas of greatest need, and continues to improve over time by finding better or lower cost ways to deliver outcomes.

**Evaluation:** The systematic, objective post-implementation assessment of the appropriateness, relevancy, process, effectiveness and/or efficiency of a program. Post-implementation evaluation is not disparate to monitoring or assurance. It is likely that monitoring of outputs, processes and outcomes or internal and external assurance functions could either constitute an evaluation or provide valuable input into an evaluation.

**Evaluation process:** the steps involved in planning and conducting an evaluation, and disseminating evaluation findings.

**Outcomes:** The short, medium and/or long term results generated as a direct result of the delivery of a program (i.e. what difference the program made). Possible outcomes of programs can include changes in awareness, knowledge, skills, attitude and behaviour, as well as economic environmental and social impacts. For the purposes of the guidelines, the terms outcomes and impacts are used interchangeably.

**Outputs:** The services or facilities provided as a result of a program's processes or activities. Outputs capture what the program does and who it reaches, rather than what difference the program made (i.e. outcomes).

**Policy:** A statement of Government intent in relation to an issue, which can be implemented through the use of policy instruments, such as laws, advocacy, monetary flows and direct actions. The development and implementation of programs is one way that Government can act in response to a policy decision.

**Program:** A structure (such as an intervention, initiative, strategy or service) created to coordinate, direct and oversee the implementation of a set of related projects and activities in order to deliver value for the agency and/or its stakeholders (including clients) in response to an identified need and policy position.

**Program design:** The process undertaken to develop a program prior to program implementation. Program design will often include development of an implementation plan, consideration of resource or training requirements, and agreement on a governance structures. Program design should also include the development of an evaluation plan. Program design is also commonly referred to as program development or program planning.

**Program logic:** A method to assist program design. It depicts the logic or pathways through which the programs processes (inputs, activities and outputs) are intended to achieve the desired outcomes. Logic models can assist in understanding how the program is intended to work, what it is trying to achieve and why. Program logic is also commonly referred to as program theory or service logic.

## OVERVIEW

Evaluation is an essential part of the management and delivery of public sector programs. Well-designed evaluations are an essential tool for public sector agencies to strengthen efficiency of program delivery and to demonstrate the effectiveness of programs in generating outcomes.

Evaluations can provide useful information for program managers on whether a program is doing the right things in the right ways, and whether there are ways to improve program delivery. Evaluations can also provide information to the public sector as a whole to inform broader decisions on the best way to achieve desired outcomes and address identified needs.

There is no single evaluation method appropriate for evaluating the range of programs delivered by the Queensland public sector. These guidelines provide a framework within which evaluations can be planned and implemented in a manner appropriate to the program being evaluated.

Evaluations require the commitment of resources and balancing the costs and likely outcomes requires the exercise of judgement. Agencies and program managers will need to take a strategic approach in determining appropriate evaluation scopes, designs and resourcing requirements. The guidelines have been designed to provide agencies with the flexibility to exercise judgement in the planning and implementation of evaluations, albeit in an accountable and transparent framework of minimum requirements.

Regardless of the program's characteristics, it is expected that program managers and public service officials understand the need for evaluating public sector programs and will have considered evaluation requirements during program development and implementation.

When considering evaluation requirements, it is expected that any evaluation of public sector programs will<sup>1</sup>:

- Specify criteria for determining the success of the program
- Focus on the key issues that will inform decision making
- Use a systematic and evidence-based approach to assess performance
- Be reliable, useful and relevant to decision makers and stakeholders
- Be timely.

To enable these criteria to be met, evaluations should be built into the program design and have the following features:

- A clear, considered evaluation plan and, where relevant, a well-drafted terms of reference
- Clearly defined roles and responsibilities
- Strong stakeholder engagement
- Evaluation milestones timed to be able to inform decision making
- Strategies in place to compensate for any potential deficiencies in evaluation design, data collection and analytical methods
- Checks and balances in place to ensure validity of evaluation findings
- Clear, transparent reporting that outlines methods, assumptions and key findings.

---

<sup>1</sup> Queensland Audit Office, 2013, Report 13: 2012-13 – Drink Safe Precincts Trial, available at <https://www.qao.qld.gov.au/files/file/RTP2012-13DrinkSafePrecincts.pdf>

## INTRODUCTION

These guidelines outline a set of broad principles that are expected to underpin the planning and implementation of evaluations for programs funded by the Queensland Government.

The guidelines have been designed to clarify expectations for the evaluation of public sector programs in Queensland, and provide a framework to guide Queensland Government agencies through the minimum requirements expected to be met for the planning, implementation and management of program evaluations.

The guidelines are intended as a resource for those responsible for developing, designing and implementing programs and for public sector officials responsible for making resource allocation decisions within Government.

The guidelines apply to any programs, interventions, initiatives, services or trials that are being developed in response to an identified need and policy position (referred to in the Guidelines simply as *programs*). The guidelines will apply most easily to Government programs involved in service delivery, with a direct client interface, but should be able to be applied for the most part to all programs delivered by the public sector. While there may be some learnings from the guidelines that can be translated to assessing policy, it is not intended that the guidelines will be applied directly to the evaluation of Government policies.

For the purposes of these guidelines, the term *evaluation* refers to any systematic post-implementation assessment of a program. The guidelines are not designed to provide advice on the appraisal of options or preliminary evaluations of program merits prior to program implementation.

### Why evaluate?

At the broadest level, Government business is about the development and implementation of public policy in response to identified issues. Government policy is a statement of government intent and is implemented through the use of policy instruments, such as laws, advocacy, monetary flows and direct actions<sup>2</sup>. The development and implementation of programs is one way that Government can act in response to a policy decision.

Evaluation plays a critical role in the development and implementation of Government programs, providing a systematic process of collecting credible data and using it to inform decisions on whether or not a program is achieving its objectives and remains the best policy response.

Evaluation can be used to assist program managers in understanding program processes and the mechanisms by which it operates, as well as assessing whether the program is **on the right track** and **doing the right things** to achieve its objectives.

Evaluation can also be used to assist Government more broadly to assess the **appropriateness** and **value for money** of programs to influence decisions on resource allocation.

Evaluation does not have to be an onerous process. Evaluation requirements will be determined by the scale and cost of the program, the risks associated with the program and the types of information required by program managers and decision makers. For some programs, evaluation could simply involve routine assessment of a program's inputs, activities and outputs built into program reporting,

---

<sup>2</sup> P Bridgman & G Davis (1998) The Australian Policy Handbook, 3<sup>rd</sup> Edition.

while for others evaluation will need to be comprehensive and assess whether the program is appropriate, effective and efficient.

All well-designed evaluations, regardless of their scale, play a role in strengthening public sector efficiency, effectiveness and accountability by<sup>3</sup>:

- Enhancing the ability to achieve government priorities and policy outcomes
- Strengthening resource allocation, planning and decision making
- Assessing and improving the performance and impact of service delivery
- Demonstrating outcomes to key stakeholders.

Evaluation will be most effective when embedded within program development, rather than occurring as a separate process alongside program design, implementation and operation. Incorporating ‘evaluation thinking’ into program function from an early stage will enhance program development and provide greater potential for evaluation findings to be used to continually improve and refine program delivery.

To ensure evaluations can be relevant and timely, planning for an evaluation should commence during program design, and where possible, baseline data should be collected prior to program implementation. Once the program has been implemented, there should be ongoing feedback between the evaluation and the program, to refine data collection, improve program delivery and enhance the relevance of evaluation findings.

Evaluation findings should also be used to inform future program development and to drive continual improvement in the delivery of Government programs.

**Figure 1: Incorporating evaluation into program development and implementation**



<sup>3</sup> Queensland Audit Office, 2013, Report 13: 2012-13 – Drink Safe Precincts Trial, available at <https://www.qao.qld.gov.au/files/file/RTP2012-13DrinkSafePrecincts.pdf>

## How the guidelines are to be used with other Queensland Government frameworks

These guidelines are designed to complement currently available resources on program development and evaluation, and outline a set of generic principles that should underpin all evaluations, regardless of the specific evaluation approach or program design.

For example:

- The Queensland Government Contestability Framework outlines a fair, transparent and thorough process for considering how services are best delivered, and who is best to deliver them. The Queensland Government Program Evaluation Guidelines will support the evaluation of programs assessed as being contestable, to ensure that the program continues to achieve its service objectives and remains relevant after implementation. Working through the Contestability Framework may also help to evaluate the efficiency of a Government program.
- The Queensland Government [Performance Management Framework](#) (PMF) is designed to improve the analysis and application of performance information to support accountability, inform policy development and implementation and deliver value to stakeholders. The PMF ensures a clear line of sight between planning, measuring and monitoring results and public reporting. The information contained in the Queensland Government Program Evaluation Guidelines supports the measuring and monitoring results element of the Queensland Government Performance Management Framework.
- The [Project Assurance Framework](#) (PAF) is a whole-of-Government project assessment process. It establishes a common approach to assessing projects at critical stages in their life cycle, aiming to maximise the benefits returned to government from project investments. The information contained in the Queensland Government Program Evaluation Guidelines may assist in developing and undertaking the post-project implementation reviews (ie. Benefits Realisation) required as part of the PAF. While many aspects of these Guidelines will be able to be applied to program appraisal, it is not expected that these Guidelines will be able to be applied directly to the project preliminary evaluation (or options analysis) stages outlined in the PAF.
- The Queensland Government [Project, Program and Portfolio Management Methodologies](#) (PPPM) outline standard approaches for managing ICT enabled projects, programs and portfolios in the Queensland Government. The methodologies offer a consistent approach to help projects come to life – on time and within budget – and drive programs to deliver the outcomes and benefits they are intended to achieve. The Queensland Government Program Evaluation Guidelines and the PPPM Methodologies will complement one another to support and enable the effective management, delivery and improvement of programs over time.

Agencies and statutory bodies undertaking evaluation on specific subject areas should also refer to any subject specific guidelines available, such as the Department of Premier and Cabinet's [Criminal Justice Evaluation Framework](#).

