

Victorian Guide to Regulation: Evaluation Toolkit

Victorian Guide to Regulation: Evaluation Toolkit

This toolkit complements the Victorian Guide to Regulation, available at:

www.vic.gov.au/victorian-guide-regulation

You can use parts of this toolkit to help develop an evaluation strategy, and to design or improve evaluation of existing legislation or regulation.

It expands on the requirements, principles and advice for evaluation methodology set out in the Victorian Guide to Regulation, with a focus on how to design and deliver an evaluation.

Contents

Using this toolkit to support design and delivery	1
The three steps to developing an evaluation strategy	2
Getting started: Scoping a regulatory evaluation	2
Questions to consider	3
Objectives to guide evaluation design	3
Types of evaluations and their requirements	4
Step 1: Developing your evaluation strategy	5
Using evaluation questions to develop indicators	6
Developing measures for your indicators	8
Defining your baseline	8
Developing proportionate measures	9
Measuring regulator resource inputs	9
Measuring regulator outputs	10
Measuring behaviour change with regulator observations	10
External measurements of behaviour change	10
Account for both routine and program activities	11
How to undertake evaluation in specific scenarios and different types of evaluation	12

Step 2. Undertaking the evaluation	15
Decide who will conduct the evaluation and when it will be done	15
Undertake proportionate analysis	16
Ensure your evaluation is adequately resourced	16
Recognise technical topics that may require expert support	16
Remember that evaluation can involve continuous measurement	16
Account for the influence of evaluation on delivery of regulations	16
Monitor and be accountable for changes in evaluation approach	17
Monitor frequently enough to track data quality issues	17
Step 3. Writing the evaluation report	17
Supporting strategy with a program logic model	18
Take a proportionate approach that recognises complexity	18
Account for the different dimensions of the regulation	18
Account for what motivates compliance, including external influences	19
Explain connections between regulator actions and compliance drivers	19
Use a program logic to explain relationships between indicators	20
Test cause and effect of activities and outcomes	20
Define a suite of indicators rather than narrow targets	20
Appendix A: Using program logic models for regulator actions	21
Appendix B: Social research of regulated entities	23
Appendix C: Linkages to the VGR	26
Glossary	27

Using this toolkit to support design and delivery

Good policy development requires both early and ongoing planning. This includes the evaluation of current regulations, the analysis of new issues and options and the improvement of regulation to achieve outcomes.

This toolkit is designed to help you plan and deliver an evaluation and meet the evaluation requirements set out in the Victorian Guide to Regulation (VGR).

The toolkit steps through how to consider how the actions of regulators and other parties can lead to behaviour change and drive compliance with regulations, in order to achieve long-term objectives (outcomes).

You can use this toolkit for the following:

Understand and meet VGR requirements

- The VGR sets out evaluation requirements, including when mid-term evaluations are required. It provides guidance on developing your evaluation methodology and advises on how to focus an evaluation and ensure it is proportionate. These requirements are referred to in this toolkit.

Support evaluation design and delivery

- Structure the evaluation strategy in your Regulatory Impact Statement (RIS) or Legislative Impact Assessment (LIA).
- Review and deliver your regulatory evaluation, including mid-term evaluations or evaluations at the sunset of regulations.
- Document your evaluation.
- Support you with any other evaluation of regulatory settings that you undertake.

Account for previous commitments

- Review or improve an existing evaluation, account for commitments made under the relevant RIS or LIA, or other existing program or regulator performance measurements. For example, a RIS may have committed policymakers to gather data to fill knowledge gaps. This will provide a starting point for delivering or enhancing a regulatory evaluation.

Integrate with regulator performance measurement

- Regulatory evaluation often draws on, or influences, regulator activity or performance reporting. Commitments under an evaluation strategy may also support regulators to resource their performance reporting. These interactions are situation-specific and not accounted for in this toolkit. Regulatory evaluations can inform but should not limit how a regulator generally measures its performance.

The advice in this toolkit is general in nature. Its focus is primarily on overall regulatory evaluation design, accounting for the impact of regulator activities. It might also be useful for broader policy evaluation.

In some cases, an evaluation will also need to draw on specialist expertise in evaluation design, or for specific subject matters (e.g. health, education, or environment).

The three steps to developing an evaluation strategy

After you have considered and scoped the evaluation strategy, there are three steps for an evaluation. This toolkit is designed to support and guide you through these steps.

- **Developing the evaluation strategy:** This is a requirement of both a Regulatory Impact Statement (RIS) and a Legislative Impact Assessment (LIA). An evaluation strategy sets out the plan for evaluation years in advance of its full implementation. As such, your approach to evaluation should always be revised at the time of undertaking the evaluation and writing the evaluation report.
- **Undertaking the evaluation:** Performed at committed milestones set out in the RIS/LIA, at mandated evaluation milestones (e.g. mid-term or sunset), and throughout the life of the regulation as part of ongoing evaluation. It can also be triggered by external factors such as challenges to a regulation or changes in the environment.
- **Writing the evaluation report:** Undertaken at committed milestones set out in the RIS/LIA and at mandated milestones (e.g. mid-term and sunset). Continuous or periodic reporting will inform how an evaluation is documented. External factors such as changes in the environment may also instigate reviews of regulations.

In most cases, the findings and recommendations from the evaluation report should be communicated to external stakeholders.

The contents of an evaluation will differ depending on the stage of the legislative lifecycle. Key requirements and differences are summarised in the next section on Getting Started.

Note: The Commissioner for Better Regulation Victoria (BRV) monitors and notifies departments of their evaluation commitments, provides support to departments and agencies undertaking evaluations and reviews evaluation reports. Early engagement with BRV supports policy makers to plan and undertake an evaluation.

Getting started: Scoping a regulatory evaluation

The **scope of a regulatory evaluation** focuses on the effect that a regulatory intervention has on the behaviour of regulated entities and other parties. This includes how the regulations have been implemented, the broader context of the regulatory problem and how it may have changed through the course of the regulatory lifecycle.

The **evaluation strategy** you develop as part of your RIS or LIA sets out how to deliver the evaluation, but the regulatory evaluation itself may be developed or revised during the life of the regulations including implementation.

