

Real-Time Evaluation: Blurring the Line between Program, Evaluation and Market Research

David Cohan, Northwest Energy Efficiency Alliance

Richard Hazzard, Energy Market Innovations, Inc.

Robert Bordner, Energy Market Innovations, Inc.

ABSTRACT

The Northwest Energy Efficiency Alliance has been employing ‘real-time evaluation’, a type of formative evaluation rarely used in the energy-efficiency industry, to address two common situations in which traditional ex-post evaluations cannot provide adequate or timely information: (1) pilot programs when the decision to move ahead with a full-scale program needs to be made in the latter stages of the pilot, before all results are available; and (2) when small-scale, targeted market research is necessary to make a program more effective after it has been launched but the implementation staff either lack the resources or expertise for such work.

In the first situation, we have had evaluators work closely with program implementers reviewing materials and providing verbal and written program feedback on very short turnaround. This ensures that our program decision-makers have the maximum amount of evaluation information at all times. For market research situations, evaluators are generally more qualified to plan and conduct such work than implementers. In addition, it is more cost-efficient to have an evaluator who is already intimately familiar with the project do the work than to hire an outside contractor with no project experience.

The potentially negative side of real-time evaluation is that such close association with the program may put the evaluator’s credibility or objectivity at risk. This paper will describe the value real-time evaluators have brought to Alliance projects, discuss trade-offs between traditional and real-time evaluation approaches, and suggest ways to decide when real-time evaluation makes sense and how it can be used to complement traditional approaches.

Introduction

Energy efficiency programs have been evaluated since the 1970s to ensure effective use of program and utility ratepayer funds. Both the utilities that proposed the programs and the regulators who approved them based on planning assumptions required verification that the promised energy savings had indeed been achieved and that the programs were administered in an efficient manner. Short funding cycles (programs were most often approved a year at a time) and regulatory proceedings originally developed for supply-side prudence reviews and ‘used and useful’ determinations led naturally to an emphasis on what the energy efficiency industry refers to as ex-post evaluations – after-the-fact analyses used to determine if a utility should recover costs it had already incurred in running a program. The obvious advantage to ex-post evaluations is that all program data are available, allowing for comprehensive analysis and definitive results. The equally obvious disadvantage is that the information gained from the evaluation is only

available after the program has been implemented and the possibility of using this information to make mid-course improvements is limited at best¹.

In the evaluation literature, ex-post evaluation is referred to as summative evaluation. Summative evaluations examine the effects or outcomes of some object -- they summarize it by describing what happens subsequent to delivery of the program or technology; assessing whether the object can be said to have caused the outcome; determining the overall impact of the causal factor beyond only the immediate target outcomes; and, estimating the relative costs associated with the object. While process evaluations have been a staple of energy efficiency programs since they began, in many cases even they have effectively been relegated to summative status since, even when the evaluation was done while the program was still operating, no mechanism existed by which implementers of the evaluated program could receive the information or act on it.

Summative evaluations are contrasted to formative evaluations which are intended to strengthen or improve the object being evaluated -- they help form it by examining the delivery of the program or technology, the quality of its implementation, and the assessment of the organizational context, personnel, procedures, inputs, and so on². The Northwest Energy Efficiency Alliance and its contractor, Energy Market Innovations, Inc., have employed a type of formative evaluation we call 'real-time' evaluation³. The goal is to get market research and evaluation information to program planners and implementers when it will have the greatest value in supporting program decision-making. In select situations, we have pushed this approach well beyond its normal use in the world of energy efficiency program evaluation. Far enough, in fact, that the traditional boundaries between program and evaluation activities have been blurred and we have been forced to re-assess the role that the evaluator can and should play in improving programs that are in active development and implementation.

Alliance Program Philosophy

The Northwest Energy Efficiency Alliance is a regional market transformation organization funded by the electric utilities in Washington, Oregon, Montana and Idaho. The Alliance is unusual among major funders of energy efficiency programs in that it is unregulated⁴. Equally important, it is funded in five-year cycles with most individual programs receiving funding for three or more years, not including a planning phase that may last up to a year. These

¹ In the extreme, though not uncommon, situation where utility evaluations are conducted after programs have been completely shut down with no intention of reviving them, there is no possibility of evaluation findings being incorporated into improved program delivery. These have been humorously described as 'post-mortems' rather than evaluations!

² The two definitions given here are taken from the Research Methods Knowledge Base developed by William M. Trochim of Cornell University. See <http://trochim.human.cornell.edu/kb/intreval.htm>. The definitions are inverted from the original text.

³ A reviewer of this paper commented that there is a distinction between formative and real-time evaluation. We have chosen to use the latter term as its meaning is much more transparent to the people we work for and with. Further, a review of the Evaluation Thesaurus (Scriven, M., Sage Publications, 1991 Fourth Edition), shows no entry for "real-time evaluation" so we feel this term is unlikely to cause confusion.

