An Inside Look at a U.S. Department of Energy Impact Evaluation Framework for Deployment Programs

Edward Vine, Lawrence Berkeley National Laboratory
Gretchen Jordan, Sandia National Laboratories
John H. Reed, Innovologie, LLC
Jeff Dowd, U.S. Department of Energy

ABSTRACT

The U.S. Department of Energy's (DOE) Office of Energy Efficiency and Renewable Energy (EERE) is developing a theory-based approach to impact evaluation that could be used by its deployment programs for evaluating energy savings and market effects with credible attribution of impacts (DOE forthcoming). The purpose of this paper is to describe the framework and its research design. The framework also provides information for program improvement in a consistent and structured manner. It joins Everett Rogers' diffusion of innovation theory with logic models to examine linkages between program activities, target audiences, behavioral and institutional changes, and energy savings or adoption of cleaner energy sources. Using the framework's templates, a program can describe its outcome goals and program logic, as well as identify key outcome questions and indicators (metrics). Evaluators could use the framework to understand where to look within the program logic for measured outcomes such as sales or adopted technologies and practices. Finally, by using the framework a causal link between the program and outcomes can be tested and alternative explanations investigated.

Introduction[1]

The mission of the Office of Energy Efficiency and Renewable Energy (EERE) in the U.S. Department of Energy (DOE) is to strengthen America's energy security, environmental quality, and economic vitality. To accomplish this mission, DOE supports the research and development of new technologies and deployment programs aimed at diffusing energy efficient and renewable energy technologies more widely into markets. In this paper, we will refer to "energy savings" as the primary outcome rather than list all of the specific goals of EERE programs. EERE has designed an evaluation framework for use by its deployment programs in guiding the evaluation of its programs (hereafter referred to as the Evaluation Framework). The Evaluation Framework focuses on deployment programs that include activities such as: identifying promising technologies, facilitating development of supportive public infrastructure and policies/regulations, facilitating creation of business infrastructure, providing technology adoption supports (e.g., technical assistance and financial incentives), and conducting market outreach.

[1] The work described in this paper was funded by the Office of Energy Efficiency and Renewable Energy, Office of Planning, Budget and Analysis of the U.S. Department of Energy under Contract No. DE-AC03-6SF00098.

Federal program managers are being pressured by Congress and the Executive Branch's Office of Management and Budget (OMB) to demonstrate the performance of their programs and provide evidence of results. Congress and OMB want to know what outcomes or impacts can be attributed to the program that would not have occurred otherwise. Assessing outcomes and separating program-driven effects from non-program effects is difficult. Many influences other than the program cause the various market actors, sellers, distributors, buyers, policy makers, and others to change their attitudes and behaviors. Consequently, OMB wants to see that federal programs "do independent evaluations of sufficient scope and quality and indicate that the program is effective and producing results" (OMB 2005). OMB expects programs to provide empirical evidence of the linkages between program activities, outputs, and short-, interim-, and long-term outcomes, and to use an appropriate research design to assess impacts and address attribution.

An appropriate action in responding to these calls for credible impact evaluation is for Federal managers to develop a program theory explaining how their program will induce changes and then to initiate independent evaluation, measurement, and verification studies to see if the program theory is supported, including an examination of non-program influences on those outcomes. This paper describes an impact evaluation framework that EERE has built based on program theory. The framework joins Everett Rogers' diffusion of innovation theory (Rogers 2003) with logic models to examine the linkages between program activities, target audiences, behavioral and institutional changes, and energy savings. Logic models at three levels of detail provide templates that describe the basic domains of the market and describe the possible cause and effect pathways in a consistent and structured manner. In addition to energy savings and other effects on end users, the framework covers effects in three market infrastructure domains (knowledge, government, and business). As we describe the high-level logic and causal links between the program and outcomes in both end user and market infrastructure domains, the possible non-program influences on the desired outcome goals become clearer. That leads us to the discussion of the framework's research design, identifying what data to collect and how to collect it in order to assess progress, measure outcomes, and investigate alternative explanations.

Elements of Research Design for Assessing Technology Deployment Impacts

First, an evaluation needs to demonstrate that an outcome occurred, which requires (1) having a baseline of pre-program behaviors, (2) knowledge of program activities or interventions that would affect behavior, and (3) knowledge of changes in behavior of the target audience. All of this information is included in a good program theory, describing what a program intends to do and how it will affect that change. Second, the research design must be able to demonstrate that the observed changes or outcomes are due at least in part to the program and not due to other influences or mere coincidence. Demonstrating this causality (attribution) is usually the most difficult part of the evaluation.

