Ontario Energy Board

Staff Report to the Board

on the Proposed Evaluation Framework for Board Policy

December 1, 2011

Table of Contents

Executive Summary

The Ontario Energy Board (the "Board") initiated a process to develop a systematic framework to monitor and evaluate the effectiveness of its policies. The Board believes that an evaluation framework for Board policy could be an important tool to:

- assess whether the objectives of its policies have been met; and

- encourage continuous improvement (as on-going monitoring may reveal issues that can be addressed).

Board staff ("staff") has undertaken research and commissioned expert advice on the development of the proposed evaluation framework. Staff's proposed framework focuses on four areas in the monitoring and evaluation process.

- A **logic model**, which provides a structured approach to link a policy to its intended or expected outcomes/objectives.

- **Impact analysis** (or "impact assessment"), which includes any analysis that supports a decision to initiate or revise a policy and takes place prior to the issuance of the policy.

- **Performance monitoring and reporting**, which includes tracking the progress of a policy on a regular basis so that successes or problems are reported and communicated.

- **Evaluation**, which provides feedback, after the sufficient passage of time, on the success in terms of achieving expected outcomes/objectives.

Staff recommends that the application of the proposed framework should be limited to codes, rules, filing requirements, guidelines, handbooks, and Reports of the Board; and that it should be applied as follows:

- for future policies, an evaluation plan would be built into the policy making process from the outset;

- for existing policies with identified measurable objectives, the framework would be used to evaluate the effectiveness of these policies; and

- for existing policies where objectives have not been identified, elements of the framework may help the Board to report on the impacts of these policies as part of periodic reviews.

Staff also recommends that the proposed framework should be implemented on a pilot basis before being fully adopted by the Board. Staff sees merit in retaining a measure of flexibility to refine the framework as experience warrants.

1 Introduction

The Ontario Energy Board (the "Board") initiated a process to develop a systematic framework to monitor and evaluate the effectiveness of its policies. To date, various approaches have been taken to reviewing Board policy. Typically, when initiating or revising Board policy, in consultation with stakeholders, research is carried out, options are analyzed and recommendations are made as to the best alternative. The proposed evaluation framework set out in this Report outlines a structured approach to policy development including a more systematic approach to evaluation.

The *Ontario Energy Board Act, 1998* (the "Act") has a requirement for a periodic review of the Board's effectiveness. Establishing a systematic approach to monitoring and evaluating the Board's polices would be a fundamental part of this review.

The proposed evaluation framework provides a road map to help:

- develop future policies with clear, measurable and realistic objectives;

- ensure an evaluation process is built into the policy making process from the outset (management buy-in and resources are in place);

- understand and apply the different evaluation methodologies;

- develop principles to guide both the selection of which Board policies should be evaluated, and how they should be evaluated;

- establish reporting and monitoring provisions; and

- guide the development of an evaluation plan for future policies.

Board staff ("staff") recommends that the application of the proposed framework should be limited to codes, rules, filing requirements, guidelines, handbooks, and Reports of the Board; and that it should be applied as follows:

- for future policies, an evaluation plan would be built into the policy making process from the outset;

- for existing policies with identified measurable objectives, the framework would be used to evaluate the effectiveness of these policies; and

- for existing policies where objectives have not been identified, elements of the framework may help the Board to report on the impacts of these policies.

This Report differs from other Staff Reports to the Board in that the report includes both background information on evaluation concepts and staff's recommendation for establishing and implementing a framework. Staff believes that this evaluation framework would help the Board assess whether its policies are achieving the desired outcomes and encourage continuous improvement (as on-going monitoring will provide timely feedback that can be fed into the policy making process).

Organization of this Report

This report is organized as follows: Section 2 defines evaluation and discusses the setting of objectives; Section 3 outlines the proposed evaluation framework – logic model, impact analysis, performance monitoring and reporting, and evaluation; and Section 4 is the conclusion. An evaluation plan template and the advantages and disadvantages of conducting evaluations in-house vs. independent are contained in the Appendices.

2 What is Evaluation?

Evaluation involves an assessment of the strengths and weaknesses of policies, programs and/or projects to improve their effectiveness. The goal of an evaluation is to provide "useful and transparent feedback" for managing policies and to help with the decision-making. It uses systematic data collection and analysis to address questions about how well policies are working, whether they are achieving their objectives and why policies are or are not effective.

Evaluation is a process that can take place before, during and after a policy has been introduced. It is an iterative process that can be applied to all aspects of policy development and implementation. Some jurisdictions describe this iterative process as the following:

At the start of the process ... forward-looking analysis is used to support a decision to initiate a new policy or revise an existing policy. This part of the process is the impact analysis or impact assessment (and it is considered to be ex-ante as this analysis is performed at the outset); and

At the finish of the process ... backward-looking analysis is carried out when a policy has been in the market for a period of time. This part of the process is the evaluation (and it is considered to be ex-post as this analysis is performed after the policy has been implemented).

This iterative process often forms five key elements to a broad policy life cycle (as outlined in Figure 1 below). Specifically,

- **Problem (or Need) Definition –** this stage focuses on problem (or need) definition (what is the problem we are trying to solve with the policy), the scope of the problem, which stakeholders are affected by the problem, will this problem change over time, and what happens if we do nothing.

- **Policy Development** – this stage focuses on setting objectives[1] (what are we trying to achieve, what would constitute a successful outcome, how might our objectives be measured, and how might we monitor our progress in meeting these objectives) and identifying alternative policy options that might meet these objectives. This stage typically involves researching, identifying and describing costs, benefits and risks, and testing these alternative policy options through a formal consultation.

Figure 1: Policy Life Cycle
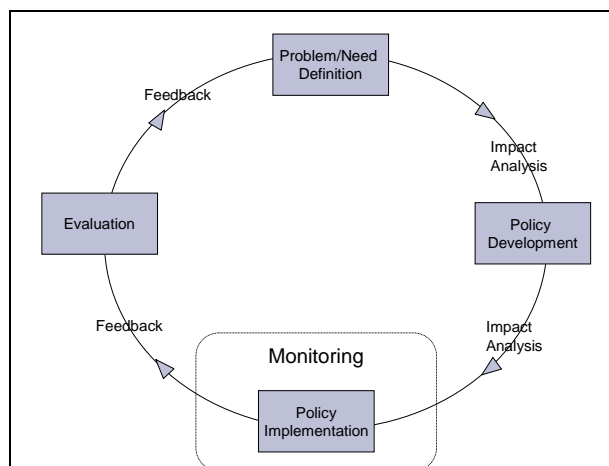
- **Policy Implementation** – this stage focuses on selecting the best option, refining the chosen option and setting out the monitoring and evaluation plan.

- **Monitoring** – this stage focuses on tracking the progress of the policy on a regular basis so that successes or problems are reported and communicated.

- **Evaluation** – this stage focuses on assessing the success or failure of the policy in achieving its objectives and to ensure that the lessons learned are fed back into the policy making process.

Generally, the importance of evaluation, whether ex ante or ex post, is apparent in the jurisdictions reviewed, particularly in the United Kingdom and Canada. These jurisdictions conduct comprehensive but proportionate[2] assessments and evaluations, wherever it is

---

[1] The terms "objectives", "outcomes", and "results" are used interchangeably.
[2] The level of effort applied to an impact assessment and evaluation should be proportional to the size and complexity of the policy, the outcomes at stake, the resources and time available, etc. This is outlined in detail in section 3.2.1.

practical for large policy initiatives and/or projects. These jurisdictions believe that impact analysis and evaluations help them avoid past mistakes by learning from their experiences. Staff notes however that in those evaluations more attention is paid to the up-front analysis (i.e., the impact analysis).

