An Overview and Example

Impact Evaluation Framework For Technology Deployment Programs

An approach for quantifying retrospective energy savings, clean energy advances, and market effects

John H. Reed
Innovologie LLC

Gretchen Jordan
Sandia National Laboratories

Edward Vine
Lawrence Berkeley National Laboratory

July 2007

Introduction and Background

The document briefly describes a framework for evaluating the "retrospective" impact of technology deployment programs and provides an example of its use. The framework was developed for the US Department of Energy's (US DOE) Office of Energy Efficiency and Renewable Energy (EERE) but potentially can be applied to most deployment programs.[1] This walk through of the seven-step impact framework process and the illustrative application provide an overview and introduction to the more detailed framework document (Reed, Jordan, Vine 2007). The framework is specifically designed to assist program managers and evaluators in Federal, state, and local governments.

Program managers in the Federal government are being asked by their management, the Office of Management and Budget (OMB), and Congress to demonstrate that the technologies and practices they are developing and their deployment activities are cost effective and produce, or have the potential to produce, the expected outcomes, such as energy savings, emission reductions, and enhanced national security. Evaluating the impact of deployment programs using this framework will help program managers obtain information to improve their programs and to effectively communicate the full range of results realized through program efforts.

[1] EERE is responsible for developing and deploying energy efficiency and renewable energy technologies. This multi-program agency addresses wind, solar, hydrogen, and biomass technologies, the efficiency and use of energy in residential, commercial buildings, industrial facilities, the use of efficient and clean energy for vehicles, reduction of energy use in government buildings and low-income residences, appliance standards, and more.

U.S. Department of Energy
Energy Efficiency and Renewable Energy
Bringing you a prosperous future where energy is clean, abundant, reliable, and affordable

A Generic Program Theory for Technology Deployment Activities

Figure 1 provides a generic high level description of how technology deployment programs typically produce impacts. At the top level, programs analyze, plan, and build market infrastructure; fund and promote the adoption of new technologies; and review, evaluate, and report. Programs target knowledge workers, public entities, market players, and end-users or some combination of these. Program delivery activities usually aim to cause one of four things:

- Create and package knowledge to make it accessible to users.
- Condition public policies and institutions to facilitate the delivery of energy efficient and renewable technologies.
- Reinforce the market to promote energy efficient and clean energy technologies and practices.
- Influence end-users to adopt energy efficient and clean energy technologies and practices.

In other words, programs, in varying degrees, conduct activities to influence audiences in four domains:

- Scientists, researchers, and engineers in the **knowledge** domain.
- State and local agencies, nongovernmental organizations, and utilities in the **public entities** domain.
- Manufacturers, distributors, retailers, engineering and design professionals, and installers in the **business** domain.
- End-users of technologies and practices in the **end-user** domain.