The Evaluation Framework makes formulating evaluation questions easier, because the logic model templates organize possible changes/effects into four domains: three related to infrastructure (knowledge, public/policy, and business infrastructure), and one related to the end users of the technology or practice. For each of these domains that apply to a program, the evaluation should describe and assess possible non-program-induced effects as well as program-induced effects. Research design and methods are the subject of the final section of this paper.

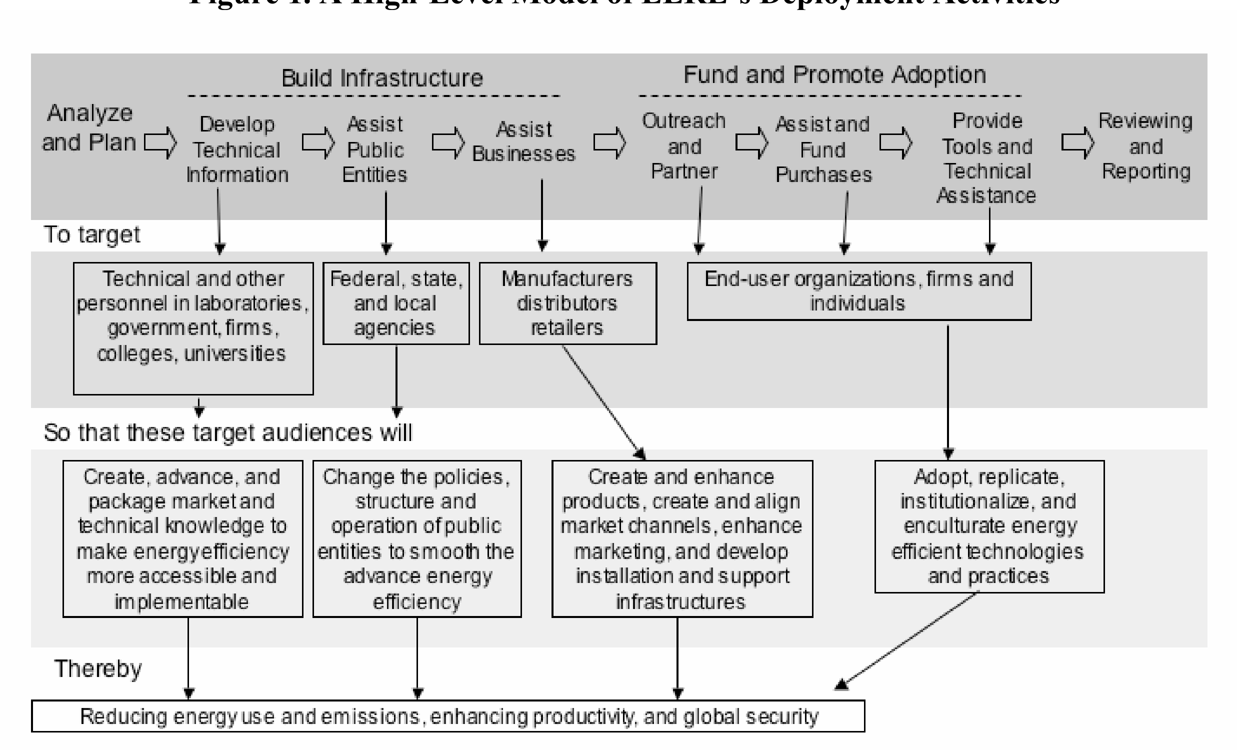
A Program Theory for Technology Diffusion

The Evaluation Framework uses logic models to build a concise description of a technology deployment program's performance path from activities to outcomes.[2] The research design focuses on the part of the logic model that explains what the customer (i.e., the target audience) is going to do in response to the product or service that the program provides. It is in this part of the performance spectrum where changes to behavior or practices that reduce energy demand occur, the availability and accessibility of energy-efficient or renewable energy technologies is increased, or the policy or market infrastructure is strengthened.

By basing the evaluation research design on well-founded program theory, one can then do the following: (1) credibly explain how program actions influence target audiences to take actions that result in long-term outcomes (i.e., energy savings); (2) measure whether the program activities are effective in actually influencing the actions of target audiences; and (3) measure intermediate progress and then attribute long-term outcomes to program actions.