⁴ The Alliance is a 503(c) non-profit corporation that was established in 1996. Its Board consists of representatives from public and private utilities, state government, energy efficiency advocacy groups, private industry, and regulatory bodies. The regulators are non-voting members.

long timelines give Alliance staff the freedom to think of programs evolving through a variety of stages over extended periods of time.

One aspect of this is the philosophy of ‘adaptive management’ that guides program offerings. On a practical level, this refers to changing program parameters as new information about the market becomes available. This philosophy has been an explicit part of the Alliance since its inception and is deeply engrained in the thinking of both the Board and the staff. Changing a program to adapt to new circumstances is seen in a completely positive light; indeed, change is an almost universal expectation each time a new evaluation or market research study is completed. As a result, there is a clear connection between the accepting attitude toward program change and the evolving use of formative methodologies.

Evolution of Program Planning at the Alliance

As a new organization in 1996, many of the first Alliance initiatives were ‘legacy programs’, inherited from or built upon utility or regional efforts that already existed in the Northwest. Though none of the legacy programs were traditional rebate programs, many still had the look and feel of utility information or product promotion programs in that they were focused on individual technologies (e.g. washing machines, CFLs), processes (e.g. wastewater facility optimization), or types of information (e.g. technical assistance for specific end uses).

Evaluators produced Market Progress Evaluation Reports (MPERs) that covered both traditional process analysis and the developing concept of ‘market progress’ or how far along the market transformation path a program had advanced. When applicable and practical, market research established a baseline at the beginning of implementation; after that an MPER was issued approximately once a year. (Evaluators initially thought to report more frequently but soon realized that market progress over a six- or eight-month period was typically too small to measure or provide meaningful program information.) While there has always been an emphasis on getting evaluation results out quickly to implementers, the typical one year cycle for evaluation activities meant that new information was only rarely available.

This system worked well for these first programs because, as mentioned, they followed from existing efforts and their ‘trajectories’ were fairly predictable. Over time, however, Alliance thinking about how to effectively transform markets has evolved and one of the concrete, if unintentional, results has been a tendency toward more complex program designs that more realistically reflect the complex, dynamic markets the Alliance is trying to influence.

The Alliance’s most advanced effort to date in this respect is the Commercial Sector Initiative (CSI), which contains over forty separate programs ranging from ideas still on the drawing board to tests and pilots to long-established efforts such as the Lighting Design Lab in Seattle. (To make things even more interesting, every one of the CSI programs purposely overlap with at least two or three others; most overlap with many more.) This blend of many individual but overlapping programs in so many stages of development is justified theoretically but, not surprisingly, presents significant difficulties in administration and implementation.

As a result, the implementation teams desire more frequent and faster feedback, especially during initial program implementation, to identify opportunities and obstacles in the market that were not apparent during program planning.

Evaluation of CSI and the Beginnings of Real-Time Evaluation

From an evaluation perspective, two somewhat conflicting responsibilities have slowly become clear as program complexity has increased. One is that the critical evaluation function of providing summative program information for Board funding decisions must continue indefinitely. The other is that the standard summative evaluation approach does not provide process information frequently enough for the needs of implementers. From the Alliance perspective, there is no benefit to having an ex-post evaluation find that a program was a failure when the failure was preventable through improved market or program information.

Confronted with this situation, the Alliance began to implement a version of real-time evaluation that consisted of two new roles where evaluators could add value: market research and quality assurance. The former is typically the domain of program planners; the latter is typically the domain of program implementers. The reality of complex programs is that planning is an on-going task that falls increasingly to implementers once programs are approved and quality assurance is rarely funded adequately in any program. Most important, program implementers often do not have the skills or perspective appropriate for the research and analysis needed for these so it makes sense to include evaluators' expertise in this process.

For the purposes of real-time evaluation, the scope of these roles is tightly defined. Market research is limited to small scopes of work where information is needed quickly. This research is intended to address a program situation where the implementer says, "If I only knew this, I could make a much better decision." Importantly, such research is not meant to be a substitute for the large-scale market research used in program planning. Quality assurance is constrained to the beginning stages of a project. For example, when a database is first implemented, real-time evaluators will determine if it is being fully populated with appropriate data and explore how users are interacting with the database.

While market research and quality assurance are traditional areas of work for evaluators, their implementation within CSI has been highly unusual – evaluators working directly with the implementation team in both the planning and implementation (though not the analysis!) of the work to ensure that implementers get the information they need when it will be most valuable to them. It is this aspect that potentially makes the Alliance real-time approach controversial because the evaluator is effectively working as part of the implementation team while being charged with the responsibility for the long-term reporting to the Board⁵.