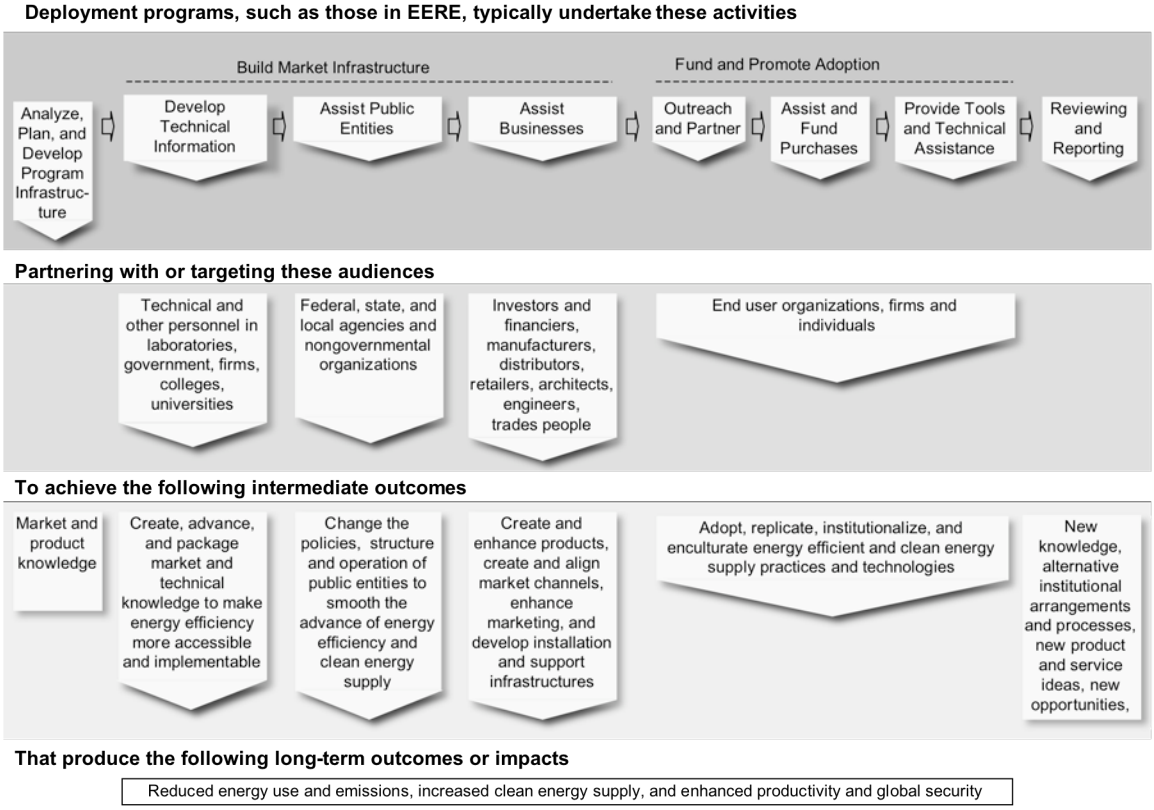


Figure 1. A High Level Generic Logic Model for Deployment Activities

The desired outcomes are for people, firms, and organizations in these domains to respond to program activities and outputs and take steps to produce the desired impacts. 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 they repeat and sustain their actions in the absence of stimulation from the program.

Why an Impact Framework?

The impact evaluation framework is designed to assist program managers and their evaluation contractors develop powerful and meaningful impact evaluations that:

- Help refine their programs,
- Increase program effectiveness,
- Help identify and make decisions about ineffective program elements, and
- Develop credible evidence to demonstrate their accomplishments and defend against skeptics.

This impact evaluation framework focuses on linking program outputs to short-term and long-term outcomes (impacts), measuring partner and target audience response to program outputs, designing sound evaluations, and taking credit for the program effects that are attributable to the program. It provides tools to step users through the development of an impact evaluation and especially to identify the linkages among outputs and short- and long-term outcomes. The framework uses well established principles from social science to clearly identify what needs to be measured, develop strong evaluation designs, and harness programs' existing data collection activities to obtain evaluation data. The framework is designed to produce common evaluation questions, measures, definitions, data collection instruments, and protocols that can be applied systematically across multiple deployment programs within a given organization.

The Impact Evaluation Framework

The impact evaluation framework is comprised of the seven modules shown in Figure 2. By following the steps, program managers and

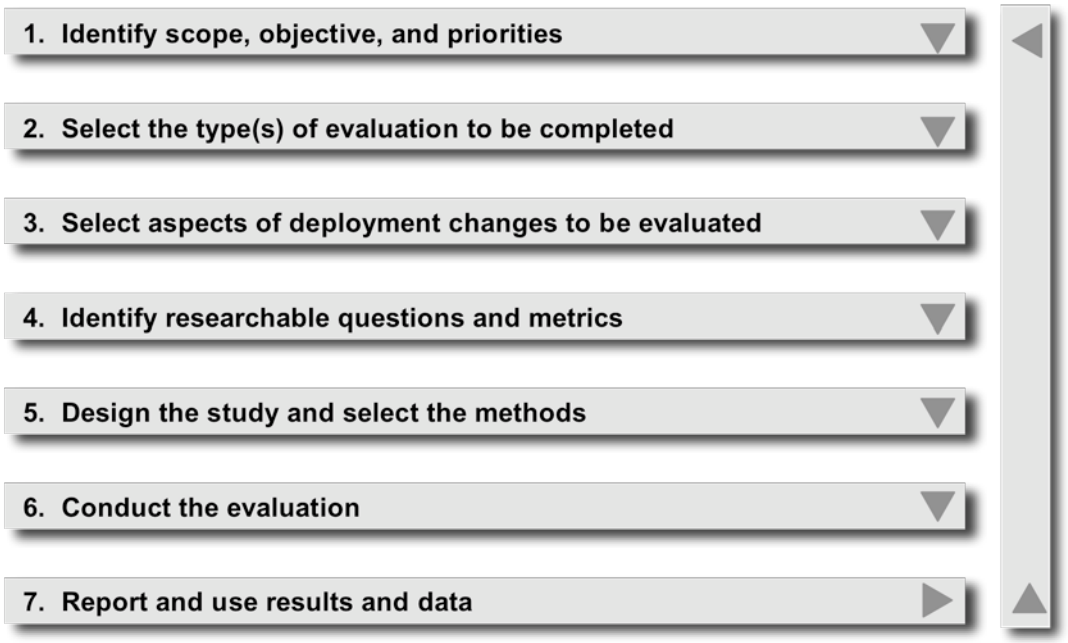


Figure 2. Overview of the Impact Evaluation Framework

evaluators can design and implement an impact evaluation of a deployment program.