## Structure of the guide

The document comprises 3 main sections.

1. [The evaluation process](#) explains the key steps that should be involved when planning and implementing an effective evaluation. The section focusses heavily on the stages involved in planning an evaluation, particularly on the specification of purpose and outcomes for both the program and the evaluation, and governance issues regarding role clarification and risk mitigation.
2. The [practical considerations](#) section deals with some of the more practical issues that will need to be considered by public sector officials when planning and conducting an evaluation, such as appropriate roles, responsibilities and resourcing.
3. Finally, a list of resources is provided as a starting point for [where to go for further information](#) to assist in planning or conducting an evaluation.

Attachments on [Evaluation Approaches](#) and [Collecting Evaluation Data](#) provide greater detail on the different types of methods that can be used to collect data and develop evaluation findings.

# 1 THE EVALUATION PROCESS

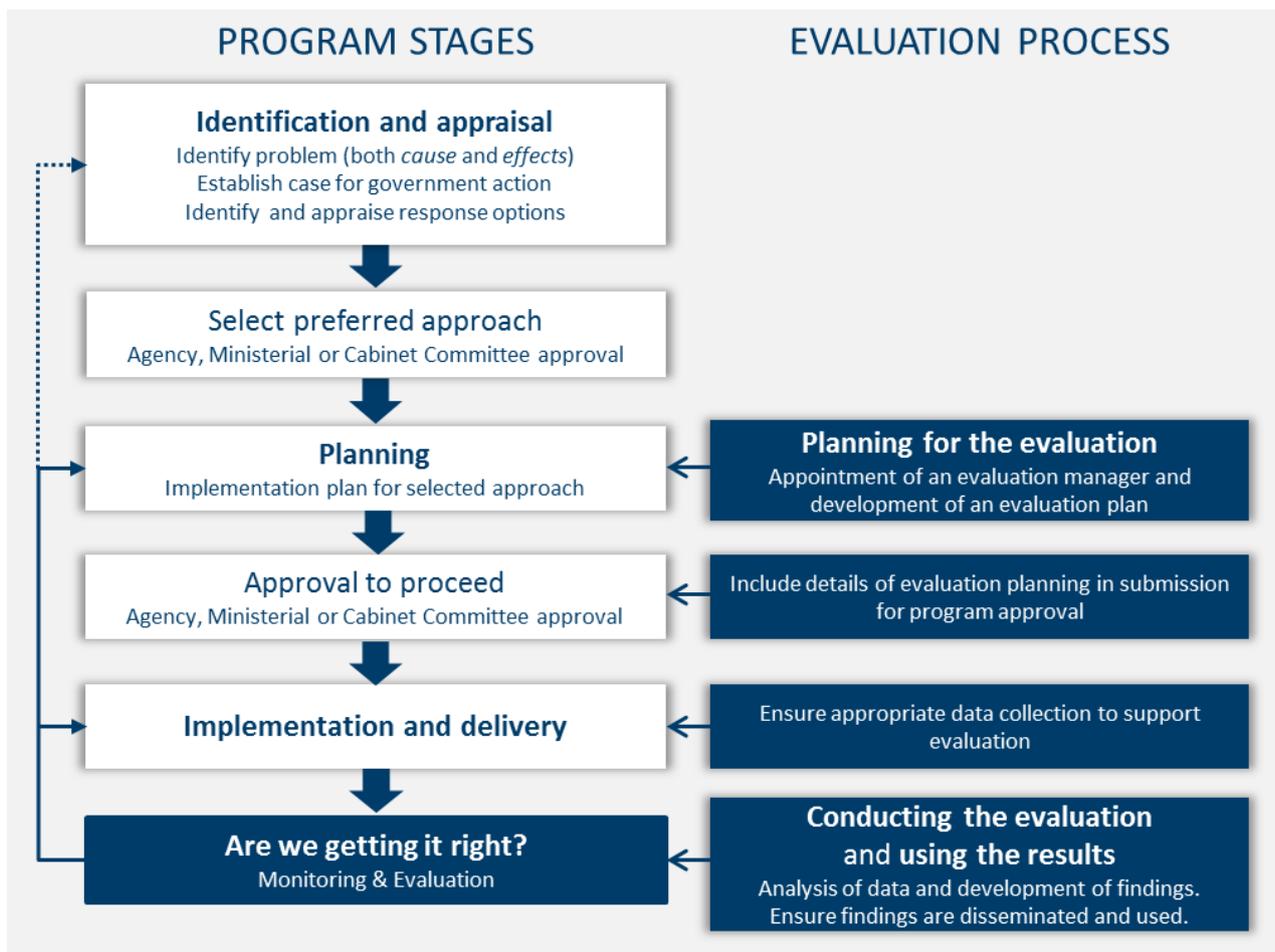
The evaluation process can be broken down into three core stages:

- Planning for the evaluation
- Conducting the evaluation
- Using the results

Planning for an evaluation should always occur during program design. Once the program has been implemented, the development of evaluation findings should be timed to inform program decisions and drive continuous improvement in program delivery.

For existing programs, program managers should endeavour to review whether evaluation arrangements are in place and, if not, make arrangements to ensure an evaluation plan is developed and implemented.

**Figure 2: The Evaluation Process**



To deliver findings that inform decision making and inform continuous improvements of programs, at a minimum, every evaluation process should:

- Be underpinned by a clear, considered evaluation plan
- Involve clear, transparent reporting that outlines methods, assumptions and key findings
- Result in the production of findings that are useful in informing decisions regarding program delivery.

## 1.1 Planning the evaluation

Arguably, planning an evaluation is the most important step in the evaluation process. Different types of evaluation provide different information and support different decisions. This makes it important to plan upfront what questions need to be answered, how they will be answered, and by when<sup>4</sup>.

Planning for an evaluation should begin while the program is being designed. Evaluations that are planned simultaneously with the plan for program implementation will be more likely to result in meaningful evaluation findings than those that are planned after the program has been implemented.

Planning is particularly important when evaluating whole-of-Government programs. Having to coordinate the evaluation across multiple agencies and potentially across multiple delivery strategies will add additional complexities to any evaluation and increase the need for a detailed and comprehensive evaluation plan.

Every evaluation plan should identify:

- **Why** – the background and rationale for the program and the evaluation, and who will use the results.
- **What** – the scope and objectives of the evaluation and questions for the evaluation to answer.
- **Who** – those responsible for managing and those responsible for carrying out the evaluation.
- **How** – the methods of gathering and analysing data to answer the questions, strategies to manage risk, and the plan for dissemination, disclosure and use of results.
- **When** – the timing of evaluation in the program cycle, key milestones and deliverables.
- **Resources** – the people, materials, infrastructure and logistics needed for the evaluation.

There are a number of key steps involved in the development of an evaluation plan, outlined in some detail below. While each step is important, they do not necessarily have to be performed in the order listed. Some steps may need to be brought forward, or revisited, to ensure the evaluation plan effectively identifies all factors listed above.

### Step 1: Identify stakeholders

Engaging key stakeholders early on in the evaluation process can help to inform an appropriate evaluation design and improve the usefulness and acceptability of evaluation findings. Stakeholders are also much more likely to engage with and support the evaluation if they are involved in the evaluation process from the beginning<sup>5</sup>.

In general, stakeholders fall into three major groups:

- **Those involved in running the program**, e.g. program managers and staff, funding agencies, and delivery bodies.
- **Those served or affected by the program**, e.g. program participants (and their families), individuals who purchase services or products delivered by the program, and the general public.
- **Those who are interested in the program and would use the evaluation results**, e.g. senior public sector managers and Ministers, as well as community, industry or advocacy groups.

---

<sup>4</sup> NSW Government Evaluation Framework, August 2013, available at [http://www.dpc.nsw.gov.au/data/assets/pdf\\_file/0009/155844/NSW\\_Government\\_Evaluation\\_Framework\\_Aug\\_2013.pdf](http://www.dpc.nsw.gov.au/data/assets/pdf_file/0009/155844/NSW_Government_Evaluation_Framework_Aug_2013.pdf)

<sup>5</sup> CDC, 2011, Developing an Effective Evaluation Plan, available at <http://www.cdc.gov/obesity/downloads/cdc-evaluation-workbook-508.pdf>

Considering the needs of key stakeholders will influence decisions on the evaluation objectives, research questions and evaluation design. For example, stakeholders can inform:

- How the evaluation can be timed to feed into decision making
- Ways to increase the effectiveness of evaluation findings, including data requirements, presentation of the results, and mechanisms for and timing of dissemination
- How robust the findings need to be to support the intended end use and what level of scrutiny the findings will be subject to.

Stakeholder engagement and management will be more challenging where there are multiple agencies or proponents working together to deliver a program. In these cases, identification and ongoing management of stakeholder expectations will be critical to ensuring that evaluation objectives are relevant and that findings from an evaluation can be used to inform decision making.

## Step 2: Understand program objectives and intended outcomes

Before designing an evaluation it is important to develop a complete understanding of how the program works (or is intended to work), what it is trying to achieve (in terms of measurable objectives), and why (the underlying policy problem). This is often referred to as *program theory*, *program logic* or *service logic*.

This information should have already been determined as part of the program design, but needs to be carried across into the evaluation plan.

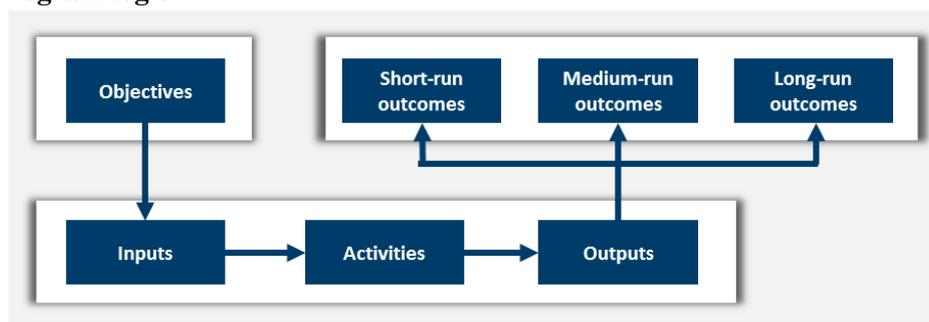
### **Tip: Understanding the purpose is critical**

Effective application of these guidelines will depend on the program having clearly defined and measurable objectives against which success can be assessed. Even the most rigorous evaluation process will fail if it is not clear what the program is trying to achieve and why.

