

Starting Over – Developing an Evaluation Framework and Protocols in Ontario

*Michael Messenger, Independent Consultant for the Ontario Power Authority
Toronto, Ontario*

Abstract

This paper tests the hypothesis that it is both possible and desirable to transfer energy savings evaluation protocols across jurisdictions or states in developed countries. An experiment is underway attempting to transfer the structure and insights from California's measurement and evaluation (M&E) protocols to the province of Ontario. The results of this experiment suggest that it is possible to transfer M&E protocols and evaluation practices across states or jurisdictions but the levels of rigor and comprehensiveness need to be revised to reflect the realities of local evaluation practices. This paper discusses the short-term results to date associated with the transfer of M&E methods and principles to the Ontario planning and evaluation system. Finally, a summary of lessons learned is provided that may be useful for other practitioners attempting to transfer M&E protocols across jurisdictions.

Introduction

The need for simple and coordinated transfer of energy efficiency program "savings" results from jurisdictions across the world has become significantly more important given the rise of potential international cap and trade systems that depend on the credibility of estimates of incremental carbon savings credits. One of the most important political challenges of the 21st century may be to create a consensus across cultures and nations that it is both possible and desirable to trade carbon emissions credits resulting from investments in energy efficiency or alternative sources of energy supply. Part of creating this consensus will be to develop transparent reporting mechanisms or counting rules that track the results of investments in terms of their impacts on energy use and resulting changes in greenhouse gas (GHG) emissions. Developing standardized energy savings protocols and reporting mechanisms across jurisdictions will be a significant challenge but it is likely to be the first significant test of any international system used to report emissions credits. (Violette, Mudd and Keneipp, 2001)

This paper tests the hypothesis that it is both possible and desirable to transfer energy savings measurement and evaluation (M&E) protocols across jurisdictions in developed countries. We define "possible" to mean that the terms of reference used in both jurisdictions can be harmonized and understood by evaluation practitioners. We define "desirable" as leading to an outcome where the accuracy of program savings estimates is increased and the ability to compare energy and peak savings impacts outcomes among programs is improved.

To test this hypothesis, this paper describes attempts to transfer key portions of California's M&E protocols [CPUC, 2006] to the Province of Ontario and the difficulties encountered in making the transition across different models for the administration, development and evaluation of energy efficiency programs. The actual impact of this transfer effort is difficult to assess because the evaluation protocols have not yet been implemented by evaluators in Ontario. In this paper, we describe the strategies used to transfer the evaluation protocol and report on the principles and protocols that have been introduced in Ontario. The transfer of information and standardized measurement systems may be an easy step relative to the more difficult transfer of protocols or trading systems from industrialized countries such as the United States to developing countries, but the lessons learned in this experiment may shed some light on the best way to handle that challenge.

Background

Evaluation protocols for energy efficiency programs have been developed in California over a fifteen year period, with the first generation being adopted by the CPUC in 1993, (Pacific Gas and Electric, Southern California Edison and San Diego Gas and Electric, 1993) and an improved or advanced version of similar protocols adopted in 2006 (California Public Utilities Commission, April 2006,). Both versions of the protocols were driven by the desire of policy makers and program managers to develop an objective and rigorous measurement system that would increase the credibility of programs designed to reduce electricity usage. While some products of this protocol development process were quickly adopted for use by other jurisdictions, most states and surrounding jurisdictions declined to adopt the California evaluation protocols. This decision was in part because the protocols were considered to be a unique product of the California regulatory environment (and tied to performance incentives) and in part because the methodology requirements in the protocols were judged by some evaluation practitioners to be too prescriptive. The author was hired to develop a set of evaluation protocols based on the work done in California but customized to the unique regulatory environment in Ontario. This paper describes the approach to developing protocols in Ontario and the preliminary results from this experiment.

Strategy to Develop and Adopt Protocols

The goal of this project was to get the Ontario Power Authority (OPA) to adopt a set of energy efficiency evaluation protocols that would accomplish two objectives:

- Increase the accuracy of reported program energy savings in Ontario, and
- Increase the likelihood that program designs could be rapidly improved by ensuring that evaluators provided useful feedback on programs that worked and those that needed to be revised.