Figure 1 is a high-level generic model of how EERE programs produce impacts (Reed and Jordan 2006). The logic model covers all of the various market actors and is generic, so it applies to any market sector. The model also builds on a strong foundation of social science literature. Thus, this program theory is comprehensive and has a theoretical base. Programs can identify or select the parts of this logic that describe their efforts.

Figure 1. A High-Level Model of EERE's Deployment Activities



Source: Reed et al. 2005; Reed and Jordan 2006

[2] Logic models are two-dimensional graphics and supporting text that describe a sequential set of activities in the performance of a program, i.e., the resources required for these activities, the outputs, the target audience, and the short-, intermediate- and long-term outcomes (e.g., McLaughlin and Jordan 1999)

Beginning at the top of Figure 1, EERE technology deployment programs analyze and plan; build infrastructure; fund and promote the adoption of new technologies; and review, evaluate, and report. EERE programs target four domains: knowledge workers, government officials, mid- and up-stream market actors, and end-users. EERE's program delivery activities are aimed at: (1) creating and packaging knowledge and decision support tools, (2) conditioning public policies and institutions to facilitate the delivery of energy-efficient technologies, (3) removing market and information barriers to promote energy-efficient technologies and practices, and (4) influencing end-users to adopt energy-efficient technologies and practices.

The knowledge domain includes national laboratory staff who provide technical assistance to software engineers who develop whole building system audit software, or to marketing staff who design brochures, websites, or labels promoting energy efficiency. The public policy and public institutions domain (government officials) includes local, state, and federal policy makers and their staffs who influence codes and standards, zoning, etc., as well as staff of government programs who implement DOE or related deployment programs. The market domain (mid- and up-stream market players) includes all the actors on the supply side: manufacturers, distributors, retailers, technicians, and the financial community that make decisions that determine what technologies are supplied, financed, installed, and serviced. The end-user domain (end users) is made up of the consumers of the technology or practice, with all their possible tastes, preferences, and income levels.

The desired outcomes are that persons, firms, and organizations will respond to EERE activities and outputs and take desired actions that produce desired long-term outcomes (i.e., energy savings). In general, deployment programs have neither the staff nor the resources to continuously stimulate actions in these domains or to touch all possible actors. Thus, it is important to create change among the actors in these domains in such a way that actors will repeat their actions in the absence of stimulation from the program. Furthermore, it is important that actors who are not touched by programs observe the actions of those who have responded and then learn about and emulate those actions. It is through codification, standardization, replication, emulation, and the internalization of the meaning of the actions that energy-efficient technologies and practices will be incorporated into the culture.

A More Detailed Look at the Theory of Diffusion

Since EERE deployment programs are about the diffusion of a new technology (i.e., an innovation), the program logic must describe how the diffusion occurs. This is aided by joining Rogers' theory of change about the diffusion of innovations with the logic models.[3] The diffusion of innovations is a comprehensive theory that describes how ideas and technologies find their way into markets and cultures. The diffusion theory might help evaluators of EERE programs determine if any causal linkages exist between the outputs of its programs and the actions of it partners and target audiences. The theory might also help evaluators determine if the actions of the partners and target audiences result in the outcomes that meet EERE's long-term goals. Furthermore, it might help to show the role that EERE plays in the marketplace and, therefore, also show possible explanations for behaviors that are unrelated to EERE activities. If EERE programs can describe their specific theory of change, the evaluators might be able to document that events correspond to the theory. Using the theory, EERE evaluators can identify,

[3] Diffusion of innovations is a term that came to prominence with the publication of the first edition of Everett M. Roger's book, *Diffusion of Innovations*, in 1962. Rogers' book is now in its fifth edition (Rogers 2003).

measure, and track outcomes, so that the linkages between EERE program and outcomes such as energy savings and market effects[4] can be accurately measured.

Rogers' theory of change provides a set of hypothesized concepts and linkages that can be used as a guide when describing how outputs generate the desired long-term results. The basic components of Rogers' diffusion of innovation theory are listed in Table 1 and are summarized as five models in Reed (2005). Very briefly, Rogers describes five stages in the adoption or diffusion of a technology, and these stages take place within a market and socio-economic culture that influences the movement from stage to stage. For instance, the characteristics of the technology (product) influence diffusion, a key point for EERE programs where the program activities are often directed at product characteristics (e.g., green power marketing programs try to make the amount of power from renewable energy observable in customer electric bills). New technologies often appear complex until people are trained to use them. Different groups adopt technologies at different rates of diffusion over time. The rate of diffusion is affected by differences in characteristics of the technology and communication about the technology.