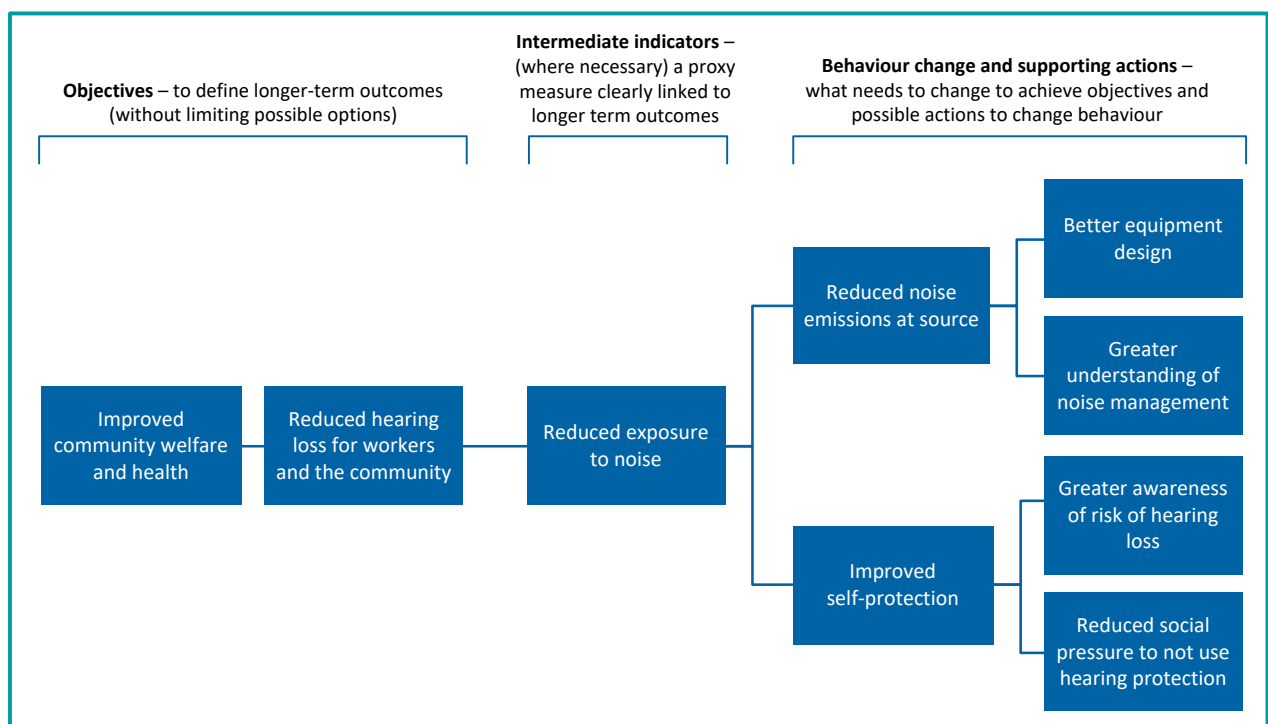
- Evaluation activities such as data gathering and developing 'baseline' measures should occur before a regulation is put in place. These activities should inform the 'problem definition' to explain the need for regulation, e.g. in a RIS.
- The evaluation strategy should set out a range of up-front activities that work out the effectiveness of the regulation over its life.
- The evaluation activities should be embedded in the delivery of the regulation
- Mid-term or other points in the evaluation lifecycle should provide for review and calibration of the evaluation approach.

Questions to consider

- What are the objectives and intended outcomes of the program or activity?
- What is the context of the regulatory lifecycle? At what point in the lifecycle is the evaluation taking place?
- What type of evaluation is required?
- Is there an existing evaluation strategy to guide the evaluation development?
- How do I identify what needs to be evaluated?
- What behaviour changes and supporting actions are needed to achieve objectives?
- What is the best way to measure and assess the outcomes?

Objectives to guide evaluation design

As set out in the RIS or LIA, objectives are established during the initial stages of policy development. These objectives guide the design of the evaluation strategy by outlining measures of long-term success, such as improved outcomes or reduction in harm. Intermediate indicators may be developed to assess progress towards achieving long-term objectives. Behaviour change and supporting actions to achieve objectives need to be identified. An example of this process is provided in the figure below.



The evaluation will need to address these elements **at a minimum**, but also account for the effectiveness of the regulation and actions of government in creating the desired change.

This exercise can be challenging, as it may involve accounting for events that did not happen (e.g. harm prevented).

However, you can use a **program logic model** that includes indicators, explains relationships between measured activities, sets out why you expect certain activities to lead to outcomes (causal assumptions), and accounts for external factors. Find out more about how to develop a program logic model in *Appendix A*.

Types of evaluations and their requirements

Evaluation strategies are required in RISs and LIAs. Depending on the context of the regulatory lifecycle, and where evaluation falls in the lifecycle, evaluations have differing requirements as set out under the VGR.

Evaluation strategies for new and amending regulations

What to do

- Set out a strategy to evaluate the extent of the change or new law or regulation at certain future milestones (e.g. 3, 5 and 10 years)
- Set out an evaluation strategy that meets VGR requirements
- Capture evaluation commitments at a high level to ensure you can do further work and relevant parties are accountable.

What to document

In the RIS/LIA 'evaluation strategy' document a clear method for evaluating the actual effects of the preferred option. The evaluation strategy will explain:

- what will be evaluated
- how it will be done
- who will do it
- when it will be done.

Mid-term evaluation

What to do

- At a minimum, deliver against commitments made in the RIS/LIA
- As required, draw on up-to-date guidance in the VGR and this toolkit, to build on the minimum requirements
- Analyse the effectiveness of the regulations and document the results
- Focus on the most important attributes of the preferred option (as set out in the impact assessment)
- Can involve both evaluation of data and qualitative feedback
- May identify how the problem has changed over time and whether adjustments need to be made (in regulations or in delivery) given what was learnt and what has changed

What to document

In your evaluation report, apply the strategy set out in the RIS/LIA as a starting point, and describe:

- the focus and scope of the evaluation
- the parties involved
- the findings of your evaluation, including any unintended impacts
- stakeholder feedback and findings (e.g. implementation issues)
- how well the preferred option performed
- external factors that influenced results
- key adjustments required to the regulations or their implementation, in response to evaluation finding.

Ongoing evaluation

What to

- Identify and respond to issues as they occur
- Build evaluation into ongoing delivery across the regulatory lifecycle

What to document

- Similar material as for a mid-term evaluation. May focus on a subset of indicators or a rolling set of evaluations that look at different aspects of the preferred option and its effects.

RIS when regulations are sunseting and under review

What to do

- At a minimum, evaluate existing regulations and set out a new evaluation strategy for the next cycle, in line with the requirements for a RIS about new or amending regulations.
- Drive improved understanding of the problem, filling knowledge gaps and addressing issues identified when making the regulations.

What to document

As an input into the RIS (e.g. problems definition), validate the current regulations accounting for:

- commitments and plans made in the initial RIS/LIA
- current guidance (e.g. this toolkit)
- other external factors.

Document a new strategy afresh for any proposed replacement regulations.

Step 1: Developing your evaluation strategy

The evaluation strategy you develop as part of your RIS or LIA should outline how to deliver the evaluation. It is typically reviewed and adapted as part of ongoing evaluation processes. The evaluation strategy sets out what will be evaluated, how it will be done, who will do it and when.

Evaluation activities such as data gathering and developing 'baseline' measures should also occur before a regulation is put in place. These activities should inform the 'problem definition' to explain the need for regulation, e.g. in a RIS

To account for these and develop your evaluation strategy you will need to consider the following:

- Using key evaluation questions to develop indicators
- Developing appropriate indicator measures and establishing a baseline
- Undertaking evaluation in specific scenarios and different kinds of evaluation
- Developing a program logic model that links regulator actions with behaviour change
- Accounting for external factors that will influence outcomes.