To gain a complete understanding of the program, the evaluator and program manager should identify:

- **Need:** why the program is required
- **Objectives:** what the program aims to achieve and why
- **Inputs:** the resources needed to operate the program (labour, materials etc.)
- **Activities:** processes, tools, events, technology and actions integral to program implementation
- **Outputs:** direct products of program activities (such as types of services to be delivered)
- **Short-run outcomes:** such as changes in awareness, knowledge, skills, and attitude
- **Medium-run outcomes:** such as changes in behaviour
- **Long-run outcomes:** such as wider economic, environmental and social impacts.

Figure 3: Program logic



Where a program involves multiple delivery strategies, the inputs and actions of each strategy should be identified separately. This will assist in identifying whether there are improvements that can be made to individual elements to increase the success of the program as a whole.

It will also be necessary to identify any economic, social or political factors that might present challenges for the program and potentially result in realised impacts differing from intended impacts.

A detailed outline of different ways of developing and representing program logic models can be found at [betterevaluation.org](http://betterevaluation.org).

### **Step 3: Decide on evaluation objectives and key evaluation questions**

The purpose of an evaluation could be to examine one, or a combination, of the following elements of a program:

- The relevance and appropriateness of the program's objectives and activities in addressing recognised needs and priorities
- Program processes governing design, implementation and delivery
- The effectiveness of the program in achieving outcomes
- The efficiency of the program in delivering outputs and outcomes.

The specific objective/s of an evaluation will vary from program to program, and will depend on the type of program being delivered, the existing body of evidence on the program, and the requirements for stakeholders to assess the program and to make improvements.

Clear establishment of the objectives (or goals) of an evaluation are vital to ensure that the evaluation produces the right information for stakeholders and decision makers. For example, an assessment of efficiency may provide information on whether the services being delivered are cost-effective, but this might be irrelevant if the services are not meeting client, stakeholder or the broader communities' needs (relevancy, appropriateness and effectiveness). For this reason, many evaluations will often involve assessment of multiple elements of a program (such as effectiveness and efficiency) to ensure that a more complete picture of the program is provided to stakeholders and decision makers. That being said, there will always need to be a balance between the desired level of information and the resources required to produce evaluation findings.

#### **Tip: Less is sometimes more**

Trying to evaluate too much at one time can lead to challenges in evaluation design, management, and dissemination and use of findings. At times, it could be better to narrow the scope of the evaluation rather than trying to spread resources too thin and risk the production of valid and robust findings.

Once the objective has been established, specific questions need to be formulated for the evaluation to address. The questions that an evaluation will need to investigate will depend on the scale of the program and its intended impacts.

**Figure 4: Possible Evaluation Questions**

<i>Program</i>	<i>Possible evaluation questions</i>	
<b>Objectives</b>	<b>Relevance and appropriateness:</b> <ul style="list-style-type: none"> <li>• Does the identified need still exist, and is the program relevant to the current need?</li> <li>• Does the program align with Government objectives / goals?</li> <li>• Are the program activities and outputs consistent with the identified objectives?</li> </ul>	
<b>Inputs, activities and outputs</b>	<b>Process (design, implementation and delivery):</b> <ul style="list-style-type: none"> <li>• Was the program delivered as planned? If not, why?</li> <li>• How can the program (delivery, outcomes, reach etc.) be improved?</li> </ul>	<b>Efficiency:</b> <ul style="list-style-type: none"> <li>• How can the program be more efficient in achieving its objectives?</li> <li>• Are the current mix of programs efficient?</li> <li>• Do the benefits of the program exceed the cost of delivery?</li> <li>• Could others provide the services more efficiently?</li> </ul>
<b>Short, medium and long-run outcomes</b>	<b>Effectiveness:</b> <ul style="list-style-type: none"> <li>• What impacts resulted from the program?</li> <li>• Were there any unintended consequences of an initiative, and if so, why?</li> <li>• How were different stakeholders impacted?</li> </ul>	

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 questions like:

- Which program initiatives are providing the greatest impact?
- Are there elements of program delivery that are more or less effective than others in generating desired outcomes?
- Is greater impact achieved when specific strategies are combined into a package of initiatives?
- What are the contextual settings in which mechanisms are triggered to achieve desired outcomes?

These questions will be particularly relevant for the evaluation of whole-of-Government programs, where multiple agencies are working towards collective objective/s.

#### **Step 4: Choose appropriate evaluation methods**

After deciding on the objectives of the evaluation and the key questions that the evaluation will need to address, the evaluator will need to consider:

- **Methods:** Which methods are going to be used to answer the evaluation questions? The choice of methods will depend on what information is needed, the program characteristics, and organisational capacity and resources. The different types of methods that can be used are explored further in [Attachment A: Evaluation Methods](#).
- **Data collection and analysis:** What information is required and from whom and how the information can best be obtained? There are many different methods that can be used to source or collect data, and as such, the evaluation plan should outline what data is required and how it is going to be sourced. Wherever possible, evaluations should incorporate qualitative and quantitative data collection methods. The plan should also identify whether there are any cultural or ethical considerations or privacy concerns that need to be taken into account in the collection and use of data, and strategies to deal with any limitations or deficiencies in data collection methods. Different methods for collecting and analysing data, and the issues that need to be

considered when selecting an appropriate method, are explored further in [Attachment B: Collecting Evaluation Data](#).

- **Reporting:** How the information is going to be used? Analysis and reporting requirements will depend on the objectives of the evaluation and the intended audience. Consideration should be given to the extent of analysis required to allow for the development of valid findings, and the best mechanism for communicating findings (such as the format, language and structure of reporting).

While the specific elements of evaluation design may be refined during the evaluation, these issues should be considered while the program is being planned, as they will impact on resourcing, timing and consultation requirements.

### Step 5: Specify criteria for measuring success

The evaluation plan should also specify explicit criteria for determining the success of a program, to ensure that objective conclusions can be drawn. Appropriate criteria may depend on the questions being asked, the methods chosen for the evaluation and the objectives of the program.

The criteria could overlap with performance measures or key performance indicators (such as the percentage of people changing behaviour as a result of the program), or could be specific to the evaluation method (such as a positive benefit cost ratio for a cost benefit analysis).

**Table 1: Success criteria**

Evaluation Questions	Criteria requirement	Example of success criteria
Design, implementation and delivery	Measure the extent to which the program delivered its activities and outputs in line with the plan for implementation	<ul style="list-style-type: none"> <li>• Number of persons assisted was within a specified percentage of a target.</li> </ul>
Effectiveness	Measure the quantifiable extent of the effect of the program (the outcomes achieved), as a result of program activities and outputs.	<ul style="list-style-type: none"> <li>• At least a specified percentage of participants changed behaviour as a result of the program.</li> <li>• The program resulted in a statistically significant improvement in behaviour.</li> </ul>
Efficiency	Measure how resources are used to produce outputs for the purposes of achieving desired outcomes.	<ul style="list-style-type: none"> <li>• The program has a positive cost benefit ratio.</li> <li>• The cost of a unit of activity is in line with, or lower than the national average or some specified benchmark.</li> </ul>

For more information on performance measures refer to the Queensland Government [Performance Management Framework](#) (PMF), particularly *A Guide to the Queensland Government Performance Management Framework* and *Better Practice Guide for Developing Performance Information – useful information and examples*.

### Step 6: Consider who should manage and conduct the evaluation

The evaluation team could be internal to the program, external, or a combination of both. The appropriate expertise and size of the evaluation team will depend on the objectives of the evaluation, the evaluation design and the scale of the evaluation required.

If the evaluation is to be outsourced, the terms of reference must cover:

- The rationale for the evaluation
- The key evaluation questions that need to be investigated
- The type of evaluation required (i.e. process evaluation, cost-benefit analysis, etc.)
- The scope of the evaluation
- The key stakeholders
- Milestones, deliverables and modes for reporting
- Expectations for dissemination, disclosure and use of results.

### Step 7: Conduct a risk assessment and develop mitigation strategies

Consider any potential risks to the effectiveness of the evaluation and identify strategies to prevent or minimise the risk from occurring.

The development of contingency plans will improve the likelihood of evaluation managers being able to quickly implement appropriate responses to emerging problems and ensure the evaluation stays on track to meet its objectives.

**Table 2: Risks and risk management strategies**

Potential risks	Possible risk management strategies
Evaluation support	<ul style="list-style-type: none"> <li>• Ensure that the need for evaluation is clearly communicated to senior officers.</li> <li>• Ensure that the findings from the evaluation will be both valid and timed to inform decision-making.</li> <li>• Maintain regular contact with Government stakeholders to ensure their needs are being met.</li> </ul>
Timing of activities	<ul style="list-style-type: none"> <li>• Regular monitoring of evaluation plan milestones by evaluators.</li> <li>• Allocate sufficient time between evaluation milestones to allow for a small amount of project creep.</li> <li>• Ensure stakeholders are informed of the time involved in carrying out different evaluation approaches.</li> </ul>
Reliable data	<ul style="list-style-type: none"> <li>• Identify and clearly communicate data requirements at the earliest possible stage during program planning to enable collection of data during program implementation.</li> <li>• Attempt to access similar data from an alternative source.</li> <li>• Modify research design if it is not possible to source appropriate data.</li> </ul>
Funding availability	<ul style="list-style-type: none"> <li>• Include evaluation budget as a separate item in the program budget.</li> <li>• Ensure funding proposals accord with established Government processes.</li> <li>• Ensure that the need for the evaluation has been appropriately communicated, and that findings from the evaluation can be used to inform decision-making.</li> </ul>

For more information on risk management, see the Queensland Government's [A Guide to Risk Management](#), which outlines the minimum principles and procedures of a basic risk management process.