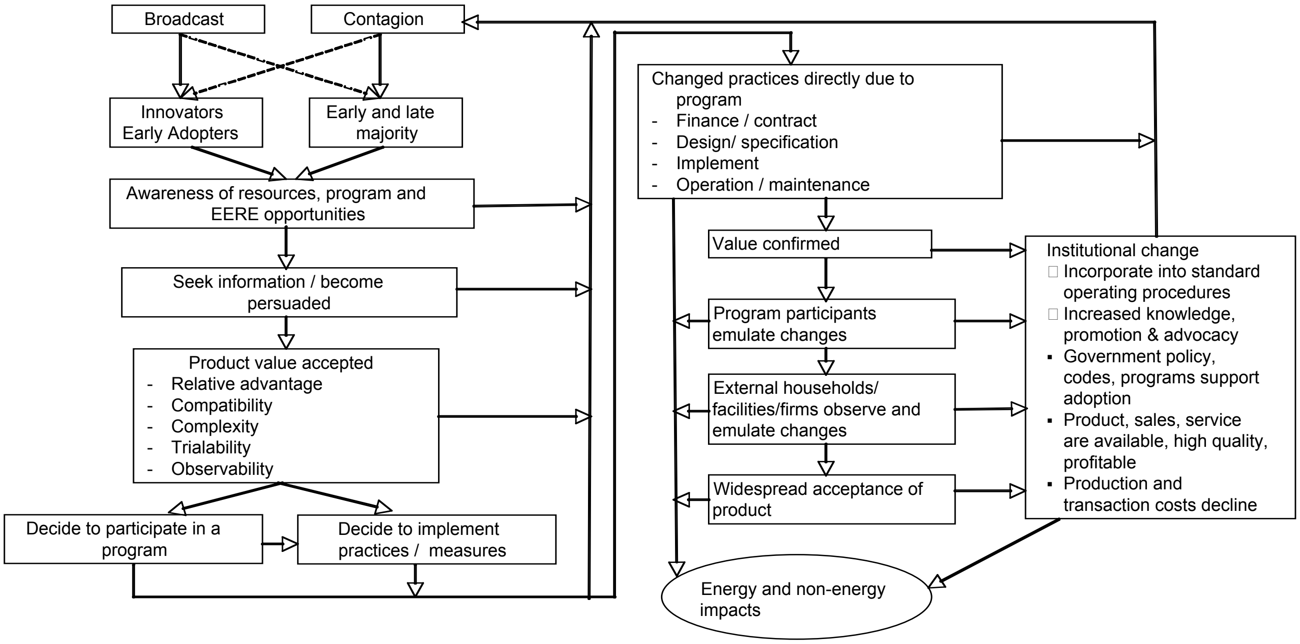
Table 1. Diffusion of Innovation Models (based on Rogers (2003))

Diffusion process model. Diffusion stages are: Awareness, Information/persuasion, Decision, Implementation, and Confirmation. The Evaluation Framework adds a 'Sustainability' stage to Rogers' five diffusion stages.
Physical, social and cultural environment model. Aspects of the market environment include market structure, market segments, prior practice, culture and norms, innovativeness, etc.
Product characteristics model. The five characteristics of technologies or innovations that determine whether the adoption will be relatively easy or difficult are: Relative advantage, Compatibility, Complexity, Trialability, Observability.
Characteristics of adopters model. The five types of adopters are: Innovators, Early adopters, Early majority, Late majority, and Laggards.
Communications model. Two basic communication methods are 'Broadcast' and 'Contagion' (word-of-mouth).

Figure 2 represents an adaptation of Rogers' diffusion of innovation models that can be used to illustrate how target audiences link outputs to long-term outcomes. Starting in the upper left portion of the logic model, the target audience receives information through broadcast and contagion processes (i.e., word-of-mouth). Innovators and early adopters are more likely to get information through broadcast methods, while the early and late majorities are more likely to get information through contagion processes. Dotted linkages are used to remind us that the correlation between communication processes and adopter types is not perfect. As a result of the communication, members of the target audience become aware of the EERE program, its products and services, and the technologies and practices that the program is promoting. Some subset of the target audiences will find the information salient, and some of these may seek additional information, or be sufficiently convinced, that little additional information is needed for them to make a decision.

[4] Market effects refer to changes in the marketplace (other than energy savings) as a result of a program intervention.