The framework can be flexibly applied in a broad range of situations. It can be used to design an evaluation of a single program element, multiple elements within a program, or multiple programs. It can be used to examine one or more delivery channels and/or one or more communications channels or some combination. The framework can be used to guide impact evaluations of end-user behaviors; the behaviors of intermediate market players such as investors, manufacturers, distributors, retailers, and design professionals; the actions of governmental and nongovernmental organizations; and actors and organizations in the knowledge industry. It is appropriate for use with resource acquisition programs as well as market transformation programs. It is applicable across multiple sectors – residential, commercial, industrial, transportation, and government. The framework is generic so that it can be customized to for a particular deployment program and evaluation need.

The impact framework integrates diffusion of innovations (Rogers, 2003) with program theory and logic models to assist program managers and evaluators develop scientifically sound descriptions of programs, design evaluations, and collect and analyze data, and report results.

A key step in an evaluation is to identify the domains that are targeted or served by a program, the actors within the domains, and the program and non-program related (external factors) interactions within and among domains that influence program outcomes and impacts. The key to an effective and useful impact evaluation is to identify and justify:

- The outcomes in each domain,
- The measures that are relevant for assessing those outcomes, and,
- The external factors that potentially provide an alternative explanation for the outcome or inhibit the realization of the outcome.

Integrating Deployment Program Logic and the Diffusion of Innovations Story Line

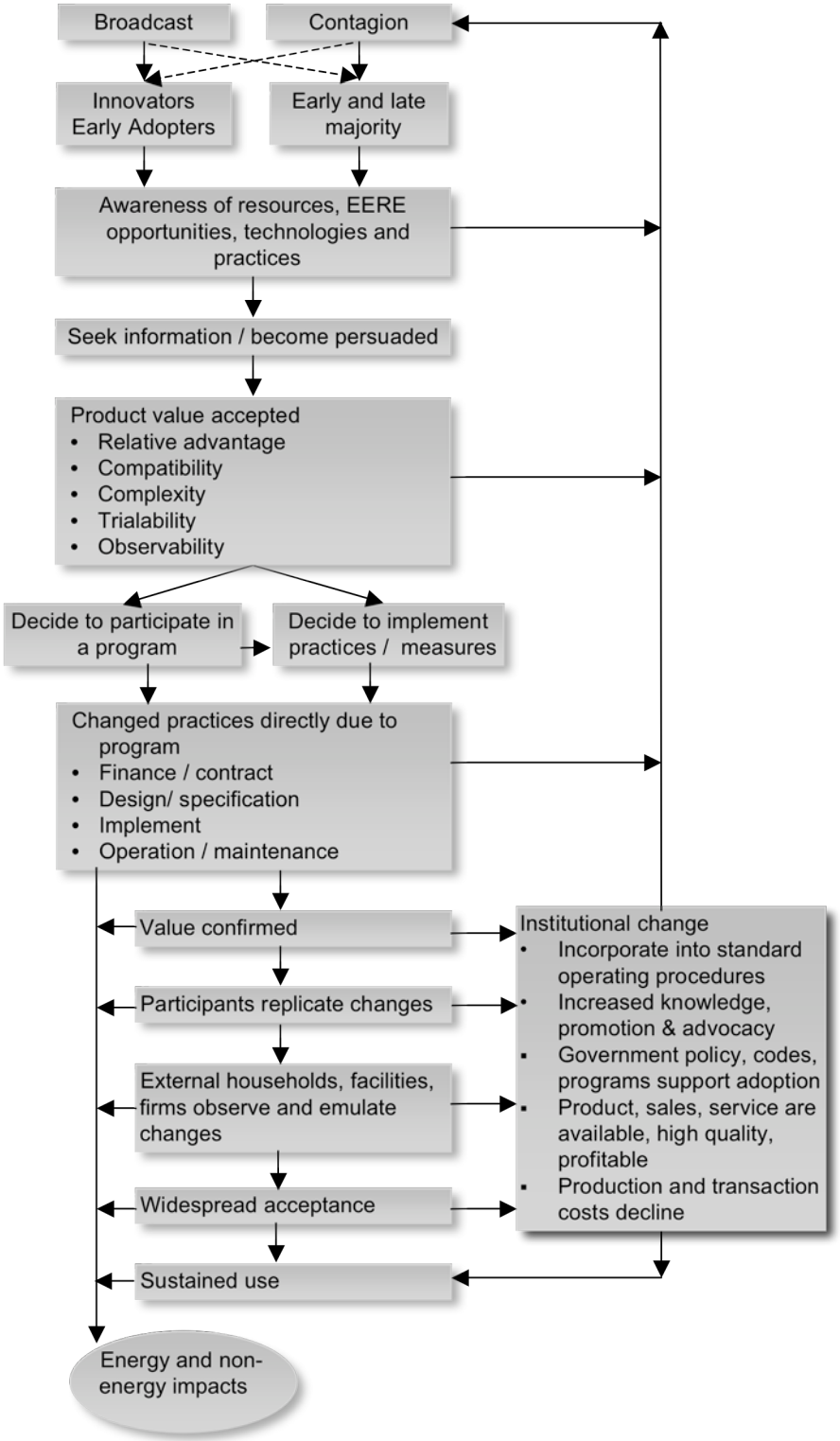
This impact evaluation framework uses diffusion of innovations as a storyboard to assist evaluators and program managers in developing well focused and useful evaluations.