## 1.2 Conducting the evaluation

It is important to note that, even when an evaluation is meticulously planned, not all evaluation processes will go smoothly. There are a number of things that an evaluator can do to help ensure the evaluation is implemented successfully:

- Ensure there is clear and regular communication between the program evaluator, program manager, and key stakeholders throughout the entire evaluation process.
- Identify early whether there are any conflicting stakeholder interests, and determine strategies to minimise tensions between stakeholders about the findings of the evaluation.
- Ensure there is flexibility in program implementation, and a co-ordination between program and evaluation objectives.

If the evaluation is not progressing smoothly, it is important to identify why, and to set it right as soon as possible. This demands good project management and strategic skills on behalf of the program manager and the evaluator, and a willingness to call a halt to the evaluation and determine what is wrong.

If the evaluation occurs over a staged process, rather than as a one-off exercise, ongoing risk management should occur. Risks need to be actively reviewed and managed throughout the entire evaluation process, to take account of changing circumstances that may impact on the success of the evaluation. This reiterates the need for a risk management strategies established as part of the evaluation plan.

## 1.3 Using the results

The findings of the evaluation should be used for decision-making, accountability and to improve existing programs. Decisions about improvements can be made at any time during the evaluation for continuous improvement, or they can be used to improve the way that future programs are designed. Evaluation findings can also be used to inform decisions about the way the program is managed.

Different audiences will have different expectations when it comes to being advised about the findings of an evaluation. As such, the approach for reporting on and disseminating the findings from an evaluation will need to be adapted to suit the audience.

For many evaluations, an evaluation report will be the main means of communication with stakeholders and decision makers about the program. A formal evaluation report should include the following<sup>6</sup>:

- A summary, and a list of findings, judgements and/or recommendations
- A brief description of the evaluation objectives, method, participants and limitations
- A brief description of the program background, description, management, participants, objectives and method
- A section on evaluation findings (including results and their sources)
- A section on the conclusions drawn from the evaluation
- A summary section, describing what was learnt from the evaluation, and who should know about this information (this could be an executive summary).

---

<sup>6</sup> Department of Environment and Conservation (NSW), 2004, Does Your Project Make a Difference? A guide to evaluating environmental education projects and programs, available at <http://www.environment.nsw.gov.au/resources/communities/040110-Project-Evaluation.pdf>

For some audiences, effective communication of learnings may require communication of findings by more contemporary means. This could include presentations and information sessions, or making use of video, audio and written stories disseminated via a wide range of media.

**Tip: It's okay to find that the program requires improvement**

A well designed evaluation may find that the program is not meeting its objectives. This should not be viewed as a problem of evaluation, but rather, a chance to find out what worked and what did not, and respond accordingly.

## 2 PRACTICAL CONSIDERATIONS

There are a number of practical considerations that need to be taken into account when planning and undertaking an evaluation. These include the need for:

- Clear and explicit clarification of roles and responsibilities
- Explicit timing requirements for an evaluation, and the impact that this will have on resourcing
- Reflection on the success of an evaluation in meeting objectives and the potential for future improvement
- A strategic approach when determining appropriate resourcing for an evaluation.

### 2.1 Clarifying roles and responsibilities

For any evaluation, there needs to be clearly defined roles and responsibilities. The roles required for an evaluation will vary depending on the nature of the program and the evaluation requirements.

Where the evaluation manager is separate to the program manager, the two parties will need to work closely with one another to ensure that the program planning and implementation supports effective evaluation, and that the evaluation findings are timed and relevant to be able to influence program decision making.

To enable this to occur, the selection of an evaluation manager should occur during program development, to allow the evaluation manager to work with the program manager to ensure that program planning and evaluation planning can occur simultaneously.

Where the program involves multiple delivery partners, consideration should be given to formalising role and responsibilities through MOUs, agreements, contracts or committees, which set out<sup>7</sup>:

- The objectives of the arrangement, including desired outcomes and timeframes
- The roles and responsibilities of the agencies, including their capacity to contribute
- The details of the activity, including specifications of services or projects to be undertaken
- Resources to be applied by the agencies and related budgetary issues
- The approach to identifying and sharing the risks and opportunities involved
- Agreed dispute resolution arrangements.

**Table 3: Roles and responsibilities in program evaluation**

Role	Potential Responsibilities
Funding providers	<ul style="list-style-type: none"><li>• Ensure there is sufficient rationale for program development and implementation.</li><li>• Encourage program developers to give consideration to evaluation requirements during program design and provide support to encourage and enable the effective evaluation of programs.</li><li>• Over time, ensure that there is sufficient evidence available to support understanding of a program's effectiveness and efficiency.</li></ul>
Program developer	<ul style="list-style-type: none"><li>• Ensure the program has appropriate and clearly defined objectives, targets</li></ul>

<sup>7</sup> ANAO *Better Practice Guide – Implementation of Programme and Policy Initiatives: Making Implementation Matter*, October 2006, available online at [http://www.anao.gov.au/uploads/documents/Implementation\\_of\\_Programme\\_and\\_Policy\\_Initiatives.pdf](http://www.anao.gov.au/uploads/documents/Implementation_of_Programme_and_Policy_Initiatives.pdf)

/ program manager	<p>and performance criteria.</p> <ul style="list-style-type: none"> <li>• Appoint an evaluation manager.</li> <li>• Ensure adequate data is collected to support any planned evaluation.</li> <li>• Work with the evaluation manager to ensure that the evaluation is timed to enable the incorporation of evaluation findings into program decisions and that the evaluation will produce findings relevant to decision-making.</li> </ul>
Evaluation manager	<ul style="list-style-type: none"> <li>• Preparation of terms of reference for the evaluation, including clarifying objectives, scope and key stakeholders.</li> <li>• Work with the program manager to ensure the evaluation is timed to enable the incorporation of evaluation findings into program decisions and that the evaluation will produce findings relevant to decision-making.</li> <li>• Ensure stakeholder involvement throughout the evaluation process.</li> <li>• Ensuring evaluation stays on track, meets its objectives, is on time and is delivered within budget.</li> <li>• Quality assurance.</li> <li>• Dissemination of evaluation findings.</li> </ul>
Evaluator	<ul style="list-style-type: none"> <li>• Conduct the evaluation.</li> <li>• Preparation of evaluation reports.</li> </ul>

On a broader scale, agencies need to ensure that there is appropriate support and resources available to enable effective evaluations of public sector programs. It is ultimately the responsibility of Government decisions makers and senior officers in portfolio and central agencies to encourage a culture that articulates and supports the need for program evaluations.

**2.2 Evaluation timeframes**

The time taken to undertake an evaluation will vary depending on the type of evaluation approach, the program’s complexity, and the strength of evidence required for the evaluation.

A simple process evaluation could take only a number of days. Complex program evaluations could take a number of years, depending on the time required to detect changes in short, medium and long-run outcomes.

For example, the process of designing a prospective impact evaluation and collecting a baseline from scratch can often take a year or more. Then, once the program starts, the intervention needs a sufficient exposure period to affect outcomes. Depending on the program, that can take anywhere from a year to five years, or more. Collecting one or more follow-up surveys, conducting the analysis, and disseminating evaluation findings will also involve substantial effort over a number of months. Altogether, a complete impact evaluation cycle from start to finish can take three to four years of intensive work and engagement, with adequate financial and technical resources required at each step of the way<sup>8</sup>.

It is important that stakeholders are realistic about the costs, complexity and time involved in carrying out different evaluation approaches.

For evaluations that extend over a longer time, it is important to have pre-determined reporting points built into the evaluation plan. These points provide an opportunity to disseminate information on how

<sup>8</sup> Gertler P.J, Martinez S, Premand P, Rawlings L.B, Vermeersch C.M.J (2011). Impact Evaluation in Practice. The World Bank, Washington DC.

the evaluation is tracking and whether there are any interim findings that can assist in improving either the program or the evaluation (such as improvements to program delivery or data collection).

### 2.3 Refining the evaluation process

Just as evaluations can help to inform continual improvement of program and outcome delivery, evaluation practices should also be subject to continual review and improvement over time.

Evaluators should reflect on whether the evaluations being undertaken are achieving their objectives, and if not, why. This involves critically analysing the evaluation process on an ongoing basis, to consider elements such as:

- **Planning** – Has an appropriate evaluation plan been developed that includes clear consideration of objectives, design, methods, outputs, governance arrangements and risk management strategies?
- **Competency** – Do the personnel assigned to the evaluation have sufficient capability (in terms of knowledge, skills, abilities and experience) to undertake the tasks required?
- **Evaluation scope and design** – Is the evaluation scope and design appropriate given the program and evaluation objectives and the resourcing and timing constraints placed on the evaluation?
- **Data and analysis** – Was the data collected sufficiently reliable for the intended use? If not, what could improve data collection in the future? Have data and processing techniques been reviewed for accuracy and reliability? Have weaknesses in data collection or analysis been identified and, where possible, corrected?
- **Findings and conclusions** – Is there sufficient supporting evidence available to underpin the evaluations' findings? Are the conclusions and recommendations supported by the findings? Have the assumptions and limitations of the evaluation been made transparent?
- **Quality assurance** – Were appropriate internal controls developed, and adhered to, in order to ensure the accuracy, reliability and applicability of findings?

A useful strategy in ensuing continual improvement in evaluation practices can be to subject the evaluation process and findings to a peer review, undertaken by a qualified individual (or team) not involved in the evaluation. The purpose of undertaking a peer review should be to determine if the evaluators made the most of the available data and used appropriate methods (given external factors, unforeseen events and constraints placed on the evaluation), and to identify actions that could be taken to improve future evaluations.