Figure 2. Adaptation of the Diffusion of Innovations Model



(Source: Reed et al. 2005; Reed and Jordan 2006)

There is a potentially a dual information and persuasion process occurring. Members of the target audience may be seeking and evaluating information about the *products and services* provided by EERE, and they may be seeking and evaluating information about the *technologies and practices* that EERE is promoting. Members of the target audience evaluate both the products and services and the technologies and practices for relative advantage, compatibility, and complexity. The opportunity to try and observe the technologies or practices will influence their decisions as well.

Members of the target audience will decide whether or not to use EERE products and services and the technology or practices that EERE is promoting. The decision could be mixed. The decision-maker might decide to utilize EERE products and services, or they may decide to bypass the EERE products and services and just adopt the technologies or practices being promoted by EERE. For example, EERE may have a software package to analyze an industrial steam system. A potential user may accept the idea of increasing the efficiency of the steam system, but the potential user may decide not to use the EERE program software in favor of a consultant's analysis.

Implementation of the decision often lags the decision, and this lag varies by technology and market sector. Depending on the targeted market actor, the form of the implementation action can also vary quite considerably. During project implementation and after its completion, the value of the project is confirmed. Like implementation, confirmation can take many forms. It may range from the impressionistic (someone casually glancing at an energy bill to confirm that the bill has gone down) to the formal (a rigorous metering and verification study with before and after metering and analysis to confirm that energy consumption has changed). Mistaken perceptions can reinforce or undermine the decision to implement a technology or practice.

Decision-makers may try something and, if it works, may repeat the decision. The early majority adopters will emulate their peers. Thus, an innovation will diffuse to other end-users, manufacturers, retailers, consultants, builders, and householders. This difference mostly takes place by word-of-mouth.

The rest of the model shown in Figure 2 focuses on sustained institutional change and is an extension of Rogers' work. Sustainability is not necessarily a matter of just persuading more and more decision-makers to adopt a technology or practice. It involves inducing long-lasting attitudinal, structural, and cultural change. Thus, this model for change calls out a series of complementary changes such as: changes to standards and operating procedures, increased knowledge and advocacy, support for governmental changes to codes and standards, improvements in the quality and reliability of products, and reduced manufacturing costs and increased profits. If such changes are the goal of the program, these effects need to be measured. Furthermore, changes in these that significantly affect the success or failure of the program need to be assessed for credible impact evaluation, in order to rule out alternative explanations for success, or to explain that the lack of success perhaps could not be avoided.

Focus on both Measured Energy Savings and Market Effects

When the program directly interacts with the end user, such as in a rebate program, the program is certainly interested in measuring energy savings. The program is able to know exactly who is reached, how many units were put in place, and the net energy savings per unit. Many programs do not directly interact with the end user, so they need a portion of their impact evaluation to focus on obtaining the measured energy savings. For example, a program to train people on how to install a complex piece of equipment so it operates at peak efficiency will need to assess how many people that received the training implemented the practices and how many times they did so in order to establish measured savings. But the program need not stop there because there may have been more outcomes, such as the trainers training other trainers.

There are two cases for a program to consider when examining market effects: to look at temporary vs. sustained change in the targeted end user, and to look for temporary or sustained changes in the three market infrastructure domains, if any of these are targeted by the program. Examples of sustained market effects include changes in retail stocking practices of energy-efficient appliance retailers, changes in attitudes towards energy efficiency, and increased enforcement of energy efficiency codes and standards. Since any number of such market changes might have occurred during program implementation, even if they were never targeted by the program, the impact evaluation should not choose to just assess just energy savings. Savings always have to be assessed in the context of market circumstances.

Indicators and Research Questions to Collect Data about the Program Theory

The generic logic model adapted to Rogers' diffusion of innovation model can be used to generate or confirm the specific logic model for the program. The following will need to be specified by the program manager and evaluator: (1) program goals and their relationship to EERE goals, (2) technologies and markets in which the program will work, (3) targeted domains and specific actor groups to be reached, (4) major program activities and delivery mechanisms through which participants will be reached, and (5) logic of how activities will influence market actors to bring about change, making specific the generic logic of the diffusion of the new technology or practice.

Once the program has described its expected path to outcomes in a logic model, it can systematically identify an explicit set of impact questions and a set of key indicators that flow from the theory of diffusion of innovation. It also identifies intermediate measures along the way to outcomes and impacts. The general questions provided in the Evaluation Framework can be used to develop the more program-specific questions. For a given evaluation study, the questions are intended to make it easier for program implementers and evaluators to identify key measures for routine data collection or for more periodic surveys of EERE programs. The general questions and indicators of the five Rogers' models can be applied across the four domains of interest: end users, manufacturers and business, pubic entities, and knowledge. Below, we provide examples of the types of questions one would ask for assessing ultimate outcomes and impact, intermediate outcomes, and why (and how) an outcome has occurred.