2.1 Guiding Principles

Staff notes that jurisdictions that have similar frameworks have adopted a set of guiding principles that create context and justification for their impact assessment and evaluation activities. In particular, as part of its Performance Measurement and Evaluation Plan ("PMEP") framework, the Treasury Board of Canada Secretariat has established guiding principles that could apply to the Board's impact analysis and evaluation activities. These principles are the following:

- ensure a clear and logical design that ties resources and activities to expected results;

- describe the role and responsibilities of the main players;

- ensure reliable and timely information is available;

- make sound judgments on how to improve performance on an on-going basis (establish the appropriate feedback loops);

- demonstrate accountability and benefit; and

- ensure that information gathered will effectively support an evaluation exercise.

In addition, to assist in the selection of which policies should be assessed and evaluated, staff suggests the following considerations:

- the size and complexity of the policy;

- the significance of the anticipated impact on stakeholders (e.g., customers, distributors, transmitters, generators, etc.);

- the level of risk that the policy might not achieve its expected outcomes;

- the relative priority with reference to other policies;

- interrelation with other proposals and projects; and

- the time and resources available.

2.2   Setting Objectives

Objectives are typically defined as specific changes in behaviour, knowledge, and/or level of functioning that are expected to result from a policy. A key to successful impact analysis and/or evaluation is a set of clear, measurable, and realistic policy objectives. If objectives are unrealistic or are not measurable, the Board will likely not be able to demonstrate that it has been successful even if it has done a good job.

A policy's objectives are the things against which the Board will be evaluating itself and against which stakeholders will be assessing the Board. Objectives therefore need to be SMART, that is:

- **Specific** – it should be clear what the Board is aiming to achieve;

- **Measurable** – there should be a clear and transparent measure of success;

- **Achievable** – the target should be stretching but it must be achievable and there should be some evidence that demonstrates what is possible (e.g., benchmarking with similar organizations);

- **Relevant** – the target should reflect what the Board is trying to achieve and not simply what is easily measurable; and

- **Time-bound** – it should be clear by when the outcome should be delivered.

It is staff's view that the set of questions listed below will assist in the development of suitable objectives. These questions are based on the United Kingdom's Green Book[3] and are the following:

- What are we trying to achieve? What would constitute a successful outcome or set of outcomes?

- Have similar objectives been set in other contexts that could be adapted?

- Are our objectives in relation to the policy consistent with the strategic objectives of the organization?

- Are our objectives defined to reflect outcomes (e.g., transparency, customer expectations are being met) rather than the outputs (e.g., issuance of staff papers and rules), which will be the focus of particular policies?

- How might our objectives / outcomes be measured?

- Are our objectives defined in such a way that progress toward meeting them can be monitored?

- What factors are critical to success?

- What SMART objectives can we then set? What objectives do we need to meet?

It should be noted that not all policies can be evaluated. A policy must meet the central requirements of evaluation for a successful evaluation to occur. These central requirements are that the initiative and the target population are clear and identifiable; that the outcomes are clear, specific and measurable; and that an appropriate evaluation design can be implemented.

The Board has a set of clearly defined statutory objectives that are prescribed in the Act. The Board's work (including policy work) must always be consistent with and in furtherance of the fulfilment of these statutory objectives. For the purpose of impact assessment and evaluation, objectives will be identified that are appropriate for a particular policy. These

---

[3] The Green Book, Appraisal and Evaluation in Central Government, TM Treasury, United Kingdom, 2003, page 14.

more narrowly focused objectives that relate directly to the policy itself must always be consistent with the Board's on-going statutory objectives. Section 2.2.1 below describes sample objectives for Board policy.

2.2.1   Sample Objectives for Board Policy

Some objectives might take several years to achieve while others might be achieved shortly after a policy has been implemented. Therefore, objectives are typically divided into short-, medium- and long-term objectives (less than one year, one to three years, and greater than three years, respectively).

Below are a number of sample objectives:

1. **Transparency** – Regulatory expectations are understood leading to improved predictability and market confidence.

2. **Consistency and Clarity** – Regulatory instruments are clear and consistently interpreted leading to fewer disputes and less need for Board support.

3. **Compliance with Policy Instruments** – Guidance and/or requirements set out in regulatory policy instruments are being followed.

4. **Customer Expectations are being met** – Customers and system users' expectations with respect to electricity and gas system services are being met.

5. **Responsiveness** – There is flexibility in responding to changing market conditions.

Staff notes that these objectives (i.e., outcomes) are not mutually exclusive. For example, a transparent process with easily understood regulatory expectations will better ensure that the regulatory instruments are clear and consistently intrepretated. This, in turn, will better ensure that the regulatory instruments are being followed and that the Board has the necessary information to review and approve applications on a timely basis.

Table 1 below outlines in detail the sample outcomes (or objectives) that staff proposes.

Table 1: Sample Outcomes / Objectives

Outcome Category	Short- and Medium-Term Outcomes (< 1 year to 3 years)	Long-Term Outcomes (> 3 years)
Transparency – Regulatory expectations are understood leading to improved predictability and market confidence	Stakeholders demonstrate increased satisfaction with market conditions, relationships, and practices	• Satisfaction with market conditions, relationships and practices • Confidence within the energy sector
Consistency and Clarity – Regulatory instruments are clear and consistently interpreted leading to fewer disputes and less need for Board support	Feedback that regulated / licensed entities interpret regulatory instruments clearly and consistently	• Market behaviour reflects understanding of regulatory instruments • Confidence within the energy sector
Compliance with Policy Instruments – Guidance and/or requirements set out in regulatory policy instruments are being followed	Regulated / licensed entities demonstrate increased following with published policy instruments	• Published policy instruments are being followed • Confidence within the energy sector
Customer Expectations are being met – Both customers and system users' expectations with respect to electricity and gas system services are being met	Customers and system users provide positive feedback about energy services	• Board determined performance standards are being met • Confidence within the energy sector
Responsiveness – There is flexibility in responding to changing market conditions	Increasing positive (formal) feedback about Board responsiveness and processes	• Continued positive (formal) feedback about Board responsiveness and processes

The list of sample outcomes / objectives (as outlined above) has been prepared on a generic basis. In practice, specific outcomes would likely be set for a particular Board policy. Staff notes that when setting specific outcomes in the short-, medium- or long-term, the needs of customers and regulated entities will always be considered. Further, as discussed in section 2.2, these specific outcomes / objectives must always be consistent with the Board's on-going statutory objectives.

In addition, staff is of the view that when developing outcomes for a specific policy, it would be inappropriate for the Board to be held accountable for an activity for which it is not

responsible (e.g., it is a distributor's or another agency's responsibility to achieve that outcome). Also, staff notes that taking on the distributor's accountability may introduce an incentive to "manage" the distributor beyond the stated statutory obligations.