The diffusion of innovations is a broad systems theory:

- That is capable of providing a detailed and unified story line across domains and programs, and
- That captures in a powerful and useful way the essence of what we know about change and implementing change from the social sciences.

Figure 3 is a schematic of the diffusion of innovations theory that represents the basic elements to its story line:

1. The socio-cultural environment — Firms and people operate in a socio-cultural environment that can aid or impede the spread of a technology or practice. In order to target the right actors, understand points where resistance to innovation may occur, and identify where new connections and institutions may be needed, it is important for deployment programs to know who the players are, the interconnections among the players, the cultural dynamics, and the rules of the system. The socio-cultural environment surrounds the elements shown in the figure.
2. The diffusion process — When deciding to adopt new practices or technologies, organizations, firms, and individuals pass through a series of stages — awareness, information and persuasion,



Source: John H. Reed and Gretchen Jordan 2005

Figure 3. An Adaptation of the Diffusion of Innovations Model

decision, implementation, and confirmation. It is possible to track adoption by examining the rapidity with which potential adopters pass through the stages.

3. Adopter characteristics — The characteristics of firms and organizations influence how quickly they adopt technologies and practices. The classic adopter groups are innovators, early adopters, early majority, later majority, and laggards.
4. Product characteristics — Target audiences find a new technology or practice more or less appealing depending on its relative advantage, compatibility, complexity, trialability, and observability.[1]
5. Actors or players within the environment send and receive communications by broadcast mechanisms (a single source to many recipients, e.g. radio, television, newspapers, e-mail blasts) or contagion processes (one-to-one transmission of information) through formal and informal social and professional networks.

Based on the general logic model (Figure 1), the framework includes four domain specific logic models, one for each one of the four domains identified above. For example, the adoption process for end-users is described in generic form in the end-user logic model. Impact framework users can use the generic form to analyze and understand how end-users of a specific program might adopt a technology or practice. If for example the targeted end-users are school districts and the goal is to get them to build Leadership in Energy and Environmental Design (LEED) certified buildings, then one can use the framework to begin to systematically analyze and describe how school districts might adopt LEED standards. If the target is business entities, one can use the generic logic model for business entities to develop the story for how business entities might begin to decide to develop, manufacture, distribute, and maintain energy efficient products. Because there are several kinds of business entities - manufacturers, distributors, service professionals, retailers, installers, etc. - the logic model might be applied several times.

By using a slightly modified version of the same story over and over again, the result is likely to be that the same types of measures will be identified and used. This potentially increases the opportunity to aggregate information within and across programs.

To increase its value, the impact framework also includes an extensive set of generic evaluation questions and measures that can be used as a guide for developing program specific questions. The idea is that once the focus of the evaluation is clear, one can work through the list of questions and measures while developing a list of evaluation questions and performance measures or one can take an existing set of questions and measures and compare them to the generic set to see if the range of evaluation issues has been adequately covered. Because the evaluation issues are generic they can be used in each of the domains or with specific target groups.

The Application of the Impact Framework to a Deployment Program: the Federal Energy Management Program (FEMP)

The application of the impact evaluation framework is illustrated with a brief walk through of the seven-step process (modules) in

[1]
- Relative advantage, that is, the perceived benefits and disbenefits of the technology or practice relative to other products or to not having the technology or practice
- Compatibility of the product with the cultural, social, and/or physical systems in which it is to be used
- Complexity or the difficulty of understanding, producing, installing, selling the product
- Trialability, that is, the opportunity to try or use the product
- Observability, that is, being able to touch, feel the product, to observe its operation or to see its effects.