Additionally, this project wanted to stimulate a sufficient number of analysts to support the development of protocols to evaluate programs, advertise the process to interest local evaluation firms, and hold workshops to discuss the pros and cons of specific M&E protocols. Fortunately, in Ontario, there was some experience with publicly funded energy efficiency programs, and the resource planning agency for Ontario, the Ontario Power Authority, recognized the need to develop a robust evaluation process to ensure that the public's funds were well spent. (Ontario Power Authority, 2006) The strategy used to seed the ground for a successful adoption of protocols in Ontario was:

1. Conduct a situation assessment to determine who has the authority and the motivation to adopt new evaluation protocols;
2. Gain the trust of key opinion leaders and program managers in Ontario by helping them develop and implement a respectable set of energy efficiency programs;
3. Create a vision of how to build up the necessary evaluation infrastructure to conduct and manage evaluations of a variety of energy efficiency or demand side management programs;
4. Build up a rudimentary program reporting structure and introduce simple tools to calculate the cost effectiveness of programs;
5. Provide examples of how program specific evaluation results from other jurisdictions can save time, money, and energy; and
6. Ensure that the resources needed for evaluating programs are part of the annual budget cycle.

The next section describes how this strategy was implemented and then assesses whether the strategy taken as a whole was successful in achieving its goal of developing a solid framework and set of protocols. Each step in the process is examined below:

Strategy 1- Conduct a Situation Assessment

In this context, a situation assessment is an analysis of the likely receptivity of government and policy organizations and businesses involved in energy efficiency programs to the introduction of standardized evaluation protocols. As an initial step, it was important to gauge the likely levels of program activity in the next five years and then assess to what extent evaluation efforts would be required to gauge the effectiveness of the programs. Below we review the likely levels of program funding and match this to an assessment of the state of evaluation practices in Ontario, the level of funding allocated to these efforts, and the expectations in the policy community with respect to ensuring program administrators are accountable for producing accurate assessment of program impacts.

The responsibility for developing, funding, and evaluating energy efficiency programs originally fell to the government owned electricity monopoly Ontario Hydro in the 1990's. This authority has now been transferred to the Ontario Power Authority to both develop its own Conservation and Demand Management (CDM) programs in response to government directives and evaluate them. The Ontario Energy Board is responsible for reviewing and approving the initial program plans and reviewing the results of the CDM programs at the portfolio level.

Ontario's energy efficiency programs and subsequent evaluations have been driven by changes in governments and subsequent swings in policy direction over the last twenty years. Ontario Hydro successfully operated programs at funding levels ranging from \$15 to over \$40 million dollars per year in the early 1990's but the advent of deregulation in 1995 took program funding levels down to less than \$ 5 million per year from 1995 to 2003. Energy Efficiency programs run by local distribution companies (LDC's) were revived in 2004 and these programs spent over \$60 million in 2006.

Energy efficiency programs were developed and implemented by Ontario Hydro between 1989 and 1992 and then fell out of favor and funding with the change in government in 1993. Efforts to revitalize energy efficiency programs began in 2003 when the Ontario Energy Board appropriated \$160 million for local utilities to implement CDM programs over a three year period. Full scale energy efficiency program efforts were formally revived with the creation of the Ontario Power Authority, and the Conservation Bureau within the Authority in 2004. Since then, efforts to build a program delivery and portfolio management structure have proceeded at a steady pace and were significantly accelerated with the government's announcement of a goal to save 1,350 MW of peak demand by 2007. The range of Program spending is projected to increase from \$ 100 million in 2006 to between \$200 and \$350 million by 2010.

During the period between 1989 and 1994 Ontario Hydro was responsible for evaluating their own program efforts but were not required to adhere to any guidelines related to confidence levels or precision in their program savings results. During this time, the utility worked closely with intervenor representatives to make sure their methodology concerns and evaluation priorities were incorporated into the research design at the beginning of the evaluation. After this, Ontario Hydro was responsible for managing the evaluation and reporting on program results. Energy Efficiency programs were placed on hold during the restructuring process in the late 1990's as utilities strove to reduce costs to be better equipped to compete in the deregulated market. After the opening of the deregulated market proved unsuccessful in reducing prices in some markets, the policy of using local distribution utilities to administer CDM programs was revived.