Using evaluation questions to develop indicators

Once objectives or outcomes are established, an evaluation strategy can be designed around key evaluation questions that help define indicators.

These address important gaps in knowledge, areas of uncertainty and change, attributes of the problem, regulation provisions, and resourcing commitments.

In order to establish indicators, evaluation questions should seek to establish and measure three components (shown below) which form the basis of the evaluation process and show if objectives have been achieved.

The three components are:

- **Effectiveness:** the extent to which the regulation achieves its intended objectives. It includes the extent to which there are adverse or unintended consequences. Evaluations involving regulator actions should often seek to measure the following aspects of regulator delivery:
 - efforts to monitor non-compliance were well-targeted, detected issues, and created a deterrent (including through related communications)
 - actions to respond to non-compliance created specific or general deterrence
 - other enabling activity (e.g. disseminating guidance) helped to increase industry understanding and performance.
- **Efficiency:** this refers to the relationship between resources and outcomes. An efficient outcome is one that produces the desired benefit (outcome) at the least cost, or maximises benefits for a given level of cost.

Efficiency involves looking at the relationship between inputs and outputs (e.g. the ratio of effort/resources applied to a given result). Common inputs include use of regulator or service agency resources (e.g. inspection efforts).

It is important to consider how those inputs were deployed or aligned (e.g. if inspections were deployed to problem areas) as well as resource efficiency (e.g. the costs of inputs to deliver a quantity of outputs).

Evaluations involving regulator actions should often seek to measure the following aspects of regulator delivery:

- resources were deployed to the right focus areas (e.g. monitoring higher-risk industry more than good performers)
 - resources were productive, including the time and budget involved (this may account for costs to industry)
 - overheads were well managed and efficient rework was minimised, for example if complaints were managed efficiently without having to be reopened.
- **Intrinsic Value:** Intrinsic values are not directly part of delivering the intended objectives but are important to measure (and possibly to maintain or promote). They may reflect procedural fairness and transparency in decision-making, effective engagement with industry, and efficient compliance processes and requirements. Consider what values are intrinsically important to government policies and activities more generally (e.g. equity in decision-making). This may be difficult to measure and may involve qualitative evidence-gathering techniques. Consider whether other objectives (e.g. efficiency) could have a negative impact on intrinsic values and account for this in your strategy.

Note: Key evaluation questions should be:

- **Specific and observable:** Indicators should provide a clear description of what you want to measure along with a focus on the desired action or change
- **Measurable:** Quantify the change and is generally reported in numerical terms using relevant data sources

Effectiveness example

These sample evaluation questions were part of an evaluation of community-based interventions for perpetrators of family violence in 2018. These interventions were in response to the Royal Commission into Family Violence. They incorporate the Department of Treasury and Finance's Lapsing Program Evaluation questions.

Sample question 1

- Are there early positive signs of change that might be attributable to the initiatives?

Indicators

- Increase in people who experience violence feeling safe and supported
- People who use violence report to understand the factors contributing to their behaviour, and how it impacts others.

Effectiveness measurements

- Changes in perception of safety and support of people experiencing violence
- Changes in participants' views on their responsibility for perpetrating violence or using force

Data sources

- Royal Commission into Family Violence
- Literature review
- Victorian crime statistics data

Sample question 2

- To what extent are the outputs being realised?

Indicators

- Uptake of programs among people who use violence and people who experience violence
- Number of people who use violence attending interventions
- Number of families involved in Aboriginal-based programs
- Extent to which the participant numbers are as expected

Data sources

Document review and program administrative data.

Efficiency example

Sample question

- Has the department demonstrated efficiency in relation to establishing and implementing the program?

Indicators

Government resources used to implement the program have not been wasted

Efficiency measurements

- Budget and FTE used to support program delivery
- Program implementation staff's views on the resources required to effectively implement and monitor the programs

Data sources

- Program documentation
- Stakeholder consultations – service providers

Developing measures for your indicators

With an understanding of the indicators that you intend to measure, and the way that you will organise results (through a **logic model**), you need to plan for what and how you will measure and report. This section takes a closer look at the measures you choose to develop. This section includes:

- Defining your baseline
- Developing proportionate measurements
- Measuring regulator resource inputs
- Measuring regulator outputs
- Measuring behaviour change
- External measurements of behaviour change
- Accounting for routine and program activities

Defining your baseline

Determine what baseline (or reference point) you will use for evaluation measurement. The approach to this will be shaped by your approach to impact assessment, as explained in the VGR. As appropriate, account for the following:

- **The base case:** What would be the outcome if there were no policy in place, as set out in your problem definition in a RIS or LIA.
- **A reference case:** In some situations, the point of reference for impact analysis is a modified form of the base case, accounting for anticipated changes or interventions beyond the regulation, other commitments, or minimum necessary elements of the regulation.
- **Projections:** What is the expected future state once the regulations are introduced, including population, environmental or industry trends. These should generally be incorporated into the base case but may be important to specifically call out in the evaluation strategy.
- **Jurisdictional analysis:** What the case may be in other states that have or have not introduced the same or a similar policy. This may be an important additional factor to build into evaluation, that may not be fully built into your problem definition or impact analysis.

Developing proportionate measures

The evaluation strategy should drive quantitative data collection that enables policy makers to explain interactions and effects. This may include investment in data-gathering activities and systems.

Use a suite of quantitative measures. These can include:

- measurement of industry uptake of regulations
- government or regulator activity to embed (implementation and compliance) regulations
- the effects of this on awareness, attitude and compliance behaviour
- the effect of compliance on intermediate and long-term outcomes.

However, policy makers can encounter legitimate barriers that make a fully quantitative approach to evaluation not practicable.

In many cases, qualitative or mixed methods of evidence gathering are preferable, to give a full view of the regulation or the changing policy problem.

When developing the evaluation strategy or delivering a mid-term evaluation, you should account for a full range of methods. For example, the role of focus groups, industry consultation, expert opinion, or reviews of specific programs or activities.

Consider what methods are feasible, proportionate and most likely to produce useful insights in line with your key evaluation questions.

Measuring regulator resource inputs

To measure regulator resource inputs, regulators will need to account for process design. Consider whether the following indicators could help to evaluate regulator resource inputs and whether these inputs are efficiently utilised.

Potential inputs

- Number of staff, or hours, allocated or budgeted to a program area
- Capital or operational expenditure, and assets assigned
- Functions that enable 'frontline' activity – including overheads, regulatory support services, staff training and quality assurance.

Potential resource efficiency considerations

- How processes are designed, e.g. if licence applications are 'streamed' into different risk or complexity categories
- Timeframes for activities to be done, and flow rates (e.g. number of applications received vs concluded, by category)
- How outliers or complex cases are managed and if processes minimise unnecessary variation and manual work.