Figure 2. Some of the material in the walk through is drawn from a 2001 evaluation of the Federal Energy Management Program (FEMP) (Hall, et. al., 2002). That evaluation predates this impact framework. If that evaluation were to be done now based on this framework, it would include many of the same elements but it would likely include some additional ones and be designed and implemented slightly differently.

Module 1: identify the scope, the objective, and the priorities of the evaluation

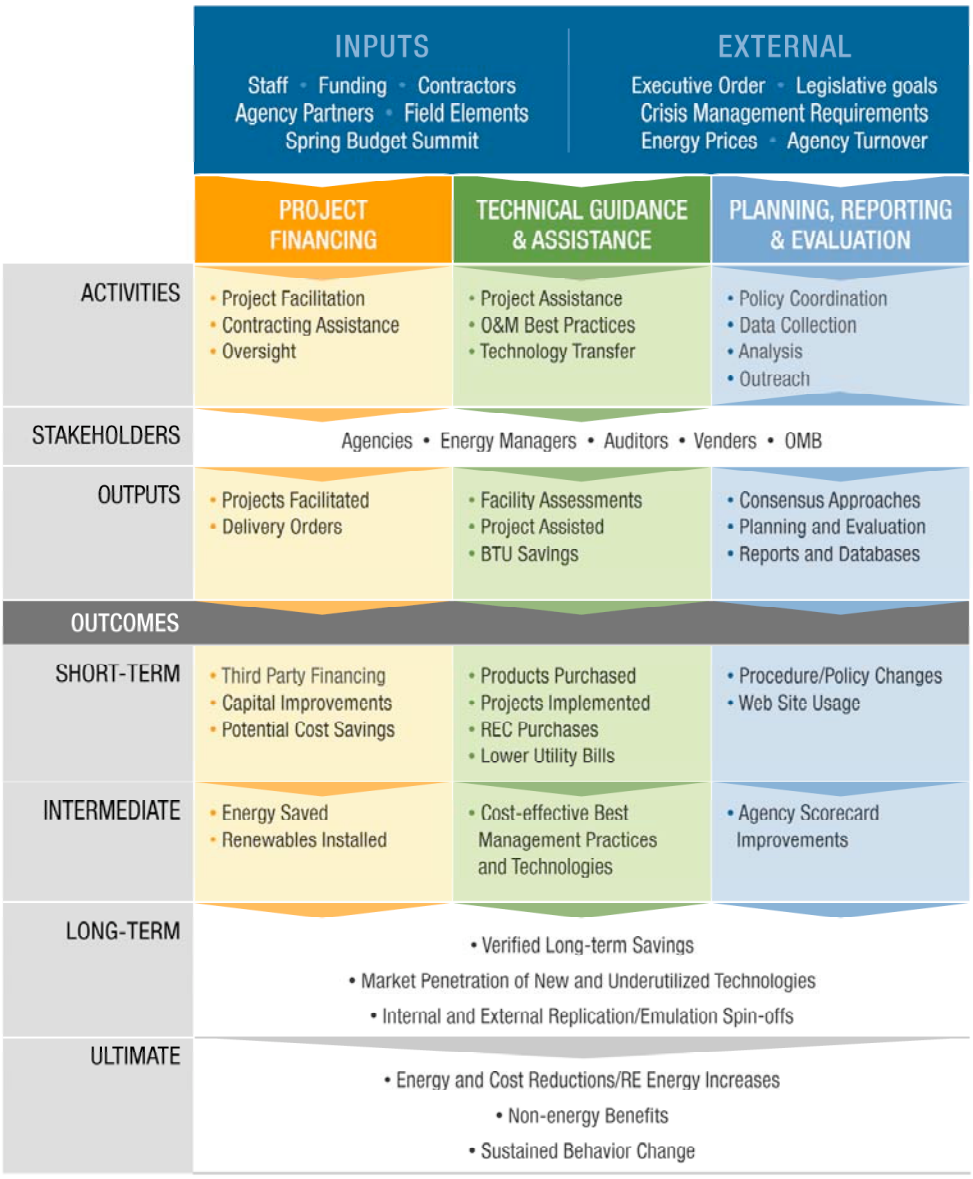
FEMP assists other Federal agencies to implement programs and measures in new and existing facilities to reduce energy and water consumption, and to develop sources of clean energy. An important part of this module is clarifying program goals. FEMP's activities result in direct savings, but the actions that it causes are also likely to be replicated and emulated so it is important to capture those effects as well. FEMP assists other agencies, and those agencies are likely to want to take credit for the actual energy and cost savings and production from clean energy sources. For the purposes of this illustration, FEMP could frame its overall goals in terms of water and energy savings and production from clean energy installations that it has influenced.