Patton [1997] describes academic evaluators who "tend to...emphasize the research purposes of evaluation, traditional standards of methodological rigor, [and] summative outcome studies" as contrasted with service evaluators who "emphasize serving stakeholders' needs, program improvement, qualitative methods, and assisting with program decisions. (Patton). Interestingly, his discussion is about the different views individuals have of evaluation. In the current case, the question is whether one person can possibly play both roles. We will return to this topic in the Discussion and Conclusions section.

⁵ While one could argue that this conflict could be completely resolved by hiring two contractors, there are both budget constraints and timing constraints that generally make that impractical.

Case Studies

To give readers an idea of how real-time evaluation has been conducted at the Alliance we provide two case studies below.

#1: Small Commercial HVAC O&M Service Pilot

Program Description

In January 2002, the Alliance funded a pilot for small (<15 ton) roof top HVAC units (RTUs). The goals of this project were to:

- Define a service protocol that could be used as the basis for an enhanced O&M service for rooftop RTUs.
- Develop appropriate training and marketing strategies to promote and deliver the new service.
- Develop a framework to document and analyze the benefits in real world examples.
- Test the market acceptance of the service.

Portland Energy Conservation, Inc. (PECI), a Portland based firm, was awarded a contract to implement this project. Energy Market Innovations, Inc. (EMI) was awarded the evaluation contract in March 2002. The project was divided into the following two phases:

- Phase I in 2002 involved primary research and development of the service protocol, various market research tasks, and limited testing of the service with participating contractors. Phase I field tests included five HVAC service providers and the servicing of 45 roof top heating/air-conditioning units (RTUs).
- Phase II in 2003 involved refining the service protocol, development of specific technical and sales training curriculum, recruiting HVAC service providers and utilities, and more extensive field-testing throughout the Pacific Northwest. Phase II included 20 service providers, 14 utilities, and the servicing of 125 RTUs.

Real-Time Evaluator Role

(1) Phase I: consumer research, Bend, OR. Market research completed during the planning phase had shown significant energy savings available from RTUs (which justified running the pilot), but provided only cursory information on the structure and function of the small HVAC market. The implementation budget had no funds allocated for market research and, because of the task schedule, PECI was focused on developing the technical protocol for most of Phase I. At the same time, they recognized that a lack of market knowledge would hamper all efforts to convince service providers to promote the protocol once it was ready. In this situation, the Alliance turned to the evaluator to conduct a market characterization study. Bend, Oregon was selected as a community that represented the overall Northwest market and, over a period of three short weeks, EMI completed a series of 60 in-depth interviews with property managers, small business owners, commercial tenants, commercial property owners, and franchise

representatives to understand their relationships with HVAC service providers and their decision-making considerations regarding HVAC maintenance and costs.

This research provided both an immediate and a long-term benefit. Immediately, the results were used to create the discussion guide for focus groups (added at the suggestion of EMI) and inform the development of the marketing materials that were used throughout the pilot. In the long-term, the research established a foundation of market understanding for both PEGI and EMI that guided all future program and evaluation activities.

(2) Tracking of participating contractors. Because the success or failure of the pilot ultimately depended on HVAC service providers' willingness and ability to sell the enhanced O&M service, it was imperative to have a thorough understanding of how well the new O&M service would fit into the service providers' existing business models. This was another area where basic information was available, but not at the level of detail needed by the implementers. Therefore, a key real-time evaluation task, initiated early in Phase I and continued throughout the pilot, was to conduct field observations and in-depth interviews with the participating service providers. These included ongoing interactions with HVAC business owners, service managers, sales managers, and service technicians. Through this process, the evaluator developed an intimate understanding of HVAC business models and the challenges that the enhanced O&M service presented. This intimacy allowed the evaluator to ask increasingly sophisticated and probing questions about how the new service could be integrated into day-to-day business operations, how the service could best be communicated to clients, what kind of clients would most benefit from the service, and what role the utility should play in supporting the program. This real time tracking, akin to both the Alliance and the contractor having an informed ear to the ground, allowed PEGI to use quantitative and anecdotal feedback to adapt aspects of program development and implementation much more quickly than they would have been able to under a traditional evaluation approach.

#2: CSI Professional Training Series

Program Description

The Alliance, under its Commercial Sector Initiative, issued a contract in July of 2003 to Price Putnam Group (PPG), a regional energy efficiency contractor, to conduct a series of professional education trainings targeting the design/build market actors in the commercial building sector in Washington, Oregon, Idaho and Montana. Professional trainings represent a significant component of the CSI strategy of promoting the benefits of high performance buildings. PPG designed and delivered the following three types of trainings:

- *Brown Bag Lunch and Learn* – 60-75 minute presentations offer architecture and engineering firms an opportunity to learn about high performance buildings in their own offices.
- *Workshop* – Half-day seminars designed as an interactive forum to provide architects, engineers, facilities and design professionals with an immersion in specific topics of high performance buildings.
- *Roundtable* – Two-hour forums for informal, open discussions among regional experts that explore specific areas of high performance design. The moderated format is

designed to facilitate open dialogue about different perspectives on new or problematic building performance issues.