Measuring regulator outputs

An evaluation can seek to understand whether regulator activities were effective and appropriately targeted, by looking at 'rates' (e.g. in percentages) and results such as the following:

- **Detection rates:** How often inspections or compliance monitoring observed minor or significant issues, and recidivism
- **Escalation rates:** How often non-compliances were acted on
- **Results of partnerships or programs:** Activities such as delivery concluded
- **Results of trials, experiments or campaigns:** Findings.

In most evaluations, these rates and results will be a key indicator to track. However, reporting may need to account for sample bias, e.g. whether inspections were representative of the regulated population.

Measuring behaviour change with regulator observations

To measure behaviour change, consider the following regulator internal measurement methods:

- Recording field or inspection observations, which can involve:
 - collecting indicators of industry engagement, willingness and ability to comply during interactions with the regulator
 - staff 'rating' interactions with regulated entities within their inspection records (e.g. scores for level of willingness to comply, and ability to comply)
 - debriefs, workshops or surveys of frontline staff about industry behaviours
- Consultations with regulated entities or communities, such as:
 - routine feedback, such as industry reference groups
 - specific engagements at an evaluation milestone.
- Tracking rates of detection of non-compliance and recidivism or recurrent non-compliance, and escalation rates against non-compliances, as a proxy for general behaviour change. Account for the population of entities sampled.

When using consultation with staff or regulated entities (e.g. workshops and surveys), take care to recognise and identify the potential biases that may be present.

External measurements of behaviour change

External measurement of behaviour change can involve social research and using industry data.

For social research surveys of regulated entities (See **Appendix B** for model questions for a survey):

- Plan and resource at key milestones such as:
 - before the regulations took effect
 - at mid-term evaluation
 - at sunseting.

- Consider how research or other feedback can be built into routine interactions e.g. follow-up- surveys
- Research can examine drivers of compliance and the effect of regulations and regulator activities on these.

Industry data may be varied and require information sharing agreements and anonymisation, such as results of:

- internal audits or compliance monitoring
- industry peak body support programs
- compliance management performed by industry on its members corporate governance and quality assurance reporting.

Account for both routine and program activities

The regulator activities that you measure will depend on the nature of the regulation, the commitments made in the implementation plan (or in related programs), and the scope and powers of the regulator. Consider whether the following may be relevant:

- Inspections conducted, or other forms of compliance monitoring e.g. audits and desktop reviews. Considerations:
 - This might include general regulator activity or specific programs, campaigns or strategies.
 - It may be more effective to evaluate the design and execution of a specific compliance monitoring campaign, than to report on the overall number of inspections.
- Education and outreach activities, which might be part of:
 - Specific inspection campaigns
 - Routine inspections by authorised officers
 - Activities by engagement officers.
- Enforcement actions, which might include:
 - In-field actions, including remedial or corrective action (e.g. issuing improvement notices or giving directions)
 - Issue of penalties or fines
 - Prosecutions or disciplinary processes conducted.
- Partnerships such as working with industry to co-design guidance, or to educate or support members in meeting requirements
- Other compliance generating or monitoring activities, which can include tracking industry incident reporting and response to these, managing queries, or tracking consumer or citizen complaints.

Common challenges to plan for in evaluation planning and delivery

Evaluation project management, delivery and the approach to stakeholder engagement should be mindful of common issues that may arise during the lifecycle of the regulation. When delivering an evaluation, challenges can arise where:

- regulatory objectives are not sufficiently articulated or focused to provide clear indicators to measure success or failure
- there were shortcomings in the design of pre and post-intervention measures, or gaps in accounting for intervening factors that may confuse the result
- documentation developed when making the RIS or LIA was not suitably stored or made accessible for future use
- filling evidence gaps, or improving data quality, was insufficiently done because additional expertise or expenses were not fully accounted (or allowed) for
- evaluation reporting highlights issues or shortcomings and changes required in delivery. This might come against organisational resistance to change e.g. in government bodies or industry. This can occur where changes may threaten the existing arrangements, positions or budgets, or parties who have benefited from the current regulatory arrangements.

Ultimately, shortcomings in evaluation can undermine an effective review of sunseting regulations, limiting the ability to improve regulations. It is important to build in enough time to investigate, assess, and respond to issues (e.g. by reviewing your program logic model or gathering alternative data).

How to undertake evaluation in specific scenarios and different types of evaluation

This section discusses how to undertake evaluation in specific scenarios and undertake different types of evaluation, including:

- Interactions with programs or projects
- Regulator reporting and performance evaluations
- Fee setting regulations
- Permissions evaluations
- Mid-term and sunseting evaluations

It is important to consider how the regulations were implemented when developing your evaluation. How successful was implementation? How were the regulations communicated and 'bedded in' with stakeholders? What approaches has the regulator taken to compliance and enforcement? What is the current level of compliance and how has this changed over time? Have there been any reviews of regulator delivery?

Interactions with programs or projects

A regulatory evaluation may interact with a service delivery or regulator program evaluation. This happens when a regulatory intervention is complemented by government activities or funding.

For example, regulations designed to reduce agricultural runoff might be complemented by an education program to reduce runoff and subsidies for equipment upgrades.

In these cases, the evaluation strategy should aim to isolate the effects of the regulatory intervention from other actions.

This can include clarifying what actions were dependent on regulation being in place. Doing so helps to ensure mid-term evaluations, or reviews of sunseting regulations, can describe the role of the regulation compared to other actions.

This promotes continuous improvement in regulatory design.

When other evaluations are taking place with related problems, such as evaluations aligned to DTF's Resource Management Framework, you should seek to draw upon relevant materials for your evaluation.

Regulator reporting and performance evaluations

Regulator activity and performance measurement and reporting can relate to how regulators deliver against specific programs, and the performance of a regulator in addressing its objectives.

Regulators will often report against a range of objectives and potentially more than one piece of legislation. This means a regulatory evaluation may be only one part of their reporting obligations.

This toolkit does not directly address how to evaluate regulator performance. However, some of the measures for measuring regulator performance may be relevant in determining the effectiveness a regulation.

You will need to keep these relationships and distinctions in mind. For example, while most input measures would only be used internally by a regulator, they might inform periodic reporting for a regulatory evaluation.

Understanding this relationship between regulator internal performance measurement and the evaluation can help ensure an effective evaluation design.

Fee setting regulations

An evaluation strategy for fee setting often needs to look at delivery efficiency and how processes can be optimised.

DTF's [Pricing for Value](#) guidelines set out 12 pricing principles for setting fees that can inform evaluation strategies. Principle 12 states that fees should be monitored and reviewed periodically.

The evaluation also needs to consider how fees affect behaviours (e.g. testing the effects of different pricing of licence applications, or how pricing might drive unintended consequences such as black-market activity).

Fees evaluations can be relatively straightforward if you have good cost data. This may require planning in advance to obtain data from the regulator or delivery agency, or to develop an activity cost model or to commission a time-and-motion study to systematically observe, analyse and measure the steps involved in regulatory activities (for higher-cost situations).