3 The Proposed Monitoring and Evaluation Framework

Figure 2: Proposed Monitoring and Evaluation Framework
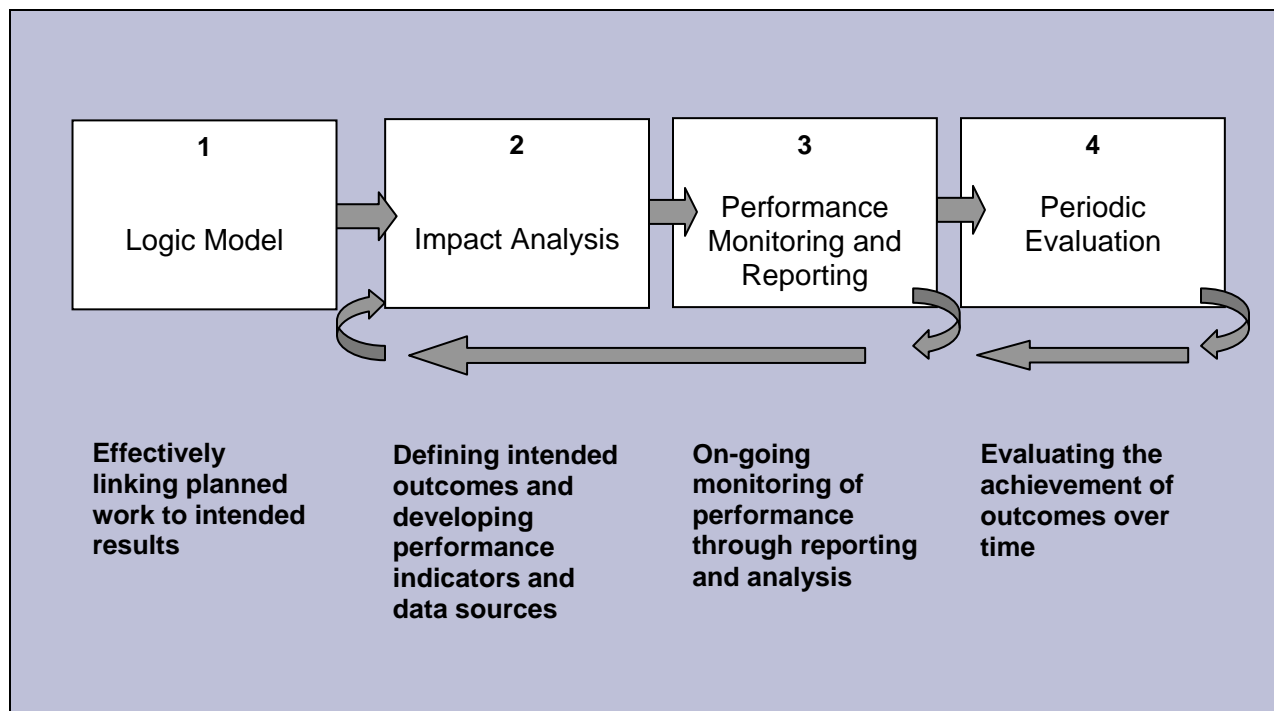
The proposed framework is structured on **four areas** in the monitoring and evaluation process:

- A **logic model,** which is a representation, often portrayed visually, of the linkages from planned activities to intended results[4]. It is a planning tool to ensure that activities and outputs are focused on achieving the desired results.

- **Impact analysis** (or "impact assessment"), which includes any analysis that supports a decision to initiate or revise a policy and takes place prior to the issuance of a policy. It is a structured approach to think through the consequences of possible and actual policy implementation: from the early stages of identifying a policy challenge, through the development of policy options, public consultation of those

---

[4] The terms "results", "outcomes", and "objectives" are used interchangeably.

options, final decision-making, and on to the review of the preferred option for implementation.

- **Performance monitoring and reporting,** which includes tracking the progress of a policy on a regular basis so successes or problems are reported and communicated. It identifies appropriate performance indicators; puts in place a plan for data collection and retention; establishes provisions for monitoring and analyzing the results of a policy; supports effective and relevant reporting on the policy's progress; and ensures that the information gathered will effectively support evaluation activities.

- **Evaluation,** which provides feedback, after the sufficient passage of time, on the success in terms of achieving expected outcomes and the relevance in terms of on-going alignment with Board priorities.

These four areas are the foundation for a comprehensive and effective approach to establishing outcomes, performance indicators, and carrying out performance monitoring and evaluations. Each of these areas is discussed below.

3.1   Logic Model

Increasingly, governments are adopting logic models to effectively represent the relationship between inputs and resources, planned activities and outputs, and desired outcomes. Specifically, a logic model is used to summarize, often in visual form, the linkages between activities and outputs, and between outputs and measurable outcomes.

As per Figure 3, a logic model describes the sequence of activities thought to bring about change and how these activities are linked to the results the policy is expected to achieve.

Figure 3: Linking Inputs to Outcomes

Planned Work		Intended Results	
Inputs / Resources	Activities	Outputs	Outcomes
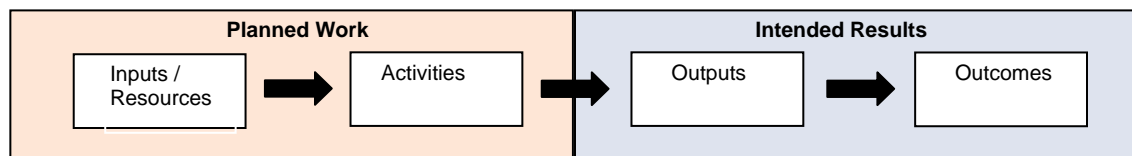
For example, a legislative change triggers the need for a new code (the "input"). To develop a new code, the Board meets with stakeholders, prepares a Staff Discussion Paper and then issues a Notice of Proposal (the "activity"). The "output" is the new code. The "outcomes" of this new code may be transparency and non-discriminatory access leading to fewer disputes and less need for Board support.

Table 2 below summarizes a generic logic model that outlines:

- some of the drivers or inputs that trigger policy activity within the Board;

- the Board's activities such as the processes, tools, events and actions that make up our work;

- the outputs which are the direct products of the Board's policy activities; and

- the outcomes which are the specific changes in behaviour, knowledge, and/or level of functioning that are expected to result from policy activities and outputs.

Table 2 - Examples of Inputs, Activities, Outputs and Outcomes

Inputs – triggers or drivers of policy work	**Activities** – processes and actions that make up the work	**Outputs –** the direct products of policy activities	**Outcomes –** changes that are expected to result from activities and output
• Ontario Legislation (statutes and regulations) • Ontario MOUs and Directives • Board Decisions • Monitoring Information • Stakeholder Feedback	• Develop or amend policy instruments (e.g., rules, codes, policy guidance and filing requirements / guidelines) for Board approval • Provide policy advice (e.g., Staff Reports to the Board, briefing notes) • Analysis of stakeholder and market information (e.g., hold stakeholder meetings / conferences)	• Rule, code, filing requirement / guideline • Board Decisions • Staff Report of the Board • Staff Discussion Papers • Reports of the Board • Staff Research Papers	• Transparency • Consistency and Clarity • Compliance with policy instruments • Customer expectations are being met • Responsiveness

Staff suggests that a logic model can be used as a planning tool to ensure that the Board's activities and outputs are focused on achieving desired results. In particular, a logic model can play an important role in the policy development process by providing a structured approach that links a proposed policy to its intended outcomes. It facilitates the planning and communication of the policy's outcomes, and as a result it creates an explicit understanding of the challenges ahead, the resources available, and the timetable in which the Board is expected to achieve its outcome.

A logic model can be tailored to any specific policy initiative. Staff notes that these linkages are currently mapped out, at a high level, in the Board's Terms of Reference. Staff suggests that a more structured approach to the Board's Terms of Reference process will ensure that a proposed policy is linked to its intended outcomes.

3.2   Impact Analysis

Impact analysis ("ex ante impact analysis" or "impact assessment") refers to any forward-looking analysis used to support a decision to initiate or revise a policy. Impact analysis is performed at the outset and is a process to help think through the consequences of possible and actual policy implementation: from the early stages of identifying a policy challenge, through the development of policy options, public consultation of those options, final decision-making, and on to the review of the preferred option's implementation[5].

Staff notes that various approaches have been taken to impact assessments. Typically, when revising or initiating policy, in consultation with stakeholders, research is carried out, options are analyzed and recommendations are made on the best alternative. The proposed framework outlines a structured approach to policy development including a more systematic approach to assessment.

Therefore, staff recommends that **an impact analysis be conducted in cross-organizational teams** to ensure policies are clear, well communicated, understood and followed. The definition of outcomes, performance indicators and data sources will benefit from a team approach. This will ensure that performance indicators and data are aligned with reporting requirements, compliance activities, complaints-handling, communications efforts and other activities that support effective policies.

Staff also recommends that **an impact analysis answer a structured series of questions** to ensure that the problem is well defined, the alternatives have been considered, and the anticipated outcomes of a policy are identified along with clear performance measures and data sources.

---

[5] UK, Impact Assessment (IA) Toolkit – Version 1.0

Below is a series of questions that reflect a common analytical approach found in other jurisdictions[6].