### 2.4 Scale of the evaluation

A strategic approach should be adopted when deciding on the scope, governance arrangements and resource requirements for an evaluation. As the scale, complexity and risk associated with a program increases, the evaluation of the program will need to become more extensive.

In general, the level of resources devoted to the evaluation should be proportional to the level of resources devoted to the program and the magnitude of the expected impacts. More resources should be committed to evaluations where the program is expensive, complex, large-scale or high risk. On the other hand, fewer resources will be allocated for an evaluation of a program that is low spend and low impact. Resourcing of evaluations should ensure that evaluations are done to an acceptable standard and that the evaluation represents value for money.

There are cases where a high level of resources might be required for the evaluation even though total expenditure on the program could be considered low. These include:

- If the findings from the evaluation will be used to inform decisions on whether to roll the program out to a wider area/group (such as with a pilot or trial).
- If the findings from the evaluation are to be generalised (used as evidence of other programs effectiveness).

In general, the programs characteristics and scale are important factors to consider in decisions on:

- Evaluation objectives and design
- Evaluator appointment
- Stakeholder consultation requirements.

**Table 4: Scaling evaluation**

Program characteristics	Evaluation design	Evaluator	Stakeholder consultation
<ul style="list-style-type: none"> <li>• <b>Low</b> risk</li> <li>• <b>Simple</b> program design</li> <li>• <b>Low</b> resource requirements</li> <li>• <b>Ongoing</b> program</li> <li>• <b>Single</b> delivery body</li> <li>• <b>Low</b> potential for behavioural impacts</li> </ul>	<p>Qualitative assessment of implementation success and program efficiency built into program reporting could be sufficient.</p> <p>Additional assessment should be undertaken where it is cost-effective to do so.</p>	<p>The evaluation can be conducted by an internal evaluator, seeking advice and assistance from experts if required.</p>	<p>Consultation may simply involve <b>project managers</b>, but could also include individuals or institutions directly impacted by the program.</p>
<ul style="list-style-type: none"> <li>• <b>High</b> risk</li> <li>• <b>Complex</b> program design</li> <li>• <b>High</b> resource requirements</li> <li>• <b>Pilot</b> or trial program</li> <li>• <b>Multiple</b> delivery bodies</li> <li>• <b>High</b> potential for behavioural impacts</li> </ul>	<p>Comprehensive evaluation design that assesses the programs’ implementation, efficiency and effectiveness from a whole of society perspective.</p>	<p>The evaluation should be managed by an external party independent of the agencies involved in program delivery.</p>	<p>Extensive consultation with both the stakeholders listed above as well as:</p> <ul style="list-style-type: none"> <li>• <b>Evaluation experts</b> on the appropriateness of evaluation design, methods and assumptions</li> <li>• <b>Central agencies</b> on the fit-for-purpose nature of the evaluation approach and potential findings.</li> </ul>

The interdependence of evaluation and program characteristics reinforces the need for developing an evaluation plan while the program is being designed, rather than after the program has been implemented, so that the program budget can take into account the resources required for an effective evaluation.

### 3 WHERE TO GO FOR FURTHER INFORMATION

#### General evaluation resources

The following organisation or groups provide a range of resources or services relating to various aspects of program evaluation:

#### [BetterEvaluation.org](http://BetterEvaluation.org)

Provides extensive discussion and resources on evaluation planning and approaches, including user contributed examples and comments.

#### [NSW Evaluation Toolkit](http://NSWEvaluationToolkit.gov.au)

The New South Wales Government has developed an online resource that provides advice and tools for planning and conducting a program evaluation.

#### [The World Bank Independent Evaluation Group](http://www.independentevaluation.org)

The IEG website provides a range of evaluation resources, including “how-to” guides, evaluation reports, and case studies.

#### [Queensland Government Statistician’s Office](http://www.qstat.gov.au)

The Queensland Government Statistician’s Office provides a broad range of statistical services to support stakeholders’ evidence base for evaluation and performance.

#### [Australasian Evaluation Society](http://www.aesoc.org)

A member based organisation which exists to improve the theory, practice and use of evaluation. Publically information includes access to journal articles and evaluation reports.

#### Resources on the evaluation process

Some useful publications include:

- HM Treasury’s [The Magenta Book: Guidance for Evaluation](http://www.hm-treasury.gov.uk/magenta_book.htm)
- Horizon Research Incorporated’s [Taking Stock: A Practical Guide to Evaluating Your Own Programs](http://www.horizonresearch.com.au/taking-stock)
- United States Government Accountability Office’s [Designing Evaluations](http://www.gao.gov/products/2005/05-102)

## ATTACHMENT A: EVALUATION APPROACHES

In broad terms, evaluation approaches can be grouped into three broad categories:

- Evaluating design, implementation and delivery
- Evaluating effectiveness in achieving outcomes
- Evaluating efficiency in delivering outcomes.

Within these three categories, there are many different methods that can be used to collect and analyse data and inform findings. The ideal choice will depend on available data, timing and resourcing constraints, and how robust the findings need to be to inform decision making.

The sections below outline some of the main methods that can be used for program evaluation and discusses some of the key things to consider when selecting a preferred approach.

### **Tip: Seek expert advice if you're not sure**

Knowing which evaluation methods and techniques are suitable for any given evaluation can require expert knowledge of the requirements and caveats of different approaches. The advice below is provided at a high level as a “place to start” when thinking about evaluation methods, rather than intended to outline all issues you might need to consider.

### **A.1 Evaluating design, implementation and delivery**

An **implementation** or **process** evaluation focuses on how a program was implemented and whether it is providing services as intended (focussing on output delivery rather than on outcomes).

A process evaluation attempts to answer the following types of questions<sup>9</sup>:

- Is the program being implemented as intended?
- Are the activities being delivered as intended?
- Are participants being reached as intended?

Information obtained through a process evaluation can help to shed light on not just what is happening, but why. Process evaluations can assist in identifying intended or unintended implementation characteristics that may have affected program delivery and impacted on program outcomes (positively or negatively).

A process evaluation might assess the adequacy of administrative process, management practices and service delivery mechanisms, as well as the acceptability of services to program stakeholders<sup>10</sup>.

A process evaluation can occur at any stage of program delivery and potentially on multiple occasions (particularly for ongoing programs), and can include both qualitative and quantitative methods. Appropriate methods will depend on the purpose of the evaluation and the level of evidence required.

---

<sup>9</sup> NSW Government Evaluation Framework, August 2013, available at [http://www.dpc.nsw.gov.au/\\_data/assets/pdf\\_file/0009/155844/NSW\\_Government\\_Evaluation\\_Framework\\_Aug\\_2013.pdf](http://www.dpc.nsw.gov.au/_data/assets/pdf_file/0009/155844/NSW_Government_Evaluation_Framework_Aug_2013.pdf)

<sup>10</sup> Gertler P.J, Martinez S, Premand P, Rawlings L.B, Vermeersch C.M.J, 2011, Impact Evaluation in Practice, The World Bank, available online at [http://siteresources.worldbank.org/EXTHDOFFICE/Resources/5485726-1295455628620/Impact\\_Evaluation\\_in\\_Practice.pdf](http://siteresources.worldbank.org/EXTHDOFFICE/Resources/5485726-1295455628620/Impact_Evaluation_in_Practice.pdf)

**Table 5: Overview of process evaluation**

Evaluation questions	Possible evidence required	Possible methods or data sources
How well has the program been established?	<ul style="list-style-type: none"> <li>• Description of program development compared with client needs, timeframes etc.</li> <li>• Quality of governance, relationships.</li> <li>• Influence of different factors and contexts.</li> <li>• Initial evidence of uptake.</li> </ul>	<ul style="list-style-type: none"> <li>• Program reports, key informant interviews, consultations with managers or service providers, program forums.</li> </ul>
How is the program being implemented?	<ul style="list-style-type: none"> <li>• Description of implementation processes by different providers and in different circumstances.</li> <li>• The extent that implementation processes met milestones and targets for outputs, timeliness, cost, participation and immediate outcomes.</li> <li>• The quality of outputs and immediate outcomes measured against standards and targets.</li> <li>• The pattern of outputs, uptake and immediate outcomes, by different sub-groups or in different contexts.</li> <li>• Client or customer satisfaction.</li> </ul>	<ul style="list-style-type: none"> <li>• Program monitoring data and other program records.</li> <li>• Observation including photography and video.</li> <li>• Interviews, surveys or focus groups with managers, staff, program clients, referring agencies.</li> <li>• Consultations with managers or service providers.</li> </ul>
Is the program being implemented well?	<ul style="list-style-type: none"> <li>• As above plus information about good practice in implementation processes.</li> </ul>	<ul style="list-style-type: none"> <li>• Expert review of program documents, or observations during site visits.</li> </ul>

Source: Department of Premier and Cabinet (NSW), 2014, Policy Makers Toolkit

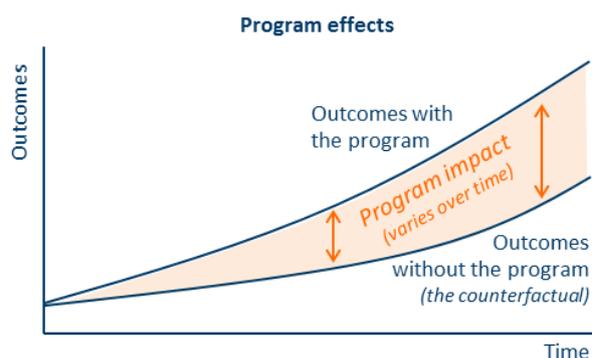
While there are many different factors that will inform whether a program has been implemented successfully, in general a successful program will have:

- Been implemented on schedule and on budget
- Delivered the agreed outputs, at an acceptable quality, to the agreed audience
- Met the requirements of stakeholders.

## A.2 Evaluating effectiveness

Assessing a program’s effectiveness requires consideration of what would have happened in the absence of the program (the counterfactual) or what has happened after program implementation in comparison to what has occurred prior to implementation.