Ultimate outcomes and impact (where enough time has passed):

- Energy savings - How much are the energy savings from technologies adopted due to direct program interaction or due to program-induced changes in the market infrastructure?
- Energy savings - How much are the energy savings related to the direct program activities? (Check differences before and after participation, participant vs. non-participant behavior.)
- Socio-cultural market effects - What are program-induced changes in market infrastructure (e.g., retail stocking patterns) and the resulting energy savings?
- Replication and emulation - How much energy was saved because program participants influenced non-participants?
- Sustainability - Have former program participants made permanent changes in their behavior?
- What other outcomes and impacts has the program generated?

Intermediate outcomes:

- Diffusion– To what extent have participants become aware of the technology or practice? Learned new skills?
- Product characteristics– What percent of people now say product benefits are observable?
- Adopter characteristics– Has adoption moved from innovators to people who are characterized as early adopters or the early majority?

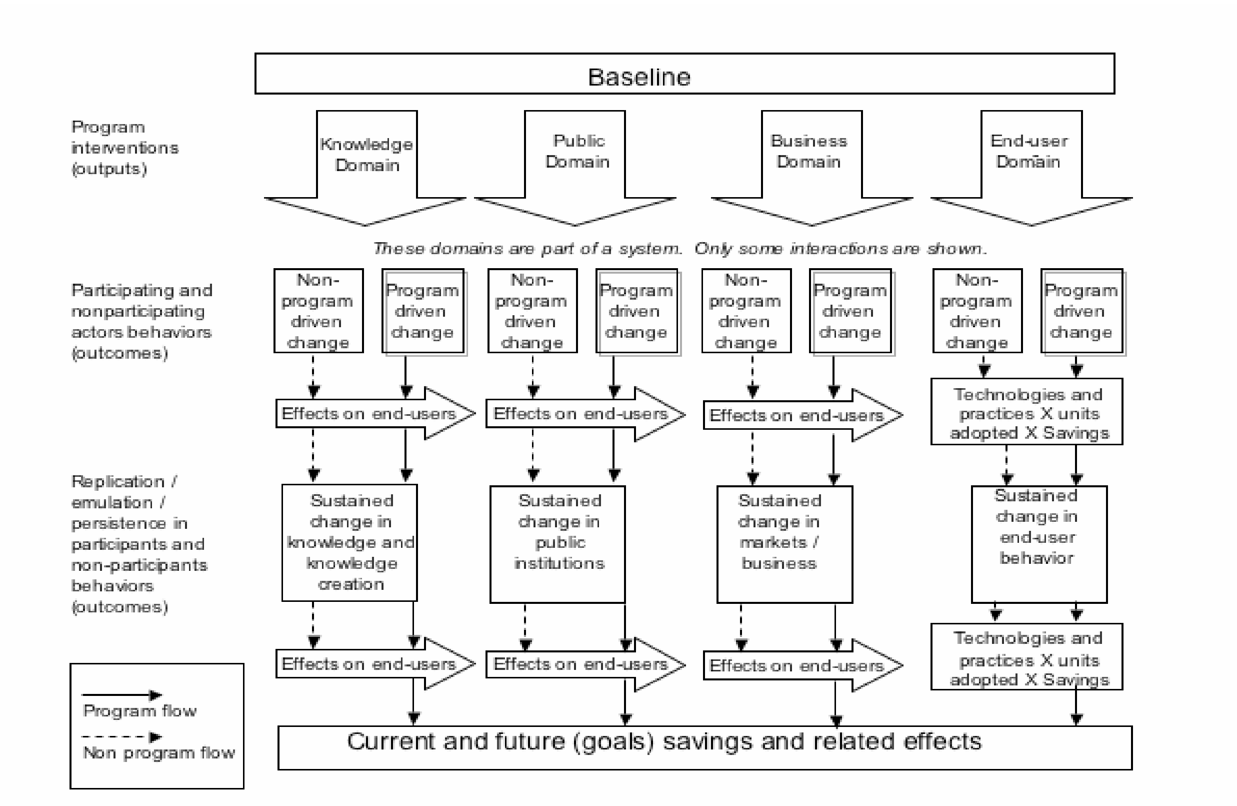
Why and how an outcome has occurred:

- Communication– How did people hear about the technology? Did one delivery mechanism work better than another?
- How has the government infrastructure changed since the program has been implemented?
- Does the use of the program's targeted technologies differ by market segment? If so, how does this affect adoption and diffusion?