- **Define the Problem (or Need)**: What is the problem (or need) this Board policy is trying to solve?

  Ideally, the statement of the problem / need should identify its magnitude (frequency, level of risk, etc.) and the incentives or behaviours that contribute to or cause the problem / need. It should examine whether the issues will decline or decrease in importance if there is no policy. Key questions that could assist with problem / need definition include:

  - What is the status quo (including current regulatory environment, markets etc.)?
  - If there is existing regulation, is there a concern with regulatory design or implementation (including compliance)?
  - What happens if the Board does nothing? Will it deteriorate in the absence of regulatory action (of some nature)? Probability? Severity of harm?
  - Who is affected and what is the scope of the problem / need? Will this change over time?
  - What are the root causes of the problem / need in the absence of any regulatory action? Market failures? Information failures?
  - Where are the incentives to change behaviour (e.g., self-protection, reputation, financial sensitivity, competitiveness)?

- **Identify the Objectives:** What is the Board trying to achieve with this policy?

  As per section 2.2, the objectives should be specific, measurable, achievable, relevant and time-bound.

- **Identify Policy Alternatives:** What alternatives have been considered and what is the best approach to take and why?

---

[6] The jurisdictions examined are the United Kingdom, New Zealand and Canada.

This should involve researching, identifying and describing costs, benefits and risks, and testing these alternatives through formal consultations. If possible, it should also identify the potential unintended consequences (e.g., unintended changes in stakeholder behaviour) that may occur and how these consequences should be monitored.

- **Determine a Performance Monitoring and Evaluation Plan:** If a monitoring and evaluation plan is deemed to be appropriate, how will the Board know that the policy is doing what it is intended to do in the short and longer terms?

  The plan should include: a) the scope of the evaluation (i.e., what is to be evaluated and why), b) the performance indicators to measure the policy's progress, c) staff and other resources necessary to conduct the on-going monitoring and the expected evaluation, d) the proposed approach to collecting the evidence to monitor and evaluate (i.e., data collection, reporting, etc.), and e) the timing and cost. It is the view of staff that the plan should be prepared in consultation with stakeholders.

  At this stage the evaluation plan would outline a general approach and this approach would be refined with more specifics (e.g., the detailed scope of the evaluation, timing, resources, data gaps, etc.) closer to the actual evaluation. Please refer to Appendix A for greater detail on developing an evaluation plan.

- **Identify Implementation and Compliance Considerations**: How would the Board implement the policy? What are the anticipated implementation costs? How can compliance with mandatory requirements be assured? How do we expect to monitor compliance or monitor whether a policy is being followed (e.g., complaints, self-reporting, audits)?

3.2.1   Level of Effort

Staff notes that the time and resources that could be devoted to impact assessment could be considerable. Staff suggests that the key to controlling costs is proportionality: effort should be proportional to anticipated impact. In other words, the analytical effort and costs involved should be proportional to the effect or importance of the proposed policy and its impacts.

Staff notes that some jurisdictions[7] offer general guidance to help organizations gauge the level of effort to commit to a particular impact assessment exercise. These jurisdictions suggest that effort applied to an impact assessment should be proportional to:

- the size and complexity of the policy;

- the outcomes at stake;

- importance to the industry;

- relative priority with reference to other policies and projects;

- interrelation with other policies and projects;

- resources; and

- timing.

3.3   Performance Monitoring and Reporting

On-going monitoring provides timely information that can be fed into the policy-making process. Monitoring is based on performance indicators (or performance measures) set during the planning phases of work (i.e., the impact analysis phase). Monitoring helps to keep the work on track, and can let the Board know when things are going wrong. When

---

[7] The jurisdictions examined are the United Kingdom, New Zealand and Canada.

done properly, it is an invaluable tool for good management, and it provides a useful base for evaluation.

To effectively monitor performance, it is necessary to:

- identify the appropriate performance indicators, which are qualitative or quantitative means of measuring an outcome, with the intention of gauging the performance of a policy;

- put in place a strategy for data collection and retention;

- monitor and analyze the results of the policy;

- support reporting of the policy's progress; and

- ensure that the information gathered will support evaluation activities.

3.3.1 Establishing Performance Indicators

Performance indicators should answer the question – How will the Board know whether the anticipated outcomes have been achieved?

The logic model in Figure 3 forms the foundation for performance monitoring. Performance monitoring is the link between the theoretical relationships in the logic model and the practical assessment of the outcomes appearing in that model. Specifically, performance indicators are the quantifiable measures used to determine whether the policy is achieving the intended outcomes as outlined in the logic model.

Table 3 below is a check list to help develop performance indicators.

	Check List	Explanation
1	**Refine outcome**	You may wish to revise or refine your outcomes and expected results as you prepare to select performance indicators.
2	Benefit from **already established indicators** and outcomes from other policies	
3	**Select at least one indicator (but no more than three)** for each outcome identified in the logic model	Ensure a mix of quantitative and qualitative indicators. Use positive indicators where possible. Quantitative performance indicators require a number and a unit (e.g., "number of complaints about XX received"). Qualitative performance indicators are more challenging to articulate and need to be expressed in a more descriptive form (e.g., "understanding of XX") and therefore require a rating scale that will allow for comparisons over time such as "strong, acceptable or needs improvement".
4	**Estimate indicator range or target** for each selected indicator	To measure and monitor the progress being made towards the achievement of outcomes requires the establishment of a target to reach (usually expressed as a number or a percentage) within a certain period of time. Ideally, baseline data should be established (i.e., information collected before a policy begins). If there is insufficient or no baseline data, explicit timelines for the targets are required. Consider other sources for comparative performance information (e.g., benchmark similar policies in other jurisdictions).
5	**Consult with stakeholders** to identify performance indicators	
6	**Identify data sources** (existing and required) and **challenges**	There are a number of existing data sources: • Existing databases such as RRRs, MPEs, CIDRA, CMIS, and annual stakeholder survey) • Statistics Canada, economic reports, third party studies Need to identify any challenges such as a data management system may not be in place yet.

	Check List	Explanation
7	**Identify frequency for collection** and staff responsible	In order to monitor the outcomes of a particular policy, the data must be collected and made available at appropriate intervals (e.g., quarterly or annually) to support monitoring needs. It is also essential that the information collected is shared with the staff responsible for monitoring and evaluation.
8	**Validate the performance indicators** selected to confirm that they meet the following criteria: ✔ Relevant ✔ Avoid perverse incentives ✔ Attributable ✔ Well-defined ✔ Timely ✔ Reliable ✔ Comparable ✔ Verifiable	**Relevant –** to what the organization is aiming to achieve; **Avoid perverse incentives** – not encourage unwanted or wasteful behaviour; **Attributable** – the activity measured must be capable of being influenced by actions which can be attributed to the organization; and it should be clear where accountability lies; **Well-defined** – with a clear, unambiguous definition so that data will be collected consistently, and the measure is easy to understand and use; **Timely** – producing data regularly enough to track progress and, and quickly enough for the data to still be useful; **Reliable** – accurate enough for its intended use, and responsive to change; **Comparable** – **w**ith either past periods or similar programmes elsewhere; and **Verifiable** – with clear documentation behind it, so that the processes which produce the measure can be validated[8].
9	**Consider the cost of collecting data** when selecting performance indicators.	It is always important to ensure that the cost of gathering a specific indicator does not exceed the benefit of having that information. Need to determine how the required information will be used.

3.3.2 Performance Monitoring

On-going monitoring will provide feedback on the selected performance indicators and the Board will be able to track the policies progress toward achieving its anticipated outcomes. This on-going monitoring will provide a series of observations. These observations will typically fall into "bands" or ranges for the indicators being tracked. It may be appropriate to

[8] Choosing the Right Fabric, A Framework for Performance Information, March 2001, written jointly by National Audit Office, Audit Commission, Cabinet Office, Office for National Statistics and HM Treasury.

have three bands within which performance indicators would fall: normal, watch, and outside the acceptable tolerances.