At the beginning of 2004, the Ontario Energy Board took steps to standardize the reporting of program results by requiring that all local delivery utilities use standardized input assumptions for a group of 100 energy efficiency measures when estimating program savings and or cost effectiveness (OEB, 2005). In

the short run, this was a good policy outcome because it introduced some discipline and rigor into the program savings reporting process. In the long run this policy may have discouraged the verification of actual energy savings from CDM programs because utilities were encouraged to rely on fixed ex ante energy savings assumptions in the master data base and implicitly encouraged not to perform any new research to update uncertain measure cost or energy savings values. This effect may have contributed to the decline of the local evaluation infrastructure because there was little demand for independent evaluation of program results once a set of stipulated energy and peak savings per measure has been adopted.

Beginning in 2006, the OPA began to develop a portfolio of Conservation and Demand Management (CDM) programs to save 1,350 MW by 2010 and sought authority and funding to evaluate the programs. This authority was granted, and the OPA began the process of developing an evaluation infrastructure with the publication of a discussion paper on CDM in December of 2006 as part of its Integrated Power Planning System Analysis (OPA, 2006). Unfortunately, the actual funding to begin the program evaluations was not available until early 2007 and formal evaluation RFP's were not released until May of 2007.

In 2006, independent energy efficiency evaluations in Ontario were not conducted very often because the province had not run and operated energy efficiency programs on a province wide scale from 1995 to 2004. Load impact evaluations were not formally required because program administrators had the option of using a standardized set of energy savings and cost assumptions to report on their results. Evaluations of energy efficiency programs focused primarily on estimating the number of measures installed and conducting process evaluations through a survey or market actors.

In addition to the situation assessment, the technology transfer literature was surveyed to determine if this offered any insights about how to transfer information intensive systems such as protocols to a new culture. While the literature offered insights into the stages of information transfer and later acceptance, most of the literature has focused on the transfer of technology to developing cultures, rather than the transfer of certification standards, software or protocols (IPPC, 2000).

Given this situation assessment, it was clear that initial efforts should be focused on rebuilding the evaluation infrastructure and reaching out to the program development teams who desired a primer on evaluation and how evaluation services could support energy efficiency programs. In order to rebuild the evaluation infrastructure, it became evident that a vision of what evaluation services would be needed and how long it would take to get there would have to be developed.

Strategy 2- Build Internal Trust

In jurisdictions where energy efficiency programs are a relatively new phenomenon, there is a premium placed on developing a process to either develop new programs designs and or transfer already successful programs from other jurisdictions. The key to developing a useful planning process is to make sure all of the relevant players are involved, agree to similar objectives, and agree on how the final decisions on program design and content will be made and by whom. In the Ontario case, the author's desire to build trust by importing some program evaluation lessons from California was not successful because utility participants in the planning process had different perceptions of what lessons were learned from some of the energy efficiency programs run in California from 2001 to 2004. Specifically, Ontario utilities and ultimately the government felt the need to increase public exposure to energy efficiency products, recommended behavior changes, and efficiency measures were more important than focusing on the cost effectiveness and savings from potential energy efficiency programs. This resulted in a decision to roll out a modified version of California's 20/20 program (called Summer Savings) across all of Ontario despite the relatively high free rider rates and thus low levels of net energy savings found in the evaluation of the California 20/20 program.

The author was more successful in building trust by working with program managers to develop evaluation plans simultaneous with the development of program designs and savings objectives. Working to develop a theory of how the program would affect target actors in the market place was perceived as very useful.

The lesson learned here is that evaluators should not insist on the use or application of planning criteria or program specific results thought to be important in their home jurisdiction to the exclusion of other important planning criteria or local policy concerns. Key local policy concerns such as the need to build political support and awareness of energy efficiency options in the new jurisdiction are often more important than evaluating the net savings from programs in the short run. It is more important to be perceived as useful to the development of an evaluation planning process than to insist that the locals use the evaluation results from other jurisdictions in their planning process.