Another important step in module one is to develop a good description of the program preferably using a tool such as a logic model. Figure 4 is a 2005 FEMP logic model that shows selected elements of the FEMP program.[2] This logic model provides a potential starting point for updating the description of the program. The high-level program logic model in the full framework document might be referenced to see what could be changed. For example, an activity describing program infrastructure development aimed at developing case studies, purchasing information, promotional material and other related types of activities might be added. An activity describing FEMP promotional efforts with Federal agencies might also be added. Also, more detail on the expected pathways through the short and intermediate outcome space in the logic model might be added.

A third step in this module is to identify the scope of the evaluation. For purposes of illustration, the evaluation might be limited to the project financing and the technical guidance and assistance activities. At this writing, these are key elements in the FEMP program. Even if the evaluation is limited to these activities, other activities cannot be ignored. For example, the impacts of project financing may depend on how well FEMP provides assistance. It may be necessary to assess awareness to fully understand the impacts.

Also, it is important to examine the domains to be targeted: knowledge, public entity, business, and/or end-use. In FEMP's case, the knowledge domain represents the activities and outputs of national laboratories and contractors that provide FEMP with software, case studies, publications, product data, and other information. For the public domain, the key public entities for FEMP are other Federal agencies such as the US Department of Defense, and the US Department of Agriculture, and others. In terms of the impact evaluation, the social and cultural environment in other Federal agencies may be more or less favorable to different forms of financing, so that it may be important to assess those environments as part of the impact evaluation. The business domain includes FEMP's private sector partners that provide performance contracting, auditing, and design assistance services, and, in this instance, utilities that may have incentives available for efficiency upgrades or purchases. Finally, it is important to identify the users and potential users of FEMP's technical assistance and financing activities. End-users are those individuals within Federal agencies, including the Government Services Administration (GSA), who build government

[2] FEMP is currently making enhancements to its program and is expected to make further modifications to its logic model.



Source: FEMP Multi-year Program Plan September 28, 2005.

Figure 4. Logic Model for the Federal Energy Management Program

buildings, lease space, or operate and manage space.

In the 2001 evaluation, FEMP wanted a broad understanding of the awareness, use, and effectiveness of FEMP services and FEMP chose to examine in more detail two delivery channels, Energy Savings Performance Contracts (ESPC) and the SAVEnergy Audits because those were the dominant activities in the project financing and technical guidance and assistance arenas.

The selection of delivery channels to be evaluated is one of the sub-steps in Module One. A well-done logic model can assist in sorting out many of the issues concerning evaluation scope, objectives, and priorities.

Module 2: Select the type(s) of evaluation to be completed

Module two is designed to assist program managers and evaluators in deciding upon a type of evaluation. The types of evaluation are market evaluation; process or output evaluation; market effects, gross impacts, net impacts and cost benefit analysis. (See Text Box of "Evaluation Types" in Section 3.2 of the full report on Impact Framework for Technology Deployment) Being clear about the type of evaluation will help shape the expectations of stakeholders about the results and help to define the types of data that will be collected.

The 2001 FEMP customer evaluation included elements of market and process evaluation and focused more specifically on the effects (market effects) of the uptake of ESPC and the SAVEnergy Audit. The assessment of gross and net energy impacts and cost benefit analysis were not pursued.

Module 3: Select the aspects of deployment changes to be evaluated

This module uses the diffusion of innovations model to assist program managers and evaluators in defining the specifics of outcomes and identifying what should be evaluated. The 2001 FEMP study used three of the six aspects of diffusion of innovations: diffusion stages, communications characteristics, and replication and emulation. The illustration follows that lead. However, if we were reconfiguring the program's services or had additional goals for the evaluation, it might be useful to categorize participants and nonparticipants by adopter type, examine the characteristics of the products (ESPC and SAVEnergy Audit), and/or conduct an in-depth analysis of the socio-cultural environment.

Module 4: Identify researchable questions and metrics

This module is designed to assist evaluators and program managers in defining research issues and developing measures. The three aspects of change identified for the FEMP program suggest three important evaluation questions (based on the 2001 study).

1. Did FEMP participants adopt or accelerate the adoption and implementation of financing energy efficiency measures (performance contracting) or have an energy audit and implement the measures or implement them more quickly than nonparticipants?
2. Were participants more likely to have heard about the program from others (through contagion or from broadcast methods?
3. Compared to nonparticipants, did a higher percentage of participants confirm and replicate their activities. Did they tell others who then emulated their actions?