Once an indicator goes to "watch", it gets a higher level of attention. This higher level of attention should include a time element. If an indicator stays in the normal band 95% of the time, but has an occasional blip in the "watch" area, this movement could still be considered to be in the normal range. However, when the indicator stays in the "watch" area for a defined period of time, this should raise a caution flag for the Board. When indicators stay in the band higher than the "watch" area for a defined period of time, then an evaluation may be warranted. In this situation, a periodic evaluation may be more appropriate because the performance monitoring is no longer suffice in providing meaningful feedback on the impacts of that particular policy.

Performance indicators can, for a variety of reasons, suddenly or incrementally rise above or drop below the baseline range in the normal "band". This could be precipitated by events such as significant federal or provincial announcements, major changes to energy market or other factors.

3.3.3 Data Collection and Analysis

As discussed previously, to effectively monitor performance, it is necessary to: a) put in place a strategy for data collection and retention, b) monitor and analyze the results of a policy, and c) support reporting of the policy's progress. To accomplish these tasks, jurisdictions with well-established monitoring and evaluation capabilities typically have the following four processes in place:

- a procedure for the collection of baseline data;
- a data collection procedure;
- a data analysis procedure; and
- a reporting procedure.

Staff notes that the Board already monitors the energy market based on market information provided by the utilities and stakeholders, and information gathered by the Board through the Reporting and Record-keeping Requirements ("RRRs"), Market Participant Enquires ("MPEs") and the complaints processes such as the Consumer Interaction and Dispute Resolution Application ("CIDRA") and the Compliance Management Information System ("CMIS").

To support formal monitoring and evaluating activities outlined in the Report, it is staff's view that processes will be required to accommodate new data requirements, on-going analysis of the performance indicators will be needed, and changes to the frequency and types of reporting may need to be implemented. This may impact the resources responsible for this work.

Each of these processes is discussed below.

Baseline Data Collection

Baseline information is the data provided immediately before or at the beginning of a period to measure. Baseline data is used to learn about recent performance and to measure progress from that point forward. It is staff's view that there will be a need to define and identify baseline data for future policies to allow for monitoring and evaluation activities, if appropriate.

Data Collection

Staff notes that there will be a need to identify the sources of the required data, design sampling procedures if and when appropriate, develop data collection instruments, determine the frequency for data collection, and collect the actual data requirements. This will require a clear description of the processes to accommodate the necessary data sources, and approaches to data collection and reporting requirements along with the staff responsible for these tasks.

Data Analysis

Staff suggests that there will be a need to establish clear processes for how data is to be analysed and packaged along with the staff responsible for these tasks to successfully monitor and evaluate its policy. For example, as discussed in section 3.3.2, on-going monitoring will establish feedback on the performance indicators. The feedback will require analysis and the results of this analysis will indicate whether the performance indicators are falling within the normal, watch, or outside the acceptable tolerances range.

Reporting Procedures

Staff is of the view that there may be a need to further refine its reporting process to effectively monitor its policy outcomes. Elements of an effective reporting procedure include any data reporting formats or templates, determining who needs the reports, determining the frequency of reporting, and determining that information is communicated to the appropriate internal and external stakeholders in a way that ensures the reporting meets the needs of these stakeholders.

For any particular Board policy, factors to be considered should include the need for which the policy is intended and the format, frequency and intended use of the data or information to be reported. The specifics will vary according to the type of policy and stakeholders.

3.3.4 Sample Performance Indicators

As noted in section 2.2.1, staff suggests five sample outcomes (i.e., objectives) – transparency, consistency and clarity, compliance with policy instrument, customer expectations are being met, and responsiveness. For each of these five sample outcomes, table 4 below provides a list of potential performance indicators that may be used to indicate whether the Board has achieved its objectives.

Staff notes that the potential performance indicators (provided below) are based on existing Board databases and have been prepared on a generic basis. In practice, specific performance indicators will likely be set for a particular policy. It is essential that the performance indicators reflect the policy's intended or anticipated outcomes.

Table 4: Sample Performance Indicators

Sample Outcome	Sample Performance Indicators
Transparency – Regulatory expectations are understood leading to improved predictability and market confidence	• Fewer requests for clarification of policy (e.g., reduction in MPEs) • Fewer compliance/informational bulletins issued • Evidence of fewer complaints and disputes (e.g., reduction in CIDRA and CMIS)
Consistency and Clarity – Regulatory instruments are clear and consistently interpreted leading to fewer disputes and the need for Board support	• Fewer compliance/informational bulletins issued • Fewer requests for clarification of policy (e.g., reduction in MPEs, received by the Board) • Evidence of fewer complaints and disputes (market data demonstrates changes in regulated / licensed entities behaviour)
Compliance with Policy Instruments – Guidance and/or requirements set out in regulatory policy instruments are being followed	• Evidence of fewer departures from policy instrument (e.g., none or few exemption applications, non-compliance to codes / rules, and not conforming to filing guidelines) • Market data demonstrates changes in regulated /licensed entities behaviour (e.g., reduction in numbers of: 1) complaints about fair market practices and 2) disputes between parties (e.g., regulated entities and their customers) • Evidence that stakeholder opinion regarding fairness of codes, rules and reports is maintained (based on annual stakeholder survey)
Customer Expectations are being met – Customers' and system users' expectations with respect to electricity and gas system services are being met	Customers: • Reduction in outages, maintain Board performance standards (e.g., SQIs), etc. • Evidence of fewer complaints and disputes • Conditions for connections Distributors: • Maintain financial ratios – liquidity ratios (current ratio); leverage ratios (debt ratio, debt to equity ratio and interest coverage ratio); profitability ratios (return on assets and return on equity) • Maintain Board approved Service Quality Requirements • Unitized statistics (average revenue from distributor, average cost of power, net fixed assets per customer, capital spending, etc.) • Safety (comply with legal requirements)

Sample Outcome	Sample Performance Indicators

Responsiveness – There is flexibility in responding to changing market conditions	• Consultation, format and pace, is appropriate for policy being developed. Process designed for:
	
	   o maximum participation
	
	   o wide range of stakeholders (representing different groups)
	
	   o different levels and format of participation (working group vs. stakeholder meetings vs. stakeholder conference)
	
	• Board determined performance standards are being met (e.g., timely approvals of applications, responding to market information such as 2010 Natural Gas Review, etc.)

3.4 Evaluation

There are many evaluation types or approaches. No one approach is best for all situations as each approach has its strengths and weaknesses. The best approach varies according to factors such as the intent of the evaluation, the nature of the stakeholders, and the available resources. In practice, two or more approaches are typically used when conducting an evaluation.

3.4.1 Types of Evaluations

Formative and Summative Evaluations

Two types of evaluation that are commonly used in government are formative and summative evaluation. **Formative evaluation**, also known as *process evaluation*, asks how the policy is operating, whether it is being implemented the way it was planned, and whether problems in implementation have emerged. It is carried out in the early stages of a policy and focuses on the processes (i.e., the development, implementation and delivery) underlying a policy's success or failure. In particular, formative evaluation reviews policy development and examines policy activities to assess whether the policy is being offered in the way it was intended and to identify areas where project administration and delivery can

be improved. This often involves addressing questions such as for whom a policy has worked (or not worked), and why.

A **summative evaluation** or *impact evaluation* examines the extent to which a policy has achieved the outcomes it set at the outset. A summative evaluation asks questions such as: What impact, if any, does a policy have in terms of specific outcomes for different groups of people? It seeks to provide estimates of the effects of a policy either in terms of what was expected of it at the outset, or compared with some other policy initiative, or with doing nothing at all.