Strategy 3- Create an Evaluation Vision

From the situation assessment, it was clear that there was a need to develop a process to ensure that program results were reported in a timely manner before analysts moved on to the next big program design challenge. It was also important to emphasize the need for independent evaluation of program outcomes and to periodically test and evaluate the assumptions used to estimate program savings. After discussing these needs with the program managers, a vision of where to go and a timeline of how long it might take to get there was created.

Table 1 shows a vision for how evaluation services should integrate with program and delivery functions over time in Ontario.

Time Line to Achieve Evaluation Vision

	2006-Q4	2007 Q1	2007 Q2	2007-Q3	2007-Q4	2008 Q1	2008 Q2	2008 Q3	2008 Q4
Level 1-Timely Reporting & Evaluation Plans in place before each Program Launch	Yellow	Yellow	Yellow						
Level 2- Independent Verification of Measure Installations and savings Assumptions within 6 months of program launch		Cyan	Cyan	Cyan					
Level 3- Verification of Gross Energy Savings using Energy Bills and Onsite monitoring			Magenta	Magenta	Magenta	Magenta	Magenta		
Level 4- Verification of Net Energy Savings and persistence of measures			Blue	Blue	Blue	Blue	Blue		
Level 5- Impact of Programs on Market Structure/Transformation					Orange	Orange	Orange	Orange	Orange

Table 1- Evaluation Vision for Ontario

This table illustrates that the focus on developing a credible program data collection and reporting infrastructure was the first priority, followed closely by developing a method to confirm measure installations and base line energy use conditions for program participants. The near-term focus was on building up the capability to evaluate measure or technology level savings first, to be followed by a later focus on estimating gross and net energy savings, and, ultimately, impacts of the program on market structures, share, and prices of more energy efficient products.

Strategy 4- Build up the Program Reporting Infrastructure

In many jurisdictions, there is a tendency to focus all available resource on developing and launching programs first and then thinking about program tracking systems later. An effort was made to counter this trend by providing standardized tracking data bases from California for Ontario's use. This effort was ultimately unsuccessful in part due to disagreement on the minimum amount of program specific data that needed to be collected from the perspective of the program managers and their delivery agents. In addition there was a desire to upgrade the data collection instruments from spreadsheets or databases used in California to allow users to upload their data to the web directly in Ontario. Both factors led to a delay in developing and implementing a universal reporting infrastructure.

Strategy 5- Provide Examples of How Evaluation Results Have Been Used to Improve Previous Energy Efficiency Programs

This strategy proved to be relatively successful because it was possible to insert a review step for program evaluation as part of the program planning process. Program managers were eager to read the results of evaluations from California and other jurisdictions that related to their proposed program designs. Most program managers intuitively understood how process evaluations that reviewed the effectiveness of working with different retail, distribution, or manufacturer organizations to promote specific products or systems might be used to improve their proposed designs in Canada. The lesson learned here was that the skill of being able to quickly survey the literature to provide specific program evaluations that are relevant for program design proposals is likely to be very important in jurisdictions without a history of developing energy efficiency programs or measurement methods, or with little knowledge of similar efforts in other jurisdictions.

Strategy 6- Ensure that the Resources Necessary to Evaluate Energy Efficiency Programs are part of the Program or Portfolio Budget

Jurisdictions or organizations with successful evaluation infrastructures can usually be traced to either a sponsor or champion who insisted on the comprehensive evaluation of efficiency programs early in their history of development, or the use of a budgeting process that explicitly budgets for the costs of evaluations using a rule of thumb relating evaluation effort to overall program expenditures. California has a history of devoting from 1 to 3 per cent of program expenditures on evaluations of program energy and peak savings impacts and closer to 4 to 6 per cent of program expenditures when expenditures needed to gather data on appliance saturation, building characteristics, energy efficiency potential and or process evaluations are included as part of the evaluation budget total. The use of this rule of thumb, "evaluation should cost between 3 to 5 per cent of program expenditures", as a budgeting tool was not particularly convincing to program and portfolio managers who faced a situation of constrained program development budgets coupled with very high expectations about the expected savings impacts of their programs. As of this writing, evaluation plans and budgets for specific programs had been approved, but the effort to develop a specific evaluation budget for the next budget year is still a work in progress. The lesson learned from this experience is that it will usually be necessary to budget for program evaluations from the bottom up with a list of specific evaluation tasks in order to gain support for an evaluation budget line item in the next year's budget.