The contractor was assigned an extremely aggressive implementation schedule that required them to assemble a team of technical and programmatic experts, develop curricula, and market and implement 60 trainings throughout the Pacific Northwest between August and December 2003. Because PPG was employing several new instructors to deliver untested content in a variety of untested forums across four states, it was imperative that they receive feedback to what was going on in the field as quickly as possible.

Real-Time Evaluator Role

Energy Market Innovations, Inc. was hired to provide real time evaluation services in September 2003 and designed a work plan that included the following tasks:

- Participation in project development and project management meetings.
- Participation in a 'Train the Trainers' seminar.
- Observing Brown Bags, Seminars and Round Tables and interviewing participants.
- Interviewing the implementation team and instructors
- Preparing Summary Reporting Memo

EMI observed ten trainings and also interviewed twenty-two training participants, six instructors and PPG's marketing and administration team. The feedback from these activities was summarized and presented to the PPG team by EMI in weekly meetings. In the absence of EMI's work PPG would have had, given the demanding implementation schedule, only anecdotal, piece-meal information concerning communications and marketing, curriculum and instruction issues, and customer satisfaction ratings. Instead, PPG was provided with comprehensive information presented in a form they were able to incorporate immediately as the program implementation unfolded.

A further benefit of the real-time evaluation work was to support PPG's development of an implementation plan for 2004. Because of the schedule of Alliance Board meetings at which projects are approved, the 2004 plan had to be developed before the work for the first contract was completed. Through the interview process EMI was able to identify services and topics the market valued that were not covered in the first contract. They also gained insights into the level of information participating firms wanted and provided recommendations for formats that had not been considered, such as a series of increasingly sophisticated presentations at one firm.

Discussion and Conclusions

As the discussion and case studies above show, the Alliance's forays into real-time evaluation represent a departure from the traditional roles evaluators have played in the energy efficiency industry. Those who view evaluation's critical role as being objective may question the intimate involvement of evaluators in implementation activities at any time. Even those who accept that evaluators might bring value to such activities may question the ability of those same evaluators to perform the formative and summative roles equally well. However, while the Alliance's experiences to date are not sufficient to make a final judgment, the outcome of the

two case studies show that the real-time role has not, in and of itself, unduly influenced the evaluation results. In the Professional Training case study, the real-time evaluators provided information that helped guide the planning for future improvements; in the Small Commercial HVAC project they provided information that led to the program being shut down.

Regarding the mixing of formative and summative roles, one could argue strongly that, regardless of how well the evaluator does the two jobs, the perception of conflict of interest is too great to overcome and the roles should therefore be split in all cases. While this argument merits consideration, the final assessment of such risk must reflect the specific situation and resource constraints at hand. There are real costs to the organization in (1) bringing a second evaluator 'up to speed' on a program, and (2) losing the detailed knowledge gained by the real-time evaluator. A compromise solution for using a single evaluator is to limit real-time activities to the earliest stages of program implementation and then insist on a complete break from program-evaluator interactions for a substantial interval before the evaluator begins summative work. Decision-makers would need to decide in advance whether this would satisfactorily overcome the perception of conflict of interest.

For all discussions of real-time evaluator-implementer interactions, it is important to note that, while the two are working together, they do not have the same job. The evaluator is never involved with actual implementation, but is tasked only with supplying information and (if desired) informing decision-makers as to how that information can be best applied. The clear overlap between the two is in trying to ensure that the program succeeds. And while there is an obvious and undeniable bias that results from evaluators knowing that this is the result they are looking for, the situation is more ambiguous from the funding organization's perspective. The Alliance only funds programs that it believes can succeed and it has an absolute interest in helping them achieve success. If hiring an evaluator early on, in any capacity, enhances the chances of achieving its organizational mission, then it would be irresponsible not to do so.

In conclusion, the Alliance acknowledges the many issues raised by using real-time evaluators but has so far found the real-time role to be beneficial to all parties – to the Alliance in having an independent observer in the trenches working on its behalf; to the implementers in having an independent resource working on their behalf to provide valuable information; and to the evaluator in gaining an increasingly sophisticated understanding of the issues and obstacles that influence the programs.

References

Scriven, Michael, 1991, *Evaluation Thesaurus Fourth Edition*. Sage Publications.

Patton, Michael Quinn, 1997, *Utilization-Focused Evaluation Edition 3*. Sage Publications.