Both summative and formative evaluations use **q*ualitative and quantitative methods*** for gathering data to support evaluation activities. ***Qualitative*** research is best suited to answering questions about how many stakeholders did something or thought about something. Qualitative research is particularly important for formative evaluation which is limited entirely to a focus on a specific context. Qualitative research uses a range of methods such as:

- Focus groups

- In-depth interviews

- Open-ended survey questions

- Forums/discussion groups

On the other hand, ***quantitative*** research uses methods adopted from the physical sciences that are designed to ensure objectivity, generalizability and reliability. These techniques cover the ways research participants are selected randomly from the study population in an unbiased manner, the standardized questionnaire they receive and the statistical methods used to test predetermined hypotheses regarding the relationships between specific variables. Quantitative measures are often most appropriate for conducting evaluations comparing outcomes with baseline data.

Quantitative research uses a range of methods such as:

- Internet, mail or telephone survey

- Process tracking forms/records

- Analysis of large datasets

- Direct measures of service reliability metrics (collected directly within the distributor such as SAIFI, etc.)

3.4.2 When to Use Formative Evaluation?

It is staff's view that there will be **less need** for formative evaluations. On-going monitoring should be able to provide sufficient feedback to determine if there are indications that a policy is being offered in the way it was intended and to identify areas where project administration and delivery can be improved. Formative evaluations are most often applied for large policies to ensure that implementation was appropriate and to make any "fine tuning" in response to formative evaluation findings.

Staff notes that the Board already has processes and existing structures that can provide feedback to support formative evaluation questions about design and implementation of policies. For example, the Board already assesses validity, reliability and accuracy of data and feedback from such analyses can provide answers to support a formative evaluation process. More specifically, an audit could explore the development (i.e., formation) and implementation of a policy. It would seek information on the processes underlying a policy's success or failure.

Staff notes that audits could be undertaken by the Board or through independent third parties contracted by the Board. Staff recommends that given the costs associated with external auditing, a more efficient use of resources, may be accomplished internally by the Board's Audit group on a sampling basis. The sample could be rotated over time for different policy initiatives.

3.4.3   When to Use Summative Evaluation?

Staff suggests that a summative evaluation is more appropriate for providing feedback on whether desired outcomes are being achieved. Staff is of the view that summative evaluations should generally be considered for major initiatives after an appropriate lapse of time (e.g., after three years to be consistent with the Board's planning cycle) to ensure that there will be evidence of outcomes having been achieved.

Summative evaluations are designed to gather deeper information, through a variety of methodologies about the policy's outcomes and to identify challenges that might be impeding the achievement of those intended outcomes. Such studies often require outside consultants and can be resource-intensive.

A summative evaluation would also include a more thorough assessment of the quality and reliability of the data being used to monitor outcomes (as compared to a formative evaluation) and could help to determine whether new indicators might be required. It would also potentially include new qualitative and quantitative data sources to gain a deeper understanding of policy impacts. The measurement of outcomes after three years should provide more mature feedback to support decisions to modify policies or indicate a need to replace them with new measures.

3.4.4   Other Considerations

The timing of an evaluation would normally be established in the planning that occurs at the impact assessment stage. Staff notes that jurisdictions that conduct evaluations typically prioritize their evaluation activities depending on the size and complexity of the policy, the significance of the anticipated impact on stakeholders, the relative priority with reference to other policies, and the time and resources available.

Furthermore, outside factors, such as Ontario government priorities and changes in North American energy markets could influence Board priorities which in turn could affect the

necessity and timing of evaluations. As well, there could be potential triggers for an evaluation that result from performance monitoring (i.e., observations of indicators outside the normal expected range [often called "outliers"]) and the level of risk impacting the policy's ability to achieve its expected outcomes.

Below are additional criteria to help prioritize evaluation decisions:

- **Intrusiveness** – how much will the proposed evaluation intrude on the stakeholders?

- **Data Credibility** – how reliable is the information available to support an evaluation? Are there data gaps?

- **Cost-effectiveness** – can the data be collected in a cost-effective manner?

- **Proportionality** – how large is the expected impact of the measure? What is the potential value of engaging stakeholders?

Where an evaluation is deemed to be necessary, the following principles should be considered in planning:

- the timelines and scope of the proposed evaluation (i.e., issues and extent of coverage);

- a completed evaluation plan should include a list of questions to address the core evaluation issues and proposed approaches to collecting the evidence to address those issues;

- costs and resources required for the evaluation; and,

- if appropriate, the identification of any data gaps.

See Appendix A for a discussion on how to construct an evaluation plan.

To develop a list of questions to address the core evaluation issues and proposed approaches (as per the principles listed above), questions of relevance, impact and effectiveness should be considered for any particular evaluation exercise. Therefore, when planning an evaluation, it is necessary to determine which questions to prioritize. For

example, on-going relevance of a policy may be of greater or lesser importance than the impacts or cost-effectiveness, and choices can be made as to where the focus of the evaluation should be.

Typically, an evaluation plan should include concrete evaluation questions to address relevance and performance.

Relevance:

- Issue #1: Continued need for the policy: the extent to which it continues to address a demonstrable need and is responsive to the needs of Ontario energy markets;

- Issue #2: Alignment with government priorities: evidence of strong links between the policy's objectives and (i) Ontario government priorities and (ii) the Board's business plan;

Performance (effectiveness, efficiency and economy):

- Issue #3: Achievement of expected outcomes: an assessment of progress toward expected outcomes with reference to indicators, impact on targeted stakeholders, and policy design, including the linkage and contribution of outputs to outcomes; and

- Issue # 4: Demonstration of efficiency and economy: an assessment of resource utilization in relation to achieving expected outcomes.

Furthermore, as with an impact assessment, a series of questions that reflect a common analytical approach for evaluations was found in other jurisdictions[9]. The difference is that evaluation questions look backward at what happened as opposed to impact assessment

---

[9] The jurisdictions examined are the United Kingdom, New Zealand and Canada.

questions which look forward at what is expected to happen. The following questions may also assist when planning an evaluation:

- Is our policy working as it was expected to work?

- Have any risks been identified that have interfered with the ability of the policy to achieve its expected results (e.g., any events that could lead to an impact on outcomes and/or achieving objectives)? Is there sufficient time to mitigate the risk(s)? Are there specific risks related to performance measurement, data collection and reporting?

- What are the impacts of the policy? Were these impacts expected, taking into account the earlier analysis of proposed impacts?

- Are any unexpected impacts positive or negative? Quantitative or qualitative?

- How have stakeholder views evolved over the course of policy implementation?

- What segments of the population have been affected by the policy?

- How well was the target segment identified?

- If appropriate, how has behaviour changed?

- Is the policy still needed in order to obtain this result? What would be the effects if it were cancelled or rescinded?

3.4.5 Data Collection Methodologies for Evaluations

There is a range of methodological approaches to support evaluation activities. Staff suggests that more than one method of data collection (discussed below) should be used because this will strengthen the validity of results. Applying only one method would lack the depth and breadth needed to credibly assess the impact of Board activities and outputs. An important principle of evaluation is that multiple lines of evidence provide greater credibility to the findings.

When considering evaluation methodologies, some factors should be analyzed to ascertain:

- the level of intrusiveness of the Board on the source providing the data;

- the credibility (likelihood of collecting relevant data with an acceptable level of reliability and validity for data to be considered credible) of the methodology, given the indicator recommended;

- the cost of executing the methodology (balanced with the acquisition of valid and reliable data); and

- the potential value of engaging stakeholders associated with that data source to participate in an evaluation as a means of improving/enhancing transparency and relationships with stakeholders.

Using multiple data collection methodologies, qualitative and quantitative data can be collected in order to answer specific set of questions. Typically, there are seven different types of data collection methodology to support evaluation needs:

- document and data review (e.g., analysis of large datasets, process tracking forms/records);

- literature review;

- internet, mail or telephone surveys;

- interviews or forums/discussion groups;

- comparative studies (e.g., cohort studies, case-control studies and experimental studies);

- expert panels; and

- analysis of cost-effectiveness.