Consider IT system requirements for monitoring related activities and determining a cost base. Consider whether these costs should be recovered.

Permissions evaluations

Design of permissions (e.g. licences, permits and registrations) and their evaluation should follow the Victorian Permissions Framework: www.vic.gov.au/victorian-permissions-framework-guidance.

Guide 2, stage 6 sets out specific evaluation concepts for permissions. For example, evaluating whether the coverage and scope of permissions was effective, whether costs were justified, and whether entities complied with requirements, or developed workarounds or misleading reporting to give the false appearance of good performance. The guide also sets out example key performance indicators to support design of evaluation measures, as well as general regulator activity reporting.

Mid-term and sunseting evaluations

Key evaluation questions should address major areas of uncertainty and impact. When developing a mid-term evaluation, or evaluating sunseting regulations, consider the areas of the RIS or LIA where there was uncertainty or unresolved questions. Consider the following when developing your evaluation strategy:

- **Problem definition** – do you have gaps in knowledge or understanding of the problem, and can you fill them?
 - If the nature of the problem is dynamic – has it changed? Have there been changes in technology or the characteristics of the industry? Have the regulations kept pace?
- **Objectives of action** – are you meeting them?
 - What were the intermediate outcomes, how are they tracking?
- **Options** – what options were analysed in the RIS or LIA and what was the preferred option? In hindsight was the preferred option the best option?
- **Impact analysis** – What were the key drivers of estimated benefits and costs of options in the previous RIS or LIA? In hindsight, were these the key drivers? Were the assumptions made and data used appropriate? Have estimated benefits been realised? Have costs been higher or lower than estimated? Have there been any unintended consequences?

Step 2. Undertaking the evaluation

When planning evaluation delivery, you will need to consider the following:

- Decide who will conduct the evaluation and when
- Undertake proportionate analysis
- Ensure your evaluation is adequately resourced
- Recognise technical topics that require expert support
- Remember that evaluation can involve continuous measurement
- Account for the influence of evaluation on delivery of the regulations
- Monitor and be accountable for changes in evaluation approach
- Monitor frequently enough to track data quality issues

Some milestones are mandated and should help anchor project planning:

- Mid-term evaluations completed 3–5 years from commencement
- Sunsetting evaluations completed prior to regulations sunsetting after 10 years.

You need to start evaluation activities well in advance of these milestones. For instance, to ensure that the outcomes of the evaluation can inform any proposed remake of regulations at the sunset date.

Typically, preparing for evaluation 'reporting' can take well over six months, or longer where you rely on external parties to provide data.

Considering this, develop other milestones, such as for securing agreements with stakeholders, making investments, sharing preliminary results, or establishing governance and reporting for periodic measures.

Decide who will conduct the evaluation and when it will be done

Define the body responsible for conducting the evaluation. This may be different to those implementing the regulations, due to resources, capability, and actual or perceived conflict of interest. For example, while regulators may report on many key indicators, they may not be best placed to conduct the overall evaluation.

You may need to develop agreements for data gathering, reporting and evaluation roles. Prompts can include:

- Review what external data or evaluation sources or research bodies may be available and can be drawn on or partnered with
- Work out which partners you want to work with, including evaluation experts
- Account for working with regulators on their activity reporting
- Consult with stakeholders and consider factors, such as resources, capability and access to the necessary information and data.

Consider if the required datasets are developed or need to be developed, and document agreements to ensure they will be available on time. Ideally, data will be gathered to develop baseline measures before the regulations take effect, and further data will be gathered over months or years before the evaluation reporting occurs.

Undertake proportionate analysis

The depth and breadth of the evaluation should be proportionate to the impacts of the regulations. The evaluation should focus on analysing how effective the regulations have been and how the problem has changed over time. The evaluation should account for residual policy questions or uncertainties acknowledged at the time the regulations were made. The evaluation report should complement and build from the RIS or LIA, but will be shorter and more targeted in focus.

Ensure your evaluation is adequately resourced

The scope of your evaluation will determine the level and types of resources needed. Once you are undertaking the evaluation, new questions and decisions around budget, resources and consultation effort are likely to arise.

Often, these new questions and decisions arise after the regulations are made (sometimes several years later) and might involve a different decision maker. As a guide to making these decisions, consider the depth of analysis contained in the initial RIS or LIA, the evaluation strategy itself, and the topics that would benefit most from evaluation and scrutiny now that the regulations are in place.

Recognise technical topics that may require expert support

In some evaluations, you may need expert support to develop your theory of change, account for external factors, measure long-term outcomes, or define intermediate outcomes that show progress against ultimate objectives.

Topics where specialist support is often required include environmental impacts, health impacts, community behaviour change, and essential services market analysis.

In many cases, evaluation of long-term outcomes in these areas can be supported by evaluation of the intermediate outcomes enabled by government action – such as the activities that increase compliance in industry.

Remember that evaluation can involve continuous measurement

An evaluation strategy may involve key periods where specific research is conducted or results are reported. However, evaluation will often also involve ongoing or continuous measurement, such as from regulator activities. Be careful to account for the timeframes for both, and not to treat evaluation as a standalone event.

Account for the influence of evaluation on delivery of regulations

Evaluation activities can help to inform ongoing delivery of a regulation and promote continuous improvement. Mid-term evaluations can also lead to changes in regulatory settings or regulator operations.

Account for the activities that may need to occur to improve delivery of regulation in response to evaluation findings. Also consider the balance between continuous improvement and maintaining stability and certainty in regulatory requirements.

Monitor and be accountable for changes in evaluation approach

You may need to revise the scope of your evaluation based on the resources you have available (e.g. rather than conducting expansive focus groups you may choose to send a survey to stakeholders). However, note the impact this will have on the quality and value of your findings.

Make sure key decisions around resourcing (e.g. change in investment, or trade-offs in investment) are planned for, well-documented and justified.

Monitor frequently enough to track data quality issues

Data quality is a common challenge in evaluations, and can diminish the accuracy of ongoing measurement, or undermine periodic assessment of results.

If data quality is not attended to proactively, it may be too late to have adequate data at key milestones. Project planning should account for the systems for monitoring and improving data quality at appropriate stages.

Step 3. Writing the evaluation report

In writing your evaluation report, set out your evaluation strategy and any key changes made to it over time (e.g. since the evaluation strategy was developed for your RIS or LIA).

Quality documentation and record-keeping when preparing the RIS or LIA will assist you to scope your report. One way to present results is against the key evaluation questions.

Areas to consider in your report include the following:

- Are there any external factors that could have contributed to the results?
- How could the legislation or regulation be improved?
- What is the impact of any unintended consequences?
- Which stakeholders are most affected?
- Is there enough flexibility to enable proportionate and risk-based enforcement?
- What are the next steps?

The evaluation report should build from and complement the relevant RIS or LIA but will be more focused and shorter than these documents. In the case of a sunseting RIS, the evaluation will likely be embedded in the RIS itself and will inform the problem definition.