Most observed outcomes are affected by a range of factors, not just the program in question. As such, a rigorous evaluation of program effectiveness will make a concerted attempt to isolate the effect of the program from all other



potential influences<sup>11</sup>. For example, an evaluation of a program aimed at getting the unemployed back to work would need to consider if the participants finding employment was a result of the program or from other factors, such as changing labour market conditions.

An **impact** or **outcome** evaluation focusses on the outcomes of a program and attempts to ascertain whether a program was successful in achieving its intended impact (focussing on outcomes rather than outputs).

An impact evaluation should ask:

- Did the program achieve its stated objectives?
- How much of the changes could be attributed to the program as opposed to other factors?
- What was the magnitude of the impact achieved by the program?
- Were there any unintended consequences, and if so, how significant were they?

Impact evaluations may be carried out using a variety of methods, which can be broadly classified into three categories, depending on the level of control over variables and the ability to determine a causal relationship between a program's activities and observed outcomes:

- **Experimental design** – based around the selection of a comparison and treatment group, through random assignment, prior to the implementation of a program. The treatment group is then exposed to the program, while the comparison group is not. A randomised-control trial is an example of an experimental design. Experimental design requires a very high level of control which can be very difficult to achieve outside clinical studies.
- **Quasi-experimental design** – consists of the evaluator selecting a comparison group, that is matched with the treatment group on certain characteristics (either before or after a program has been implemented) to provide a comparator for the treatment group. Quasi-experimental design offers less control than experimental design, in that subjects are selected from pre-existing groups; e.g. criminal offenders or people with a disability. Difference-in-difference, regression discontinuity and matching samples are examples of quasi-experimental designs.
- **Non-experimental design** – will not demonstrate a cause-and-effect relationship, but instead seek to understand how and why a program is effective (not just whether it is effective). Non-experimental designs include a very wide range of methods, including realist, theory-driven and participatory approaches, which rely on interpretation, observation or interactions to come to a conclusion.

The most appropriate method will be determined by the resources and time available for evaluation, the feasibility of implementation and the strength of evidence required from the evaluation findings.

Some **key considerations** when selecting an appropriate method for analysing impact include:

- **Determining causality** – There are many factors that can compromise the ability of an evaluation to demonstrate that a relationship exists between a program's actions and the observed outcomes. These factors are known as internal threats to validity, and include history, maturation and design contamination. Implemented appropriately, experimental and quasi-experimental designs, measuring prior to and throughout program implementation, will provide the best means for controlling these factors.

---

<sup>11</sup> HM Treasury, 2011, The Magenta Book: Guidance for evaluation, available online at <https://www.gov.uk/government/publications/the-magenta-book>

- **Generalisability of results** – If the evaluation results are going to be used to inform decision around the likely effect of the program on different individuals, contexts or outcomes, the evaluation will need to control for factors that might reduce the generalisability of the evaluation findings. These factors are also known as external threats to validity.
- **Sample size** – In general, the larger the sample size, the easier it will be to separate out the impacts of a program from other random variations. It may be easier, or more cost effective, to have larger sample sizes for some methods over others.
- **Cost** – Different evaluation methods will require different time, resourcing and levels of expertise. Evaluations with controlled or quasi-experimental design will almost always cost more than non-experimental evaluations. However, the additional cost can sometimes be quite low depending on the type and availability of data to be collected. Moreover, findings from a more extensive evaluation may lead to future cost-savings, through improved outcomes and more efficient use of resources. Nevertheless, program managers must anticipate these additional costs when budgeting for program implementation<sup>12</sup>. This reinforces the need for the evaluation plan to be completed during the design phase of the program, and for the program budget to separately identify evaluation costs.
- **Type of information provided** – Experimental and quasi-experimental designs explicitly seek to demonstrate **cause and effect** relationships (i.e. evidence of whether a program is effective in delivering outcomes). While non-experimental or observational designs are less likely to be able to demonstrate the causal link between program actions and outcomes, they can often add value to an evaluation through their ability to inform an understanding of **why** events or outcomes unfolded in a particular way<sup>13</sup>.
- **Complexity of the program** – For programs with multiple delivery strategies, the evaluation may need to examine whether individual elements of a program are effective, and how they contribute to the effectiveness of the program as a whole. A realist-driven approach to program evaluation may be required in such circumstances. Realist evaluation seeks to understand which initiatives appear to be effective, for whom and under what circumstances, recognising that there can be multiple pathways to any given outcome. Such an approach seeks to link the mechanisms triggered by a context to produce a particular outcome.

**Tip: Build a body of evidence**

It is highly unlikely that one evaluation method or technique will be able to isolate the impacts of a program from all other effects, or to provide a complete picture of program impacts. Where possible, evidence collected from a range of evaluations should be pulled together into a body of evidence to form a comprehensive view of program impacts.

<sup>12</sup> USAID, 2010, Performance Monitoring and Evaluation Tips: Rigorous Impact Evaluation, available online at [http://pdf.usaid.gov/pdf\\_docs/pnadw119.pdf](http://pdf.usaid.gov/pdf_docs/pnadw119.pdf)

<sup>13</sup> Department for International Development, 2014, Assessing the Strength of Evidence, available online at [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/291982/HTN-strength-evidence-march2014.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/291982/HTN-strength-evidence-march2014.pdf)

**Table 6: Overview of impact evaluation methods**

Evaluation design	Design features	Considerations
<p><b>Experimental</b> (such as a randomised-control trial)</p>	<p>Involves gathering a set of equally eligible individuals willing to participate in the program, randomly assigning them to either a treatment group (those affected by the program) or a control group (unaffected by the program).</p>	<p>Attempts to carefully control confounding factors<sup>14</sup> in the design and hence improve the ability to attribute changes to the intervention itself.</p> <p>Can be resource intensive and difficult (or unethical) to provide an intervention to one group and not another.</p> <p>Can also be very difficult to achieve the required level of control outside of clinical settings.</p>
<p><b>Quasi-experimental</b> (such as difference-in-difference, regression discontinuity and matching)</p>	<p>Consists of constructing a comparison group using matching or reflexive comparisons either before or after program implementation (prospective or retrospective evaluation).</p>	<p>Can allow evaluators to draw on existing data sources, making this method quicker and cheaper to implement.</p> <p>The matching methods can be statistically complex, and the methodology may not completely solve the problem of selection bias, which can reduce the reliability of the results.</p> <p>Makes some attempt at controlling confounding factors in the design to improve the ability to isolate intervention effects.</p>
<p><b>Non-experimental</b> (methods that rely on interpretation, observation or interactions such as realist, theory-driven and participatory approaches)</p>	<p>Includes a very wide range of different types of evaluations and other studies that do not compare results for the same group over time or different groups of participants.</p>	<p>The evaluation design can be relatively cheap and quick to implement.</p> <p>A number of difficulties in ensuring the validity of findings mean this design tends to provide weak evidence about the effectiveness of a program.</p>

### A.3 Evaluating efficiency

A reliable impact evaluation might be able to demonstrate and quantify the outcomes generated by a program, but will not provide information on the extent to which the program is efficient and provides value for money.

The concept of efficiency has a number of dimensions: technical, allocative and dynamic efficiency. For the purposes of program evaluation:

- technical efficiency requires that the program is delivered at the lowest possible cost
- allocative efficiency requires that the program is delivered to the areas of greatest need, and provides the types of outputs or outcomes that recipients value most (for the given set of resources available)

<sup>14</sup> Confounding factors are any factors (or variables) other than the program which can affect the outcome of interest.

- dynamic efficiency means that, over time, the program continues to improve, by finding better or lower cost ways to deliver outcomes.

The extent to which it will be possible to examine all three types of efficiency in a quantitative framework will depend on the scope of the evaluation and the type of program being delivered. Where quantitative assessment of a particular type of efficiency is not possible, the efficiency aspects can be addressed through a qualitative discussion.

### **Program specific scope**

When designing a program, the program manager should give consideration to whether there are ways to build in evaluation of a program's efficiency in delivering outcomes.

At a basic level, examination of a program's efficiency should consider:

- Whether changes to processes could result in more efficient service delivery.
- Whether there are elements of the program that are currently delivered by Government that could be delivered more effectively by the private sector (refer to the Queensland Government Contestability Framework).
- Whether there are different solutions (such as alternate technologies) that could deliver the same outcomes more efficiently or effectively, or improved outcomes for the same cost.
- Whether there are similar programs being delivered by other agencies or bodies, with potential to more strategically target efforts or pool resources.
- Whether the program is being delivered to the areas of greatest need.
- Whether the program is delivering the types of outputs desired by program recipients (such as phone or internet options for payments, as opposed to cheque or money orders).

An examination of efficiency should also involve consideration of whether there are any factors impeding the achievement of more efficient service delivery, and investigating ways these could be overcome (to ensure the program can remain efficient over time).

If a more cost-effective solution is identified and approved for implementation, the program manager would need to ensure that the implementation of the alternate solution was successful and that the effectiveness of the program in achieving outcomes was not unduly compromised (which may instigate the need for a revised evaluation plan).

### **Wider (whole-of-economy) scope**

For programs that are large, complex, expensive, associated with a high degree of risk or with the potential for large behavioural impacts, it may also be necessary to consider whether the benefits of the program outweigh the costs.

**Economic evaluation** takes a wider scope than the program specific considerations discussed above, aiming to address the broader question of whether the program is the most appropriate way of addressing the problem which triggered the introduction of the program. Economic evaluation considers the financial, environmental and social costs and benefits generated by the program against other delivery options to determine if the program provides the best solution from a whole-of-society perspective.

For example, a program manager who has been asked to deliver a treatment program could evaluate whether the program is cost-efficient and effective, while Government more broadly should evaluate

whether a treatment program is the best option, or if a preventative program may provide greater benefit or value for money.

Economic evaluation can be used to inform decisions on whether:

- A existing program should be continued or ceased
- The current mix of programs is efficient and provides value for money.