Research Design to Credibly Demonstrate Impact of Deployment Programs

The objective of an impact evaluation is to measure a net change: e.g., to measure how an EERE program changed energy use – in particular, the change in customer behavior attributable to the program. As shown in Figure 3, the Evaluation Framework includes both temporary and sustained changes in behavior, changes in three domains of market infrastructure, and changes in the behavior of the end user of the technology that influences the ultimate goals of energy savings and more permanent changes in the markets related to energy use. Figure 3 also shows where non-program-induced changes need to be considered.

Figure 3. Program Impact = Current – Baseline – Non-Participant and Non-Program Driven Effects



In the case of energy impacts, the deployment program's impact is determined by measuring energy use after program participation and subtracting that energy use from baseline (pre-program) energy use for both program participants and non-participants. The difference for non-participants is then subtracted from the total to arrive at program-induced (net) energy savings.

A program's strategy for generating ultimate outcomes would be interventions in one or more of the four domains. In each domain that is relevant to the program, the impact evaluation must examine changes in behaviors of participating actors and attempt to separate program-driven change from non-program driven change. This is often done by examining the behavior of non-participants (a comparison group). If the program has targeted one of the infrastructure domains, the evaluation study needs to determine the changes that occur in that domain and how those changes affect the end users of the technology. Since the source of the changes in energy savings is adoption of more energy-efficient technologies or practices, information on this adoption is required.

Figure 3 includes another layer of assessment of the outcomes that result from the replication of technology adoption by participants, persistence of changes, and emulation of adoption or other changes that have occurred in program participants. The same types of assessment must occur at this level in terms of measuring program versus non-program-induced changes and direct and indirect effects on end users' adoption of more efficient technologies. This assessment focuses on sustained changes in each of the market domains that apply to the program.

As stated in the *EERE Guide for Managing General Program Evaluation Studies* (DOE 2006), the evaluation research design is the research strategy that permits defensible findings to be deduced from the evaluation data. It consists of:

- The questions and indicators for which data will be collected.
- Inventory of existing data and identification of data gaps.
- The method and timing by which the data will be collected.
- The populations from which the data will be collected.
- The choices of research accuracy, sampling precision and confidence level, and degree of defensibility for the results.
- The method of analysis used to produce the evaluation results.
- The method of reasoning from the results to answers to the questions.[5]

The major requirements of the research design outlined in Figure 3 are: (1) measured energy savings from both direct and indirect program effects over a specific period of time, (2) assessment of the replication and sustainability of those effects, and (3) separation of program-induced effects from non-program-induced effects (attribution). There are several methods that can be used to carry out each of these requirements. Before turning to the attribution questions, we will speak briefly about the first two requirements.

The Evaluation Framework emphasizes the collection of data directly from or about program participants on what actions they took after participating in the program, or after being affected by some indirect program-induced market infrastructure change (such as technology demonstrations, use of an energy audit software, or more capable technology services). Data would include information on specific technologies and practices, what was replaced, extent of use, etc. Sources of data include metering data, monitoring and verification data, and sales data. Other data collection methods include surveys on attitudes and behaviors. The routine collection of basic information about program participants is very useful in the Evaluation Framework, as well as the review of secondary data and the collection of primary data.

Assessment of replication and sustainability effects can also be done by observation or surveys. For example, former participants in a program and their actions need to be tracked and

[5] Two other reasonably easy-to-understand publications are available that provide guidance on preparing an evaluation design: (1) *Designing Evaluations* by the U.S. General Accounting Office, especially chapters 2 and 3 (GAO 1991), provides guidance in selecting a research design for different evaluation objectives - its terminology is slightly different than that used in the EERE Guide, but the guidance is relevant; and (2) the *California Evaluation Framework* that helps energy-efficiency evaluators choose research elements for different kinds of program evaluations (TecMarket Works Framework Team 2004). A third easy-to-understand publication provides guidance and four options for measuring energy and load impact savings: *International Performance Measurement & Verification Protocols: Concepts and Options for Determining Energy and Water Savings. Volume I* (DOE 2002b).

observed over, or after, a period of years in order to see if they have continued to practice a particular behavior and, even more importantly, if they have shared that information with others. It is also possible to observe whether or not practices become codified into standard operating procedures, building codes, etc. Some of this information can be determined by surveys, asking if practices have been repeated, shared, or recommended to others.