For mid-term evaluations, BRV can provide feedback on the approach to the evaluation and on draft evaluation reports, including proportionate approaches to meet the best-practice standards in this toolkit.

Departments are required to provide the final evaluation report to the Commissioner, who provides advice when the evaluation report is complete.

Based on your program logic developed during the evaluation strategy (and included in the evaluation report), you can explain the results of your evaluation.

This will include how the results demonstrate the extent of the desired behaviour change and how the change leads to the ultimate outcome (i.e. your objective, or if not possible your intermediate indicators).

Supporting strategy with a program logic model

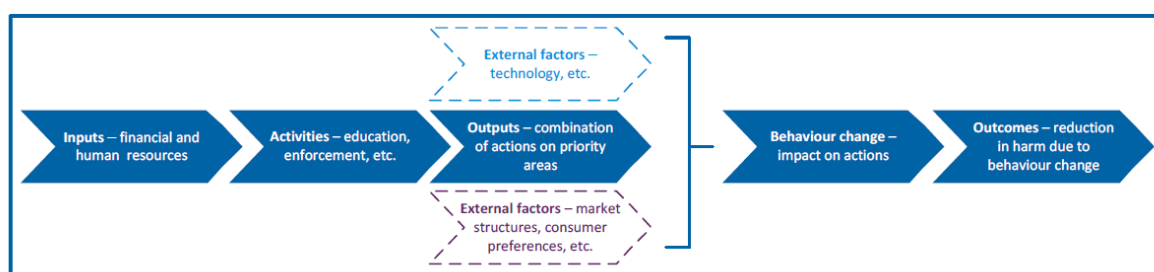
Drawing on the RIS or LIA *problem analysis*, the *objectives of action*, and *options assessment*, you should set out a hypothesis of why and how you expect the regulation (and any related activities) will bring about the desired outcomes.

For example, why following a new or changed rule should lead to reduced harms, or why new standards are expected to change behaviours.

Thinking through these connections, you should also account for the other enabling actions, such as those in your implementation plan. You should also account for external factors and the counterfactual, i.e. what might happen if the regulations did not exist. For example, account for the effect of changing industry characteristics or technology on a given harm or problem.

These concepts are illustrated in figure 5, taken from the VGR.

Figure 5 – Elements of a regulatory program logic model



Take a proportionate approach that recognises complexity

The level of effort taken to explain the theory of change will depend on the complexity of the regulation and the external factors, and what is proportionate given the degree of impact or cost of the regulations. This includes when:

- causality is difficult to establish
- timeframes for changes in ultimate outcomes are very long, or
- change is subject to significant external factors.

The evaluation may need to rely on measuring behavioural or other intermediate outcomes and explain why these should lead to long term outcomes.

Account for the different dimensions of the regulation

To help define intermediate outcomes, work through how the regulations are expected to bring about change. For example:

- key provisions with an impact on behaviour
- new or changed obligations on regulated entities or government
- implementation activities (establishment and embedding)
- broader activities (e.g. the impact of an education or compliance campaign, or the introduction of regulation with funding).

Account for what motivates compliance, including external influences

For evaluations where there are existing rules and issues with the rate of compliance, there are often both government and non-government factors that determine whether regulated parties comply. For example:

- their knowledge and acceptance of rules
- the economic advantages of non-compliance (e.g. reduced costs)
- perceived risks of being reported, inspected and detected
- perceived risks of sanction and severity of sanction.

An evaluation may need to examine these drivers of compliance. This helps to properly attribute the effect of the regulations and the regulator's actions and identify other areas to change or acknowledge when setting indicators and goals for effectiveness.

The Problem Analysis Toolkit sets out how to examine drivers of compliance, which can be built into your evaluation.

Account for external (non-government) factors that are likely to influence compliance, such as economic pressures, structural changes in industry, influence of other regulatory or self-regulatory actions, market or environmental changes, or non-government parties monitoring compliance and publicising instances of non-compliance. These factors may make regulated parties more or less likely to seek to save costs by avoiding their obligations.

Also consider key risks that may be outside of the regulator's control (or government's ability to influence) such as market shocks or environmental pressures, which may impact the effectiveness of the regulation.

Explain connections between regulator actions and compliance drivers

Drawing on your problem definition, it may be important to expressly set out why you expect regulator actions will drive or increase compliance. For example, if a reason for non-compliance is suspected to be:

- a *low threat of detection*, you might evaluate if regulator inspection targeting and communications increase perception of detection
- a perception that the *costs of compliance outweigh the benefits*, you might evaluate if sufficient penalties are applied for non-compliance or whether the regulations are unnecessarily burdensome
- a perception that the *consequences of non-compliance are low*, you might evaluate sanctions issued, and how these were communicated.

Take time to explain these connections when designing your evaluation. See **Appendix A** for further information.

Use a program logic to explain relationships between indicators

Whilst every context and evaluation differs, there are common approaches you can take to assess the relationship between regulator inputs (resources used), activities (actions taken) and behavioural change.

Many of these approaches may be applied by regulators within their own performance measurement and reporting activities. Consider how you can selectively draw on, or help develop, these to provide key indicators for a regulatory evaluation.

Figure 1 illustrates measures that can help evaluate connections between regulator activities and outcomes, which can help you to develop a **program logic model**. This model should also account for external influences on behaviour, and clarify the scope or limits of the regulator’s role. **Appendix A** provides specific guidance on how to develop a model for regulator actions.

Test cause and effect of activities and outcomes

It is unlikely that regulator activities are directly responsible for achieving behavioural outcomes alone. However, the evaluation might seek to understand relationships between activities (e.g. inspections) and observed behaviours (e.g. observed compliance), and infer whether relationships were causal or not (accounting for external factors which may have caused a change in behaviour or compliance).

Exploring these relationships can help to test and refine your program logic model. Your findings can also help to ensure that resources were deployed as intended and in a targeted way.

Define a suite of indicators rather than narrow targets

Figure 1: Example of related indicators involving regulator activity (adapted from M. Sparrow, the Regulatory Craft, p119)

Resource inputs and efficiency	Regulator activities (actions taken)	Regulator outputs (effectiveness/targeting)	Behavioural outcomes (attitudes and actions)	Effects, impacts and outcomes (to the harm)
<ul style="list-style-type: none"> Agency resources Risk-based streaming of processes Variation management Timeframes Flow/turnover 	<ul style="list-style-type: none"> Inspections Education and outreach Enforcement actions Partnerships Other compliance generating or monitoring activities 	<ul style="list-style-type: none"> Detection rates Escalation rates Partnership results Engagement results Measured effect of trials 	<ul style="list-style-type: none"> Non-compliance rates Change in awareness Change in capacity Change in practices Risk reduction Voluntary action 	<ul style="list-style-type: none"> Decline in incidents Increased quality Environmental condition

Remember that when building a program logic model, it is usually better to have a suite of indicators that help to provide an overview of activities and their possible effects over time, rather than setting a narrow set of indicators that act as performance targets.