Some of the available methodological approaches are more applicable to summative evaluations, while other approaches such as the qualitative approaches (that take place earlier in the policy process) are more applicable to formative evaluations. For example, formative evaluations are likely to use focus groups, in-depth interviews, and stakeholder analysis that are capable of providing feedback on the efficiency of a particular policy

initiative and the effectiveness of the process by which it is being managed. Typically, in a formative evaluation, the evaluator`s role becomes one of a facilitator who engages stakeholders in a form of evaluative inquiry. Staff proposes that these evaluations could be carried out in-house.

In a summative evaluation there will be more concerns with gathering and analyzing data and communicating evaluation findings. For example, summative evaluations will use comparative studies and data review (to compare outcomes with baseline data), and standardized questionnaires or surveys where stakeholders are selected randomly from the study population in an unbiased manner. In some cases, the ideal data collection method for summative evaluation needs to be balanced against the expected credibility and value of the data, the intrusiveness of a methodology, cost considerations, and other intangible factors such as the value of stakeholder consultation and engagement. This may involve making choices and trade-offs when considering options. Staff notes that such evaluations often require outside consultants and can be resource-intensive.

Table 5 below provides a summary of the information needs required for evaluation activities and the different methods for collecting the data necessary to perform the evaluation.

Table 5: Evaluation Needs and Data Collection Methods[10]

Data Collection Methodology	When to Use
External Documentation: use of data collected by other institutions or agencies	• Need information about a specific context • Need historical information • When comparing policy data to comparable data
Internal Documentation and Administrative Data: data collected for management purposes	• Need information on administrative practices, implementation, and key stakeholders

---

[10] Evaluation Guidebook for Small Agencies, Treasury Board Secretariat, Canada, 2004.

Data Collection Methodology	When to Use
Literature Review: review of past research and evaluation on a particular topic	• To identify additional evaluation questions or issues and methodologies • Need information on a specific issue • Need information about comparable policies, best practices
Interviews: a discussion covering a list of topics or specific questions, undertaken to gather information or views from an expert, stakeholder, and/or client; can be conducted face to face or by phone	• Complex subject matter • Sensitive subject matter (in-person interviews) • Flexible, in-depth approach • Smaller populations
Focus groups: a group of people brought together to discuss a certain issue guided by a facilitator who notes the interaction and results of the discussion	• Depth of understanding required • Weighted opinions • Testing policy ideas • Limited number of issues to examine • Where interaction of participants may stimulate richer responses (people consider their own views in the context of others)
Case studies: a way of collecting and organizing information on people, institutions, events, and beliefs pertaining to an individual situation	• When detailed information about a program is required • To explore the consequences of a program • To add sensitivity to the context in which the program actions are taken • To identify relevant intervening variables
Questionnaire or Survey: a paper or electronic list of questions designed to collect information from respondents on their knowledge and perceptions of a program	• Useful for large target audiences • Can provide both qualitative and quantitative information • Different types of surveys available
Expert panels: the considered opinion of a panel of knowledgeable outsiders	• Experts can share lessons learned and best practices • Where outside validation is required • Where diversity of opinion is sought on complex issues • Where there is a need for expertise
Comparative studies: a range of studies which collect comparative data (e.g., cohort studies, case-control studies, experimental studies)	• For summative evaluations

3.5   Successful Evaluations

In summary, staff notes that on the subject of evaluation, the literature offers many suggestions for increasing the probability of a successful evaluation such as:

- View the evaluation as an integral part of the project and plan for it at the outset. The evaluation should be costed and resourced as part of the project.

- Secure commitment from senior managers.

- Involve all key stakeholders in its planning and execution.

- Develop relevant criteria and indicators to assess policy outcomes from the outset of the project.

- Put in place mechanisms to enable monitoring and measurement of progress.

- Foster a learning environment to ensure lessons are heeded.

- The potential value of an evaluation will only be realized when action is taken on the findings and recommendations emanating from it. Processes are needed to ensure that this happens.

- The Evaluation Plan should be a live document. It should be kept under constant review. The existence of an Evaluation Plan allows changes in the policy objectives and other important parameters to be explicitly noted.

- Roles and responsibilities in analyzing the data and writing the report should be clearly defined in the Evaluation Plan.

- The findings of the evaluation should be validated. It is good practice to share drafts of the report with those who have provided the information (i.e., the study participants). It is important to ensure their views have been fairly and accurately taken into account[11].

---

[11] Review of Evaluation Frameworks, Saskatchewan Ministry of Education, March 2008.

4 Conclusion

Staff recommends that the application of the proposed framework should be limited to codes, rules, filing requirements, guidelines, handbooks, and Reports of the Board; and that it should be applied as follows:

- for future policies, an evaluation plan will be built into the policy making process from the outset;

- for existing policies with identified measurable objectives, the framework will be used to evaluate the effectiveness of these policies; and

- for existing policies where objectives have not been identified, elements of the framework may help the Board to report on the impacts of these policies.

Staff also recommends that the proposed framework be implemented on a pilot basis before being fully adopted by the Board. Staff sees merit in retaining a measure of flexibility to refine the framework as experience warrants. Staff notes that the framework was generalized to apply to any policy. In practice, specific objectives and performance indicators will be set for a particular Board policy.

To support the implementation of the proposed framework on a pilot basis, a structured approach to impact analysis and monitoring and evaluation will include the following:

- **Develop logic model for future policy**. The policy logic model as outlined in section 3.1 has been generalized to apply to any policy. The logic model, however, should be made more specific when developing appropriate outcomes and performance indicators for a particular policy. Some of the general categories can assist with determining appropriate outcomes and indicators, but specific outcomes and indicators should be developed for future policy to assist with effective monitoring, and determining future evaluation plans. Staff notes that a more structured approach to the Board's Terms of Reference process will ensure that a proposed policy is linked to its intended outcomes.

- **Conduct (proportional) impact analysis on future policy.** Answering a structured series of questions provides an opportunity to ensure that the problem (or need) the policy is trying to solve is well defined, the alternatives have been considered, and the anticipated outcomes of the policy are identified along with clear performance indicators and data sources. Staff suggests that stakeholders should play a role in this process.

- **Refine, as appropriate, processes and responsibilities for data collection, analysis and reporting to support monitoring and evaluation**. Effective linkages and robust processes may need to be enhanced to ensure alignment of data collection, reporting, and analysis responsibilities. Staff notes that existing responsibilities and processes should be leveraged.

- **Conduct evaluations on future policy (where appropriate)**. Not all policies will require an evaluation. Jurisdictions that conduct evaluations typically prioritize their evaluation activities depending on the size and complexity of the policy initiative, the significance of the anticipated impact on stakeholders, the level of risk, the relative priority with reference to other policies, and the time and resources available. Also, it will be necessary to identify the policy-specific evaluation questions (building on the questions discussed in section 3.4.4).

5 Appendix A: Constructing an Evaluation Plan

As indicated throughout this Report, evaluation planning is a key component of ensuring success.

The Board's Overall Plan

There are two types of plans for evaluation. The first is an overall organizational plan. For this, the Board would go through a priority-setting exercise and rank the importance of conducting evaluations for its range of new policies. That global plan would list all evaluations intended to be carried out over the planning period (three years from the inception of each policy) and estimate the resources required to conduct each evaluation.

Underlying the plan would be the performance measurement approach adapted from the guidance in the proposed framework. For each planned evaluation exercise, it would be necessary to identify the timing and expected resource requirements. As well, broad evaluation questions should be identified in the global plan. These would include continued relevance to the Board`s priorities, the details of the methodology and specific questions would be left to an individual evaluation plan.

Preparing for a Specific Evaluation

The second is a specific evaluation plan. This evaluation plan should be established at the impact assessment stage prior to the issuance of a new or revised policy. This appendix provides guidance in developing a detailed evaluation plan. Because the plan will be specific to the policy in question and the prevailing circumstances, no actual policy has been chosen for this purpose.