An economic evaluation involves the process of systematically identifying, valuing and comparing the costs and consequences of two or more alternative programs or interventions in a unified framework. Where appropriate, it is desirable to quantify the economic impacts of the program to the extent possible, while this information can then be combined with consideration of broader issues such as affordability, equity, acceptability, feasibility and the strength of the evidence base<sup>15</sup> to determine whether a program should be implemented, continued, or ceased.

There are three main types of economic evaluation that assess the economic efficiency of programs: cost-benefit analysis, cost-effectiveness analysis and multi-criteria analysis.

**Table 7: Overview of economic evaluation methods**

Evaluation design	Design features	Assessing method suitability
Cost-benefit analysis	<p>Compares the financial, economic, environmental and social costs of a program from the perspective of society as a whole.</p> <p>Involves the comparison of one or more program scenarios with an alternative scenario based on estimates of what would have happened in the absence of the program.</p> <p>Requires the monetisation of both costs and benefits to the extent possible.</p>	<p>The most commonly accepted method for assessing the economic efficiency of a program.</p> <p>Quantification of the full range of costs and benefits can be resource intensive.</p> <p>Can ignore equity concerns, although such measures can also be assessed or discussed in qualitative terms as part of a cost-benefit analysis.</p>
Cost-effectiveness analysis	<p>Takes a similar approach to cost-benefit analysis, but compares the costs of an intervention (in dollars) to a common physical unit of outcome (such as “number of lives saved”).</p> <p>Requires the same outcome measure to be relevant for all programs being compared.</p>	<p>Can only be used in situations where programs have shared goals; when the primary objective of the evaluation is to identify the most cost-effective strategy from a group of alternatives that can effectively meet a common goal<sup>16</sup>.</p>
Multi-criteria analysis	<p>Includes a broad range of techniques that can be used to assess alternative options according to a variety of criteria that have different units (e.g. dollars, tonnes, and kilometres).</p> <p>Involves assigning weights and scores to each criteria, to indicate the relative</p>	<p>The method allows for both qualitative and quantitative criteria to be considered, however the subjectivity of weighting and scoring means there is a high likelihood of bias.</p>

<sup>15</sup> Vos T, Carter R, Barendregt J, Mihalopoulos C, Veerman JL, Magnus A, Cobiac L, Bertram MY, Wallace AL, ACE-Prevention Team, 2010, Assessing Cost-Effectiveness in Prevention (ACE-Prevention): Final Report, University of Queensland and Deakin University, available online at <http://www.sph.uq.edu.au/bodce-ace-prevention>

<sup>16</sup> CDC, Cost Effectiveness Analysis Tutorial, available online at [http://www.cdc.gov/pcd/issues/2012/11\\_0324.htm](http://www.cdc.gov/pcd/issues/2012/11_0324.htm)

	importance.	
--	-------------	--

Regardless of the type of method selected, it is expected that an economic evaluation will include the following features:

- A broad view of costs and benefits; including environmental and social impacts and longer-term effects.
- Identification of all costs and benefits, even those that can't be monetised, including consideration of both indirect and non-market costs and benefits.
- Assessment of the impacts on different stakeholder groups (such as those directly impacted by the program, the Government, and the wider community).
- Sensitivity analysis: information on all costs and benefits is rarely known with certainty, necessitating the need to test the sensitivity of the results to changes in key parameters.

The findings of an economic evaluation will generally be heavily dependent on the assumptions chosen to underpin the analysis. As such, any evaluation report should clearly set out the assumptions chosen and explore the impacts that the assumptions have on the findings (the results of sensitivity analysis). Ensuring that all assumptions made during the evaluation process are transparent is essential for the explanation and use of the evaluation findings.

Economic impact analysis (EIA), while potentially complementary, is not a substitute for the economic evaluation methods outlined above (or for an impact evaluation). The method is only useful in the case of large scale reforms where an evaluation needs to understand wide scale economic implications. If EIA is used to complement an economic evaluation, then a robust methodology framework, such as Computable General Equilibrium (CGE) analysis should be used.

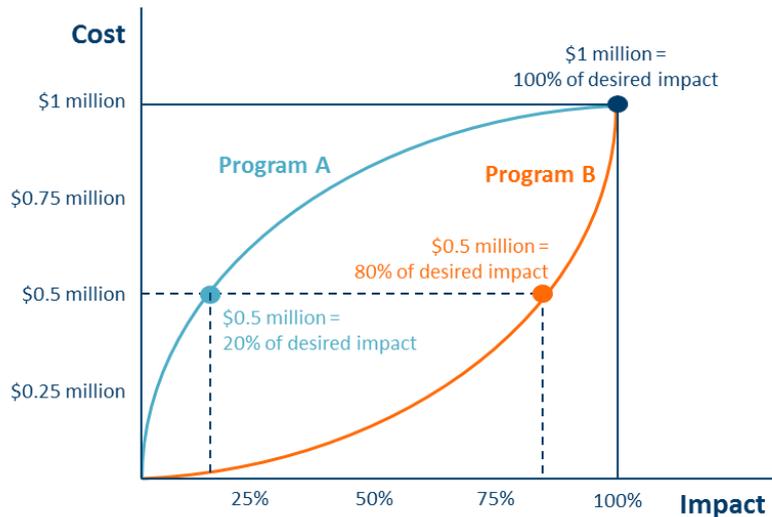
Input-output or “multiplier” analysis should not be used in any form of appraisal or evaluation. The methods rely on a range of simplistic assumptions and will produce biased estimates of the potential or realised costs or benefits of a program.

### **Incremental or marginal cost**

The default position for most evaluations will be to compare the **total costs** of delivering a program with the level of output or impact provided by the program, to determine whether program benefits outweigh costs.

However, considerations of **incremental or marginal costs** of programs can provide significantly more information for decision makers. Consider two programs; program A and program B. Both programs have been tasked with delivering a particular outcome. Upon examining the total costs of each program, both are shown to be equally cost-effective. However, program A only achieves the majority of desired impact once a significant portion of program expenditure has been outlaid, while program B achieves the majority of impact with lower levels of expenditure.

**Figure 5: Example of programs with different marginal costs**



Rather than the “yes” or “no” type information generated by considering total costs (or average costs) in isolation, information on incremental or marginal costs provides additional detail on the costs of expanding or reducing a programs reach or scope and the level of impact that can be expected to be achieved by doing this. This type of information can be particularly useful to help to inform decisions on program delivery in tight fiscal environments.

#### **A.4 Further resources**

Some useful publications include:

- The World Bank’s [Impact Evaluation in Practice](#) and [Evaluating the Impact of Development Projects on Poverty: A Handbook for Practitioners](#)
- The Asian Development Bank’s [Impact Evaluation: Methodological and Operational Issues](#) and [A Review of Recent Developments in Impact Evaluation](#)
- Commonwealth Department of Finance and Administration’s [Handbook of Cost-Benefit Analysis](#) and [Introduction to Cost-Benefit Analysis and Alternative Evaluation Methodologies](#)
- Victorian Department of Treasury and Finance’s [Economic Evaluation for Business Cases: Technical Guidelines](#)

## **ATTACHMENT B: COLLECTING EVALUATION DATA**

Data is the means by which you develop credible evidence that your program is successful, or that you have uncovered and are addressing limitations.

Failure to pay due attention to data requirements is one of the primary reasons for evaluation failure. Consequences from improperly collected data include an inability to generate valid findings, wasted resources and compromised decision making.

As such, a critical part of any evaluation is consideration as to what data will be required to support the chosen method, and how it will be collected and analysed.

The collection and analysis of data for the purpose of an evaluation, regardless of the type or method of collection, must take into consideration the sensitivity and confidentiality of the information being obtained. Information privacy and ethical issues should be considered when determining what data will be collected, and whether and how participants should be informed of how that data will be used<sup>17</sup>.

### **B.1 Types of data**

At the broadest level, data comes in two main forms: quantitative and qualitative. Quantitative data are measures of values or counts and are expressed as numbers, while qualitative data consist of virtually any information that can be captured (and often categorised) that is not numerical in nature.

Analysis of quantitative data has benefits in that it lends itself to various forms of statistical techniques, which can provide objective information and a solid foundation for description and analysis. Quantitative data analysis often aims to measure the prevalence of behaviours or opinions in the population of interest. Benefits can include the ability to measure changes in prevalence within the population. It is important, however, to ensure that the available quantitative data is relevant to the question of interest. The amount of quantitative data required should also be appropriate, as sometimes a smaller, well-targeted dataset can be of more value than a large general dataset. In addition, the methods chosen for analysing the data should be appropriate to the question of interest and to the type of data available.

Analysis of qualitative data has benefits in richness of detail. Findings are grounded in reality, and the method provides tolerance of ambiguity and contradictions, and allows for the prospect of alternative explanations or conclusions. That being said, qualitative information is often less generalisable than quantitative information, and analysis can often be time consuming or risk the possibility of decontextualizing or oversimplifying meanings or explanations.

The use of both qualitative and quantitative approaches within an evaluation can provide a more complete picture and improve the accuracy of findings (by compensating for relative strengths and weaknesses). For example, qualitative data may provide context and explanations to complement quantitative data. The trade-off comes in the resource cost involved in undertaking greater analysis, and in needing to amass additional expertise. There is also the chance that the use of different methods may result in information or findings that are conflicting, and evaluators will need to have strategies in place to help deal with this if it occurs.

---

<sup>17</sup> Information on the operation and application of Queensland's Right to Information and Information Privacy legislation is available from Queensland's [Office of the Information Commissioner](#).

## **B.2 Collection methods**

Some common methods for collecting qualitative and quantitative data for program evaluation are outlined below<sup>18</sup>.