As noted above, Section 5 of the main impact framework report contains a general list of questions and measures that can be used to design more specific detailed questions and metrics. Table 1 shows a set of questions and measures that can be used to transform the three questions into more specific ones.

Consulting the questions in section 5 of the full impact framework report suggests the following kinds of questions for replication and emulation.

- Are people/firms replicating the use of the financing mechanism/ESPC?
- Have people told other people or firms about financing mechanisms/ESPC?

The measures associated with these questions are the percentage of participants that are replicating their use of the financing mecha-

nism/ESPC and telling others about it. The savings from replications are another effect. Comparison is greater continuing use of financing mechanisms/ESPC by participants than nonparticipants and more communication to others about financing mechanisms/ESPC among ESPC participants.

Module 5: Design the study and select the methods

This module is designed to assist evaluators and program managers to develop a suitable study design. The goal is to devise a way to determine the effects attributable to the program.

If it were possible to randomly assign and select employees in Federal agencies to participate or not participate in programs like ESPC or the SAVEnergy Audits and to "isolate" them, the actions of the group exposed to ESPC or the SAVEnergy Audits could be compared to those who were not. Random assignment helps to assure that employees in both groups share non-program related characteristics that might otherwise influence participation.

However, random assignment is not possible. Agencies and employees self-select for participation in program activities like ESPC or the SAVEnergy Audits. Thus, the participants and nonparticipants may have different characteristics, some of which could account for decisions to adopt. It is still important to randomly sample program participants and non-participants from the respective populations to make sure that there is no bias in terms of selection for participation in the study. However, additional steps may be required to minimize any effects of self-selection.

A key evaluation challenge is developing a population from which to sample. For instance, in 2001 there was no government wide list of individuals responsible for managing or making decisions about buildings. Thus, a list had to be constructed from available lists. It is likely that nonparticipants were underrepresented in the lists. Further, the participation status of people on the list had to verified using screening questions at the beginning of the survey.

Specifying the approach to data collection is part of this module. For the 2001 FEMP study, the primary form of analysis was to compare changes in key measures among participants, aware non-participants and unaware non-participants. For a study such as this one, a participant and nonparticipant survey is a good choice. This requires designing and testing a questionnaire. A good source on questionnaire design is Don Dillman (1978, 2007). Also, users will want to consult the *EERE Guide for Managing General Program Evaluation Studies* (Barnes and Jordan, 2006).

Table 1. Measures and Impacts or Effects

Measures	Impact or Effect
Awareness of financing mechanisms/ ESPC	Change in percentage of participants/nonparticipants aware of financing mechanisms/ESPC
Information seeking financing mechanisms/ESPC	Change in percentage of participants/ nonparticipants seeking information about or considering financing mechanisms/ESPC
Decision to implement financing mechanisms/ESPC	Change in percentage of participants/nonparticipants deciding to implement financing mechanisms/ESPC
Implementing financing mechanisms/ ESPC	Change in percentage of participants/nonparticipants implementing

Module 6: Conduct the evaluation

An independent contractor is likely to be hired to conduct an evaluation like the one done for FEMP. The conduct of the evaluation should be based on a work plan developed by the evaluation contractor in conjunction with the program manager. There should be an independent review committee to examine the work plan, survey instruments, and interim and final reports. The evaluation contractor might use in-house resources or hire a firm to design and conduct the surveys.

Module 7: Report and use results and data

This module addresses reporting. Table 2 is an example from the FEMP study of how the effects might be reported for the different diffusion stages.

Table 2 shows the results for ESPC participants, ESPC aware non-participants, and ESPC unaware non-participants. For participants, the comparison is the stage of adoption for non-ESPC performance contracting prior to involvement with FEMP and subsequent to involvement with FEMP. The table shows that subsequent to FEMP involvement, ESPC participants moved to decision, implementation, and confirmation stages. ESPC aware non-participants became aware of ESPC but changed little in comparison to pre-FEMP involvement with performance contracting like mechanisms. People with no knowledge of ESPC were largely unaware of performance contracting like finance mechanisms, and only a few used them.

References

Bandura, Albert J., "On Integrating Social Cognitive and Social Diffusion Theories." In Arvind Singhal and James W. Dearing, eds. *Communication of Innovations: A Journey with Everett Rogers*, Thousand Oaks: Sage Publications, 2006. pp 111-135.