Overall Effectiveness of the Protocol Transfer Strategy – What happened?

The strategies discussed to develop an evaluation infrastructure and bring clarity on the roles and responsibilities of the various organizations involved in the development of programs were implemented in the latter half of 2006 in preparation for the development and release of evaluation protocols in 2007. The protocols were approved internally but had to wait for internal review and internal within other agencies. A plan to gather comments from the relevant stakeholders was developed and the actual evaluation protocols are scheduled to be released in the summer of 2007 (Messenger, 2007).

There were some difficulties encountered in the process of developing these protocols that slowed their development and approval. These included:

1. Lack of an established evaluation business unit and budget to support its activities
2. Lack of an established program tracking and reporting process

Traditionally program managers are charged with the responsibility for evaluating their own programs although this is not the case in some jurisdictions like California. When or if the program management recognizes the need to create some form of independent evaluation, the task of deciding where to place the evaluation function within an organization's structure is a key question for organizations charged with planning and delivery DSM programs. In Ontario, it was decided to create an evaluation function outside of both program design and delivery divisions. While this choice created the conditions necessary for independent evaluation, it made it more difficult to get program design and evaluation staff to work together and it complicated the process of developing a budget for the evaluation group since funding was initially seen as a zero sum game. Fortunately this problem has been resolved.

In addition to difficulty in getting funds to support evaluation activities, it was also difficult to convince program planning staff of the need to develop a comprehensive and consistent program reporting framework or process because of difficulties in getting agreements on what data must be tracked and what was optional or only needed for evaluation purposes. Given these difficulties, the development of evaluation protocols dependent on a program reporting structure took much longer than expected and the final product needed to be simplified to be useful for Ontario. Below we present an overview of the contents of the protocols and note that the final requirements are likely to change after the public review process is completed.

Audience and Contents of the Evaluation Protocols

The protocols are designed to affect three main professional segments: program designers, portfolio and program managers and evaluation professionals. The overarching intent of the protocols is to require all evaluation professionals to conduct a series of planning and analyses that is likely to improve the overall quality of both the evaluation design and reduce the uncertainties in the final estimates of program energy and peak savings results. The required evaluation reports are displayed in Table 2.

Protocol Report	Intended Audience
Draft Evaluation Plan	Program Design, Delivery and Evaluation professionals
Review of Ex-ante Energy Savings Assumptions	Program Designers and Managers
Financial Audit of Program Expenditures	Portfolio Managers and Delivery Managers
Verification of Measure Installations and Baseline Energy Use Conditions	Program Designers, Managers and Resource Planners
Gross and Net Energy Program Savings-kWh	Design and Portfolio Managers
Peak and Average Demand Program Savings =kW	Portfolio Managers, Resource Planners, and General Public
Process Evaluation	Program Designers

Table 2- Protocol Reports and their Intended Audience

The draft protocols for Ontario require that each program development effort is supported by the development of a draft evaluation plan that includes key program objectives, a program theory, and draft research objectives. After this step, a series of program evaluation reports are suggested over the life cycle of two program types: resource planning and capability building programs.

Each of the seven major protocols include a description of the purpose of the protocol, the expected types of results to be presented in each report, and a list of acceptable methods to use in developing the evaluation results. Confidence and accuracy guidelines for the development of sample designs are suggested but not mandatory. Each protocol also includes a section on the intended audience and uses of the anticipated results, who is responsible to complete the report, and the process to be used in reviewing draft and final reports. (Michael Messenger, 2007).

The ultimate impact on the evaluation protocols will not be known until after the draft protocol undergoes public review and revisions. What is clear is that the OPA program design and delivery divisions have accepted the responsibility of preparing evaluation plans for all programs and following a process to determine each program's energy and peak impacts as well as their effectiveness. The OPA management is committed to adopting a set of protocols to ensure that the impacts of CDM programs are well understood and capable of being compared to the contributions made by other sources of new supply being purchased by OPA.