### **Interviews**

An interview is a method of asking either qualitative or quantitative questions of key participants or proponents. Interviews can take many forms: structured, semi-structured or unstructured, one-on-one, or group. Interviews can be conducted in person or by telephone. The main advantages to collecting data through interviews are that the method is highly flexible and tends to have high response rates. Large scale structured telephone interviews using closed questions can be used where standardised data is required across the population of interest and are useful where prevalence measures are required. Where more open ended questions are asked, the method can provide a depth of information and contextual insight into issues that other, less personalised forms of data collection may not.

Interviews can be time-consuming (and therefore resource intensive). Unstructured interviews can also produce non-standard responses (which can make analysis challenging). Success of the method is also largely dependent on the skill of the interviewer and the context or setting of the interview, and in the case of structured interviews, the quality of the questions. The method is also susceptible to interview bias, which can result in people giving altered responses (what people say they do differing from what they actually do).

### **Self-completion surveys**

Self-completion surveys involve gathering written responses to questions from individuals. They can be administered by mail or over the internet and can produce both qualitative and quantitative data (through the use of open or closed questions). Self-completion surveys can be a relatively inexpensive way to obtain information from a large number of respondents in dispersed or geographically diverse locations.

Self-completion surveys are often associated with poor response rates, and carry greater potential for non-response bias compared to other methods of data collection. Depending on the type of information required and the intended audience, it can also be challenging and time-consuming to design a questionnaire that will ensure information collected is relevant and accurate. Respondents must be able to read and understand questions posed and questionnaires should be kept as straightforward as possible.

### **Administrative data**

The data to day collection of administrative data often leads to the generation of large amounts of information that are stored by program providers and government departments.

As these datasets already exist, it is often considered a relatively inexpensive way of obtaining data for analysis for the purposes of evaluation. For example, data on the number of participants in a training program may be recorded on a regular basis and may be an indicator of success. A potential benefit of administrative data is that it can cover time periods prior and post the implementation of an intervention.

---

<sup>18</sup> Information in this section has been adapted primarily from CDC Evaluation Briefs, available online at <http://www.cdc.gov/healthyouth/evaluation/data.htm>, and Denscombe, M, 2007, *The good research guide: for small-scale research projects*, 4th ed.

However, as the data has not generally been collected for research purposes, it is also common for important contextual or other details necessary for a valid analysis to be missing. The data may be inaccurate or inconsistent, causing difficulties in analysis.

At the start of an evaluation, it may be possible to negotiate for the collection of data in addition to the usual information collected for everyday needs, although the additional burden that this would place on frontline staff would need to be considered.

### **Observation**

Observation is a way of gathering data by watching behaviour or events, or noting physical characteristics, in their natural setting. Gathering information through observation can be a particularly useful form of data collection when data collection directly from individuals is not a realistic option. Observation can help if you want to become aware of aspects of a program that may not be consciously recognised by participants or staff or when what people say they do might be different from what they actually do.

This method is a relatively objective mechanism for collecting data, and can be efficient if observations can be collected in a short period of time. However, the method only provides information on what, rather than why certain behaviours, activities or events occurred. Observation can also be susceptible to observer bias (when people perform differently because they are under observation) and can present challenges in interpretation and categorisation of observed behaviours.

### **Case Studies**

Case studies involve the depiction of one or a few instances of a process or outcome with a view to providing an in-depth account of events, relationships, experiences or processes occurring in that particular instance. The use of case studies work best when an evaluator wants to investigate an issue in depth and provide an explanation that can cope with the complexity and subtlety of real life situations. In this sense, case studies can be a powerful means to portray the program to outsiders. However, case studies can usually be quite time-consuming to collect, organise and describe, and while they do provide a depth of information, they don't provide much in the way of breadth.

### **Document review**

Document review is a way of collecting data by reviewing existing documents that are either internal or external to the program. Documents can include government publications and official statistics, promotional literature, records of meetings, memos, diaries, newspapers and websites.

Document review can be particularly useful to gather background information and to help determine if implementation of a program reflected planned intentions. A document review is unobtrusive (and will not generally require any interruption of the program) and can bring up issues not detected by other data collection methods.

The main advantages of undertaking a document review is that it leverages off information already available, and so is relatively cost-effective. That being said, there are a number of things to watch out for when undertaking a document review, such as:

- **Relevancy** – whether the information is up-to-date and pertinent to the program in question.
- **Credibility** – whether the documents are free from bias or errors. Credibility can be affected when findings are subject to judgement, when information may have been presented in a

particular light to suit vested interests, or when there is little incentive for authors to ensure the accuracy of the information.

- **Representativeness** – whether the information in the document be generalised. This will depend on whether the document is complete and contains sufficient context, and whether the document is a typical example of the information it is portraying.

### **B.3 Selecting the right method**

Selecting the method for data collection is an extremely important decision when planning and implementing an evaluation. Often, judgement will be required to balance the trade-offs inherent in the different data collection methods.

Some questions that can help guide the decision on a suitable collection method are listed below<sup>19</sup>.

#### **Population and accessibility**

- **Is there a complete list of the population?** Methods such as self-completion surveys or telephone interviews often rely on having access to a complete listing of the units that will be sampled.
- **Is the population literate?** Methods such as self-completion surveys require respondents to be able to read and understand the questions being asked.
- **Are there language issues?**
- **Will the population cooperate?**
- **What are the geographic restrictions?** It will be less feasible to do research that requires personal engagement, such as face-to-face interviews, or observation with respondents if they are widely dispersed.

#### **Response issues**

- **Are response rates likely to be a problem?**
- **Is response bias likely to be a problem?** Response bias occurs when people adapt their response to fit what they think is expected from them. Interviews and self-completion surveys are particularly susceptible to response bias.
- **Is non-response bias likely to be a problem?** Non-response bias occurs when there is some systematic difference between those who choose to participate in a data collection and those who do not respond. Self-completion surveys, for example, are particularly susceptible to non-response bias.
- **Can interviewer or observer distortion be controlled?** The presence of an interviewer or observer can alter responses, behaviours or events. If this effect is likely to be significant (and carries a risk of not being able to be controlled) then it may be better to use data collection methods that don't involve contact with the population, such as document review or a case study.
- **Are accuracy or false responses likely to be a concern?**

---

<sup>19</sup> Adapted from <http://www.socialresearchmethods.net/kb/survsel.php>

## Administrative issues

- **What is the cost?** Considerations include costs such as training and time for interviewers or observers, postage for mailed questionnaires, length of time required to collect observations, or costs of travelling to various program locations to collect information.
- **Are there sufficient facilities and equipment available?** Considerations include access to phone surveying facilities for telephone issues, recording and transcribing equipment for verbal interviews, or software requirements for different types of data analysis.
- **Is there sufficient time to conduct the collection?** Different types of data collection will take longer than others.
- **Is there trained and available staff?** Different types of surveys make different demands of personnel.

**Table 8: Strengths and weaknesses of data collection methods**

Strategies	Particularly useful when:	Advantages	Disadvantages
Self-completion surveys	<ul style="list-style-type: none"> <li>• Gathering data about knowledge, beliefs, attitudes and behaviours</li> <li>• Resources are limited and you need data from large numbers of respondents in many locations</li> <li>• Collecting fairly straightforward and uncontroversial information</li> </ul>	<ul style="list-style-type: none"> <li>• Wide coverage</li> <li>• Relatively inexpensive</li> <li>• Eliminate interviewer bias</li> </ul>	<ul style="list-style-type: none"> <li>• Poor response rate</li> <li>• Potential for response bias</li> <li>• Can be difficult or time-consuming to design</li> <li>• Cannot check the truth of answers given</li> </ul>
Observation	<ul style="list-style-type: none"> <li>• Gathering data on individual behaviours or interactions between people</li> <li>• Gathering information on topics that program staff or participants are unwilling or reluctant to discuss</li> <li>• Data collection from individuals is not a realistic option</li> </ul>	<ul style="list-style-type: none"> <li>• Does not rely on people's willingness or ability to provide information</li> <li>• Relatively objective</li> <li>• Efficient</li> </ul>	<ul style="list-style-type: none"> <li>• Susceptible to observer bias</li> <li>• Can be difficult to interpret and categorise behaviours</li> <li>• Records behaviour, not intentions (what, not why)</li> </ul>
Case studies	<ul style="list-style-type: none"> <li>• Examining processes and relationships</li> <li>• Examining issues that deal with the complexity and subtlety of real life situations</li> </ul>	<ul style="list-style-type: none"> <li>• Can provide a full example of a programs' process, from inputs through to outcomes.</li> <li>• Can be a powerful means to portray the program to</li> </ul>	<ul style="list-style-type: none"> <li>• Can be time-consuming to collect, organise and describe</li> <li>• Represents depth of information, rather than breadth</li> </ul>

		outsiders	
Documents	<ul style="list-style-type: none"> <li>• Gathering background information</li> <li>• To help determine if implementation of the program reflects program plans</li> <li>• When answering questions on “what” and “how”</li> </ul>	<ul style="list-style-type: none"> <li>• Vast amount of information available</li> <li>• Cost-effective</li> </ul>	<ul style="list-style-type: none"> <li>• Authenticity</li> <li>• Credibility</li> <li>• Representativeness</li> <li>• Relevancy</li> </ul>
Interviews	<ul style="list-style-type: none"> <li>• Gathering information on knowledge, beliefs, attitudes, behaviours, opinions, feelings, emotions and experiences</li> </ul>	<ul style="list-style-type: none"> <li>• Insights</li> <li>• Flexibility</li> <li>• High consent rates</li> <li>• Validity</li> </ul>	<ul style="list-style-type: none"> <li>• Time-consuming</li> <li>• Non-standard responses</li> <li>• Reliability issues</li> <li>• Susceptible to interview bias</li> </ul>

#### B.4 Further resources

The following organisations provide a suite of publications outlining different methods for collecting and analysing data:

- University of Wisconsin-Extension publications on [collecting and analysing data](#), including [Collecting Evaluation Data: An Overview of Sources and Methods](#).
- Centres for Disease Control briefs on [data collection and analysis](#) for program evaluation.

The [Queensland Government Statistician’s Office](#) can also provide advice on proposed methods of data collection.