Appendix A: Using program logic models for regulator actions

If the regulation involves compliance requirements on industry or the community, and regulator actions to promote these, then it is often useful for your program to consider whether (alongside external factors):

- Regulator action is deterring regulated entities from non-compliance
- It is easy to comply with regulations.

These concepts tend to apply regardless of the long-term outcomes intended by the regulation (e.g. the environmental, social or economic objectives), and are linked closely to the concept of drivers of compliance.¹ They are illustrated overleaf in figures 2 and 3, which show the parts of a program logic related to regulator actions and their immediate outcomes.

The figures show how regulator activities work together to achieve outcome, and why accounting for these holistically can be important. For example:

- **For evaluating deterrence effects:** recognise that to effectively measure a deterrence on a group of regulated entities, you need to assess both the perception that there is a credible threat of being detected, and whether there is a perception that the consequences for non-compliances are proportionate. Indicators may come from regulator compliance monitoring activities (e.g. inspections) and surveys of how they are publicised and perceived.
- **For evaluating ease of compliance:** recognise that you may need to gather data on engagement activities with industry and feedback on guidance design and accessibility, combined with tracking campaigns or inspections to educate industry or assess the level of compliance, and associated perceptions of whether rules are fair, relevant to industry operations, and straightforward to follow.

¹ See the **problem analysis** toolkit, which summarises Parker and Nielsen's Regulatory Theory Chapter 13 – Compliance: the 14 Questions – www.jstor.org/stable/pdf/j.ctt1q1crtm.21.pdf

Figure 2: Evaluating deterrence in a program logic approach (to link to outcomes)

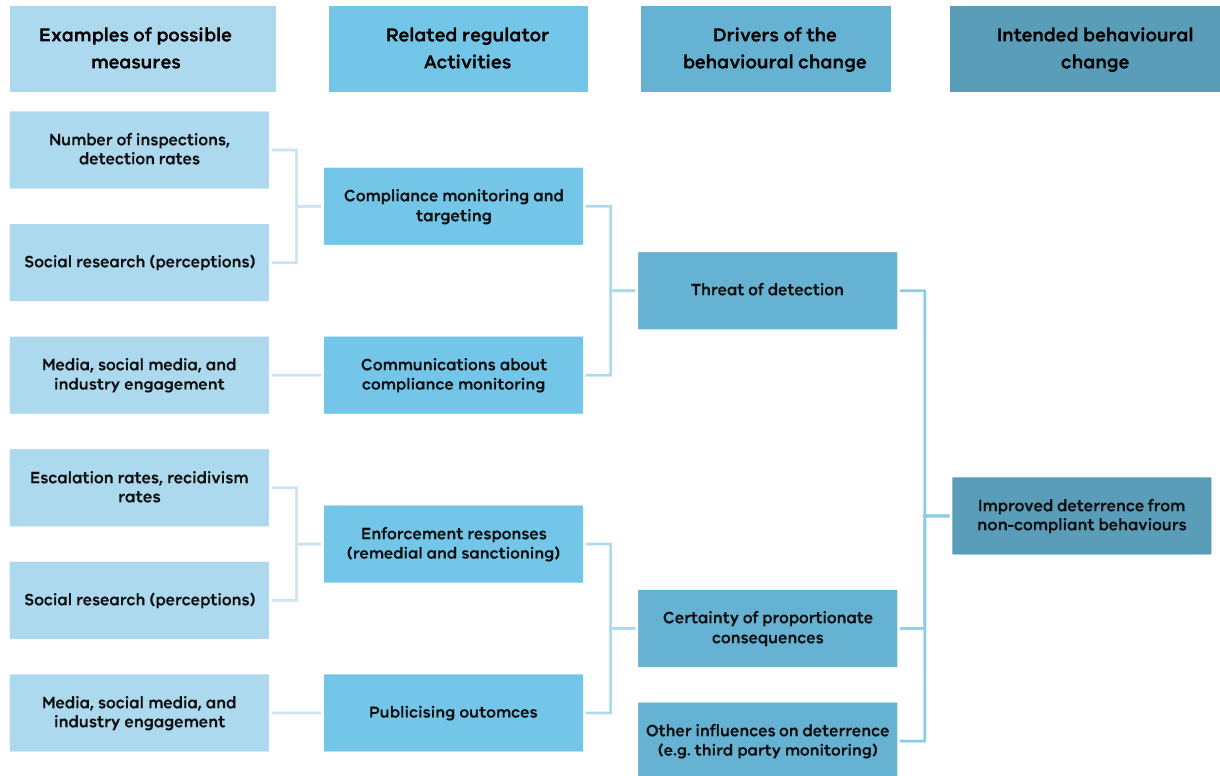
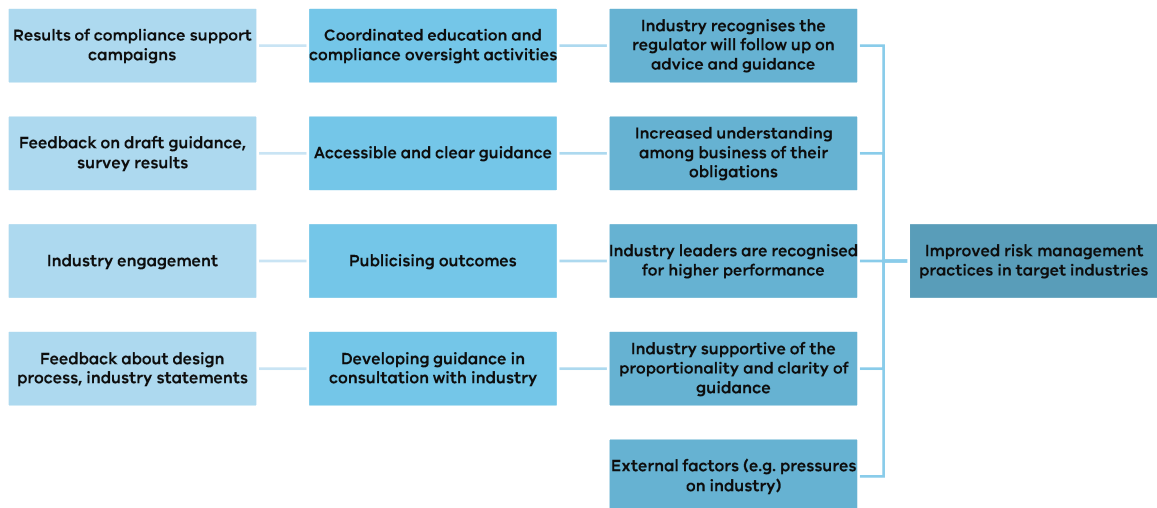


Figure 3: Evaluating improved performance in a program logic approach (to link to outcomes)



Appendix B: Social research of regulated entities

Social research, whether in forums, surveys or interaction specific 'polling', is a key tool that may provide insights into the behavioural factors that motivate compliance or non-compliance.

Many evaluation strategies of regulatory performance will survey those subject to a set of rules to understand their perspectives (regulated entities, or 'duty holders').