Conclusions

Pursuit of the strategies described in this paper have demonstrated that it is possible to transfer the concept of evaluation protocols from a state with many years of energy efficiency evaluation experience to a province where an energy efficiency evaluation infrastructure had largely vanished. Experience on this project suggests that program managers and local officials remain skeptical that the cost of completing

rigorous evaluations will be worth the benefits in terms of improved program design and more robust estimates of program impacts on electricity demand. However, the OPA program managers have been convinced that it is essential to simultaneously draft program implementation and draft evaluation plans before the programs are launched.

Overall success in this transfer effort will be measured by the extent private evaluation firms and the OPA begin to incorporate some or all of the more advanced evaluation techniques discussed in the California protocols into their evaluation reports in Ontario. It should also be possible to test whether the adoption of protocols leads to increased levels of confidence in estimates of energy and peak savings from programs by regulators and program administrators in Ontario because the protocols require the reporting of confidence intervals around all estimates of energy and peak savings.

Lessons Learned

There were four key lessons learned with respect to working with program staff and managers to organize the simultaneous launch of a portfolio of energy efficiency programs and an evaluation structure to support them. These were:

1. It is important to develop a shared vision of what benefits different types of evaluations can provide to local market participants and how quickly the infrastructure to support them can be developed.
2. It is critical for evaluators to leave the policy goals from their previous jurisdiction at the border rather than unconsciously try to introduce them into a new jurisdiction. Programs should be evaluated based on the policy goals and criteria adopted by the local jurisdiction, rather than the policy goals from other jurisdictions.
3. It is important to review the evaluation literature for all of North America to access evaluations of similar program designs for lessons learned as part of the program planning process. This service is an important added value that evaluation professionals can provide in jurisdictions with new program administrators or program delivery agents.
4. It is difficult to transfer rules of thumb with respect to the size or magnitude of the budgets that will be needed to support evaluations for a portfolio of programs. Simply citing the experience of other jurisdictions to justify program evaluations usually cost 3 % to 5% of program expenditures was not enough. In most cases, program planning authorities must go through their own evaluation budgeting process on a bottom up basis, rather than relying on experience from other jurisdictions.

References

1. California Public Utilities Commission 2006. **California Energy Efficiency Evaluation Protocols: Technical, Methodological and Reporting Requirements for Evaluation : Professionals** (San Francisco, CA, California Public Utilities Commission, .April 2006)
2. Chief Energy Conservation Officer 2006. **Ontario- a New Era in Electricity Conservation- the Annual report for 2006** (Ontario Power Authority, November 1, 2006)
3. Daniel Violette, Christina Mudd Hagler Bailly Services Inc.) and Marshall Keneipp (Summit Blue Ventures) 2001. **An Initial View On Methodologies For Emission Baselines: Energy Efficiency Case Study** (OECD Environment Directorate and the International Energy Agency, Paper # 2 : 2001)
4. (IPPC) Intergovernmental Panel on Climate Change, Working Group III, 2000. **Special Report Methodology and Technical Issues in Technology Transfer**, Geneva, Switzerland: Intergovernmental Panel on Climate Change.
5. **Michael Messenger 2007. Proposed Measurement and Verification Protocols for Ontario's Conservation and Demand Management programs** (Consultant report to the Ontario Power Authority: June , 2007)
6. Ontario Energy Board, **Total Resource Cost Guide** (Toronto, Ontario : September 8, 2005)
7. Ontario Power Authority. **Discussion Paper 3: Conservation and Demand Management Revised**, Section 4.4 Evaluation, Measurement and Verification: (December 2006)
8. **Pacific Gas and Electric, San Diego Gas and Electric, Southern California Edison, California Energy Commission, Office of Ratepayer Advocates (CPUC) and Natural Resources Defense Council, 1993: Protocols and Procedures for the Verification of Costs, Benefits, and Shareholder Earnings from Demand Side Management programs** (Adopted by the California Public Utility Commission , in Dec #93-05-063 and revised in March 1998)