Attributing outcomes to the program is arguably the major challenge of an impact evaluation. Some consider the ideal test for this to be a *randomized controlled trial (RCT)* to demonstrate the actual program impact. In the trial, a random sample of people would participate in the program and a random sample of people would not participate. Then one would compare changes in energy use, behavior, or attitudes in each group after the program has been implemented. However, RCTs can only be employed under very specific circumstances. Consequently, the evaluations of energy programs typically use alternative evaluation methodologies.

The most common alternative to RCTs is the *quasi-experimental method* where program participants are compared to a comparison group of non-participants. However, the control activity (comparison group) is not randomly assigned. Under certain circumstances, well-matched comparison group studies can approach the rigor of RCTs. However, the use of comparison group studies does increase the risk of misleading results because of the difficulty in eliminating bias in the selection of the control group. For example, differences between the participant groups and control groups that affect their changes in energy use between the pre- and post-participation periods will bias the estimate of net program savings (selection bias). Awareness of this risk is crucial to the design of such evaluations.

The quasi-experimental approach has been frequently applied to the evaluation of impacts of energy programs and used in regulatory proceedings. We will mention two common methods. One fairly easy way to compare participants and non-participants is to compare changes in measured energy use in different geographical areas where the target population and market characteristics are similar but the program application (e.g., technology demonstrations) was different. Another method is to compare behaviors and program activity using time series data. If behaviors increase or decrease at the same time program activity increases or decreases, assuming some appropriate lag factors, and activity rises and falls or not as other potential causal factors rise and fall but at different times, then one can argue that program activity is causally related.

As stated previously, well-founded program theory-based evaluations provide evidence of program-induced outcomes by assessing the presence of a logical chain of sequential events from program intervention to outcomes and by looking for side effects. The Evaluation Framework recommends that the evaluation always examine alternative explanations for the measured outcomes by investigating what changes occurred in the four domains that could explain the outcomes. For example, did energy prices change dramatically? Or were new regulatory requirements implemented? This can be determined through secondary data collection or from interviews or surveys with program partners, implementers, and participants.

Conclusions

In this paper, we have described a theory-based approach to impact evaluation that could be used by DOE's deployment programs for evaluating energy savings and market effects with credible attribution of impacts. Evaluators could use the framework to understand where to look within the program logic for measured outcomes such as sales or adopted technologies and practices. Finally, by using the framework a causal link between the program and its outcomes can be tested and alternative explanations investigated.

References

[DOE] U.S. Department of Energy. 2002a. *Strategic Plan*. Washington, D.C.: U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy.

[DOE] U.S. Department of Energy. 2002b. *International Performance Measurement & Verification Protocols: Concepts and Options for Determining Energy and Water Savings. Volume I*. DOE/GO-102000-1132. Washington, D.C.: U.S. Department of Energy. Available online at: www.ipmvp.org.

[DOE] U.S. Department of Energy. 2006. *Guide to Managing General Evaluation Studies.* Washington, D.C.: U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy.

[DOE] U.S. Department of Energy. Forthcoming. *Framework for Assessing the Impact of Deployment Programs*. Washington, D.C.: U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy.

[GAO] General Accounting Office. 1991. *Designing Evaluations.* GAO/PEMD 10.1.4. Washington, D.C.: General Accounting Office.

McLaughlin, J. A. and G. B. Jordan. 1999. "Logic Models: A Tool for Telling Your Program's Performance Story." *Evaluation and Program Planning*, Vol. 22, No. 1, February.

[OMB] Office of Management and Budget. 2005. *Guidance for Completing the Program Assessment Rating Tool (PART)*. Washington, D.C.: Office of Management and Budget.

Reed, J. 2005. "Diffusion of Innovation: An Overview." Working paper prepared for U.S. DOE/EERE, February.

Reed, J. and G Jordan. 2006. "Generic Logic Models for Delivery and Diffusion of Innovation," paper submitted for publication.

Reed, J., J. Mortensen, and G. Jordan. 2005. "Generic Logic Models for Federal Program Delivery and Diffusion of Innovation." Presentation at the Joint American Evaluation Association/Canadian Evaluation Society Conference, October 26-29, 2005, Toronto, Canada.

Rogers, E. 2003. *Diffusion of Innovations*. Free Press, NY.

TecMarket Works Framework Team. 2004. *The California Evaluation Framework*, Oregon, WI; TecMarket Works.