The following principles should be considered in planning:

- the timelines and scope of the proposed evaluation (i.e., issues and extent of coverage);

- a completed evaluation plan should include a list of questions to address the core evaluation issues and proposed approaches to collecting the evidence to address those issues;

- costs and resources required for the evaluation; and,

- if appropriate, the identification of any data gaps.

Closer to the actual time for conducting an evaluation, it will be appropriate to revisit the original plan and revise it as necessary if there have been changes in any underlying assumptions or data availability.

The following questions should guide the preparation of an evaluation plan:

- What policy is going to be evaluated? (The name and a brief description of the policy under consideration)

- What are the reasons for doing the evaluation? (For example, proportionality considerations; part of the Board's planned review cycle; large and important policy, etc.)

- What it is that you want to learn about this policy? (This would include specifying the questions to be investigated in the evaluation, such as impact on a particular segment of the industry; on-going relevance in terms of the Board's priorities, or any other consideration deemed to be important)

- What is the level of support from management? (Is there sufficient interest at the Board level to warrant carrying out the study?)

- What time, resources, and staff are available to help conduct an evaluation? (The budget would be refined and a choice of strategy for conducting the evaluation should be discussed)

- Who are the stakeholders for this policy? Who is involved in policy delivery? Who is affected by the policy? Who will be the primary users of the evaluation results?

- What is the overall goal of this policy? Is there some type of industry behavior or a practice that is being targeted for change?

- What are the policy's objectives?

- What are the major activities and outputs from this policy? (Reference should be made to the Logic Model and the links between outputs and expected outcomes. Specific refinement of the logic model may help to focus the questions for the evaluation in question.)

Evaluation Plan Template

This would list the questions to be addressed, the performance indicators to measure results, the data sources/methods, the responsibility area of the Board, and the timeline for completion of the work.

Typically, an evaluation plan should include concrete evaluation questions to address relevance and performance.

Relevance:

Issue #1: Continued need for the policy: the extent to which it continues to address a demonstrable need and is responsive to the needs of Ontario energy markets;

Issue #2: Alignment with government priorities: evidence of strong links between the policy's objectives and (i) Ontario government priorities and (ii) the Board's strategic outcomes;

Performance (effectiveness, efficiency and economy):

Issue #3: Achievement of expected outcomes: an assessment of progress toward expected outcomes with reference to targets, program reach, and program design, including the linkage and contribution of outputs to outcomes; and

Issue # 4: Demonstration of efficiency and economy: an assessment of resource utilization in relation to achieving expected outcomes

Analysis of the information generated by the evaluation

Types and source(s) of information?

Are the data qualitative or quantitative?

Risk Considerations in Developing an Evaluation Plan

The Treasury Board of Canada Secretariat conducts risk assessments for evaluation and performance measurement planning. Such considerations could factor significantly in the timing and resourcing for evaluations. Below are some considerations, adapted from the Treasury Board's guidance, to assist in planning evaluations:

- Size of population affected or targeted by the policy

- Consequences if risk to policy outcomes materializes

- Probability of the risk materializing

- Degree of control or influence by the Board

- Time available to prevent, mitigate, avoid risk

- Public confidence

- Political sensitivity of the specific issue

There are additional risk considerations relating to the availability of performance information:

- Sound underlying theoretical foundation for the policy intervention

- Robust performance management framework

- Continued relevance of existing performance information

- Time available to generate required performance information

Sample Evaluation Plan

Sample Evaluation Plan Outline

I. <u>Evaluation Framework</u>

  A. What you are going to evaluate.

    1. The initial proposal (assumptions about targeted stakeholders, short-term outcomes, intermediate outcomes, and final outcomes).
    2. Objectives (stated in measurable terms).
      a. What you plan to do, when, and how.
      b. Who will do it?
      c. Targeted stakeholders
    3. Context for the evaluation.

  B. Questions to be addressed in the evaluation.

    4. Are your objectives being achieved? If not, why not (i.e., what barriers or problems have been encountered)? Implementation problems?   Market changes? Which aspects of the policy contributed the most to achieving expected outcomes?

  C. The timeframe for the evaluation.

    5.   When will data collection begin and end?
    6.   How and why was the timeframe selected?

II. <u>Evaluating Objectives – Procedures and Methods</u>

  Question 1: Are Objectives Being Achieved? If Not, Why Not?

  A. [State objective in measurable terms]

December 1, 2011

What to include:

1. Type of information needed to determine if objectives are being achieved such as your performance indicators.
2. Sources of information. Include in your plans procedures for maintaining confidentiality of the information obtained during the evaluation, if necessary.
3. How sources of information were selected.
4. Timeframe for collecting information (dates when the data collection is planned to begin and end).
5. Methods for collecting the information (i.e., records reviews, interviews, paper and pencil questionnaires, and observations).
6. Methods for analyzing the information to determine whether the objective was achieved (i.e., tabulation of frequencies and assessment of relationships between or among variables using statistical analysis, if required).

III. Procedures for Managing and Monitoring the Evaluation

What to include:

1. Procedures for training staff to collect evaluation-related information.
2. Procedures for conducting quality-control checks of the information collection process.
3. Timelines for collecting, analyzing, and reporting information, including procedures for providing evaluation related feedback to the Board.

6 Appendix B: In-house or Independent?

Evaluations can be performed in three ways:

1. Using "in-house" staff to perform the data collection and analysis;

2. Using a combination of "in-house" staff for certain components of data collection, and using evaluation consultants for other components (this may range from basic advice on the data collection and analysis, or complete data collection and analysis by the evaluation consultants); or

3. Using evaluation consultants to perform the evaluations in-full.

The table below outlines the advantages and disadvantages of conducting evaluations internally or with the input of consultants. It should be noted that these choices are situation-specific. Factors to be taken into account include the intended use of the evaluation, whether the evaluation is a tool by which to improve, the nature of the policy being evaluated, the projected cost of the evaluation, the political sensitivity of the policy being evaluated, and other factors.

Advantages and Disadvantages

	Advantages	Disadvantages
Evaluations Conducted In-House	• Potentially cost-effective (if the skill level is strong) • Complete control of the activities associated with the evaluations • May receive better cooperation as the "face" of the Board • First-hand understanding of any difficulties that occurred throughout the process • Tool by which to improve (since feedback is direct)	• Potentially less cost-effective (if in-house skill level is insufficient) • Individuals with limited knowledge of evaluation may not understand some of the potential evaluation challenges and constraints • Potentially draining on certain staff whose job consists of other responsibilities • Credibility of findings or recommendations may be suspect due to a lack of independence of the evaluation team and data collections

December 1, 2011

	Advantages	Disadvantages
Mix of In-house staff and evaluation consultants	• Allows for specialization, which can produce more credible results • Possibly more cost-effective (less expensive methods can be conducted by Board staff, while the consultants would execute more in-depth data collections) • Ability to learn from the consultants, thereby enabling the ability to carry out future evaluations in-house • Highly transparent process • Consultants will provide recommendations • Permits data gathering and analysis by independent consultants to improve credibility for some data collections and findings	• The pace of the work may vary between in-house staff and consultants • In-house staff may require training/guidance from the consultants, which may increase cost, and slow down progress
Evaluation consultants only	• May be more cost-effective (although consultants' rates are generally higher than in-house staff, specialized evaluation consultants may work at a quicker pace) • Consultants will likely do better planning as they can anticipate expected challenges • Consultants, given their experience, understand what methods work best for retrieving data that may be difficult to find • Consultants provide an independent, external perspective on findings • Consultants will provide recommendations • Enhanced credibility for data collection, analysis, findings, and recommendations through greater independence • Would provide Board with experienced "sounding board" for preparation of data collection instruments, data collection issues, analysis of results, and formation of recommendations	• Less transparency (although consultants should provide regular progress reports) • May be a costlier option • May have difficulty gaining cooperation of stakeholders, who are independent, and not affiliated with the Board • Issues surrounding confidentiality • Individuals with limited knowledge of Board policies and processes may not understand some of the potential evaluation challenges and constraints