The following attitude/compliance and experience of regulation questions are examples that may support social research. It is important to note that the evaluation strategy, via the program logic and key evaluation questions, should consider how to link the responses to these questions to the expected effects and impacts of the regulation, and the regulator's activities.

Meaningful social research relies on a research design that considers:

- what to ask, who to ask, and how to ask them
- the representativeness of the sample of people/businesses surveyed, and the ability to generalise from it (infer validity of the results to a larger population)
- possible errors and biases in results that come from how questions are asked and who participates.

This may require the expertise of social scientists, who can advise on design and different methods for collecting evidence, including dedicated research, or gathering views as part of normal interactions with regulated parties.

While social research plays an important part in the evaluation strategy, efforts and costs incurred should be proportionate to the regulations. Sometimes, the research for the regulations can be integrated with other social research such as regulator-driven surveys, or other industry surveys.

If your evaluation incorporates social research to examine the drivers of compliance, it can draw from the questions and statements below. Responses to these questions may help you establish a problem definition or determine a cause of low rates of compliance. The statements below about the experience of regulated parties may also be helpful to understand the behavioural factors that motivate compliance or non-compliance, and the influence of regulator activities on compliance.

<p>Compliance and regulator performance (expressed in 'agreement with statement' scale)</p>	<ul style="list-style-type: none"> • Regulator inspections occur so rarely that [I/my] business is not very concerned about breaking [rule] • It is hard for [regulator] to detect when I am [breaking rule] • It is easy to [break rule] of the [legislation] without knowing [it/that I have] • [In my organisation] complying with [legislation/rule] is not given a high priority • Even if [regulator] noticed that we had [broken rules], they wouldn't be able to make [me/us] do anything about it • The punishments for breaking [rules] aren't strong enough to stop [others/businesses] from breaking the laws • [Regulator] will hold you to account for the [offences you commit/harms you cause] • It is easy to get information about [laws/rules] • [Local community/users/customers / affected people] are likely to report businesses in my sector who [break rules] • Employees in my organisation (who need to know) are clear about how [laws/rules] apply • [Regulator] has a long history of systematically targeting its inspections at businesses in my industry
<p>Views of regulatory regime</p>	<ul style="list-style-type: none"> • Are [regulator] requirements proportionate to the [class/type] of risk? • Do [regulator] [licence/permit/approval] requirements drive adoption of best practice?
<p>Views of being compliant</p>	<ul style="list-style-type: none"> • Overall do you agree or disagree that the advantages to [you/your business] of meeting [rules/laws] outweighs the cost incurred to [you/your business] to meet these laws? • What type of costs, if any, do [you / your organisation] incur to meet [laws/rules]. (Accept multiples, aid/prompt if necessary) e.g. <ul style="list-style-type: none"> – Equipment or controls – Costs of systems – Administrative (e.g. paperwork) – Sourcing requirements from regulator – Cost of training staff – Expert advice (e.g. consultants, lawyers) • What, if any, business advantages come from meeting the [laws/rules? (Accept multiples, person conducting survey does not read out/prompt) <ul style="list-style-type: none"> – Image/reputation – General desire to (perform/not cause harm) – Keeping in harmony with (community/clients/users) – Community wellbeing – Improvement in this area is a key business indicator • No problems with the regulator

Experience of regulation and the regulator

- Which compliance support methods (e.g. publications, advice) were used to help duty holders understand their statutory obligations
 - Satisfaction with these methods
 - Agreement/disagreement that regulator decisions are evidence-based
 - Agreement/disagreement that regulator permit/licence conditions are consistent across industry
 - Agreement/disagreement that permitting decisions were objective and made on complete information
 - Awareness of obligations, and non-compliance, before an inspection
 - Clarity of regulator explanation of non-compliances
 - Satisfaction with the interaction, level of information provided
 - Awareness and satisfaction with broader engagement and consultation opportunities
 - Awareness of the regulator and their regulatory actions (such as inspections, fines and prosecutions).
-

Appendix C: Linkages to the VGR

The following table summarises key requirements for evaluation strategies set out in the VGR, and how this toolkit supports evaluation design and delivery.

Key VGR requirements are to:	This toolkit supports the VGR by providing:
<p>Develop an evaluation strategy linked to your objectives of action, formed around key evaluation questions.</p> <p>Evaluation questions address gaps in knowledge, areas of change or uncertainty, and the most significant attributes of the policy.</p>	<p>Guidance on the scope of a regulatory evaluation compared to other forms of evaluation – to guide your questions and areas of focus.</p> <p>A summary of how to approach regulatory evaluation in different contexts – to allow for situations beyond preparing regulations.</p>
<p>Before implementing the preferred option, set out an appropriate organising logic to answer the evaluation questions, such as a program logic model.</p>	<p>Guidance on planning a regulatory evaluation, with a focus how to build out a program logic model, and what to measure within the program logic, including behaviour change and the supporting actions of regulators and delivery agencies.</p>
<p>Meet principles for regulatory evaluation. In summary, that evaluation: is core to government delivery; improves understanding of the problem; builds an evidence base; improves knowledge for all parties; and recognises uncertainty and complexity.</p>	<p>Considerations for specific kinds of evaluations – including for <i>fees regulations</i>, <i>permissions regulations</i>, and <i>mid-term evaluations</i>.</p>
<p>Meet minimum requirements, e.g. to have a clear evaluation design and logic, well-defined accountabilities, and proportionate investment (including for mid-term evaluations).</p>	<p>Tips including on identifying stakeholders, managing, an evaluation and common challenges. These can help meet minimum requirements by proactively addressing common issues.</p>
<p>Develop an appropriate methodology, tracking a range of measures and not just specific targets, examining cause and effect, drawing on diverse information sources and filling data gaps.</p>	<p>Steps for an evaluation – <i>planning, undertaking, and writing up an evaluation</i>. This is a high-level framework, useful for understanding when to develop and apply your methodology and how to write it up.</p> <p>Guidance on indicators and measures, helping you to identify the data you need.</p>

Glossary

Regulatory evaluation involves a range of concepts and frameworks. Key terms are described below or introduced within this toolkit.

Concept	Description
Behaviour change theory	An explanatory framework for understanding why desired, or undesired behaviours occur, and how to shift awareness, attitude and action to bring about desired behaviours, including compliance with regulations
Drivers of compliance	A framework for examining the factors that increase or decrease compliance, which is used to examine regulator and other activities.
External factors	Things outside the control of the regulation or regulator or government activities but will still have an impact on its success of the regulation
Indicators	The state or level of something, that shows the progress made towards achieving a specific output or outcome
Outcome	The effect of outputs on desired results or objectives
Output	What was produced or accomplished to achieve outcomes
Program logic model	A model that shows how a regulatory intervention works and draws out the relationships between inputs, activities, outputs and outcomes
Proxy measures	An indirect measure of the desired outcome which is itself strongly correlated to that outcome
Social research survey	A research method involving the use of standardised questionnaires or interviews to collect data about people and their preferences