Barnes, H, and Gretchen Jordan, 2006. EERE Guide for Managing General Program Evaluation Studies. Washington DC: US Department of Energy, 2005.

Dillman, Donald A. Mail and Telephone Surveys: The Total Design Method. New York: John Wiley and Sons, Inc., 1978.

Dillman, Don A. Mail and Internet Surveys: *The Tailored Design, Second Edition—2007 Update*. New York: John Wiley. New York: John Wiley and Sons, Inc., 2007.

Table 2. Movement of FEMP ESPC Participants and Non-Participants through the Adoption Cycle

	Percent of participants (N=101)			Percent of aware non-participants (N=117)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Diffusion stage	Before hearing about FEMP	Since involvement with FEMP	Change from before to after	Before hearing about FEMP	Since involvement with FEMP	Change from before to after	Percent of unaware non-participants
Unaware	24	0	-24	21	0	-21	63
Aware	27	10	-17	31	40	+9	24
Persuasion	12	7	-5	9	10	+1	5
Decision - no	10	7	-3	14	16	+2	1
Decision - yes	3	21	+18	8	9	+1	1
Implementation	7	24	+17	4	6	+2	4
Confirmation	18	32	+14	14	18	+4	2

Hall, Nick, John Reed, Tom Talerico, Jeff Riggert, Andrew Oh, and Gretchen Jordan. 2001 *FEMP Customer Survey Study Report*. Oregon WI: TecMRKT Works, 2002.

Reed, John, Gretchen Jordan and Edward Vine. *Impact Evaluation Framework for Technology Deployment Programs: An Approach for Quantifying Retrospective Energy Savings, Clean Energy Advance, and Market Effects*. Washington DC: US Department of Energy, 2007. (Main Report)

Rogers, Everett. *Diffusion of Innovations*, Fifth Edition. New York: New York Free Press, 2003.

Electronic Versions

A complete electronic version of *Impact Evaluation Framework for Technology Deployment Programs: An Approach for Quantifying Retrospective Energy Savings, Clean Energy Advances, and Market Effects* can be found at http://www.eere.energy.gov/ba/pba/km_portal/docs/pdf/2007/impact_framework_tech_deploy_2007_main.pdf. An electronic copy of this document can be found at http://www.eere.energy.gov/ba/pba/km_portal/docs/pdf/2007/impact_framework_tech_deploy_2007_overview.pdf.

Acknowledgements

This work was completed for the U.S. DOE Office of Energy Efficiency and Renewable Energy (EERE) by Sandia National Laboratories, Albuquerque, New Mexico, USA, under Contract DE-AC04-94AL8500. Sandia is operated by the Sandia Corporation, a subsidiary of Lockheed Martin Corporation.

John H. Reed (Innovologie LLC), Gretchen Jordan (Sandia National Laboratories), and Edward Vine (Lawrence Berkeley National Laboratory) developed the impact framework and prepared the document. Jeff Dowd, DOE's Office of Energy Efficiency and Renewable Energy, Office of Planning, Budget and Analysis (PBA), was the DOE Project Manager.

Charles Bailey, Moria Morrissey, and Jeff Riggert of the Innovologie Team assisted in creating this document. The authors also thank Yaw Agyeman of TMS Inc. for assistance in managing the review process and preparing the document.

Other US DOE EERE Publications on Evaluation

1. *Impact Evaluation Framework for Technology Deployment Programs: An Approach for Quantifying Retrospective Energy Savings, Clean Energy Advances, and Market Effects*. 2007. (Main Report) Prepared by John H. Reed, Gretchen Jordan and Ed Vine. Washington DC: US Department of Energy. http://www.eere.energy.gov/ba/pba/km_portal/docs/pdf/2007/impact_framework_tech_deploy_2007_main.pdf.
2. *Overview of Evaluation Methods for R&D Programs.* 2007. Prepared by: Rosalie Reugg and Gretchen Jordan. Washington DC: US Department of Energy. http://www.eere.energy.gov/ba/pba/km_portal/docs/pdf/2007/evaluation_methods_r_and_d.pdf
3. *EERE Guide for Managing General Program Evaluation Studies*. 2006. Prepared by: Harley Barnes and Gretchen Jordan. Washington, DC: US Department of Energy. http://www.eere.energy.gov/ba/pba/km_portal/docs/pdf/2007/evaluation_mgmt_guide_final_2006.pdf
4. *EERE Peer Review Guide*. 2004. Prepared by: Sam Baldwin, Jim Daley, Jeff Dowd, David Howell, John Ryan, Alan Schroeder, Frank Wilkins, and Gretchen Jordan. Washington DC: US Department of Energy. http://www1.eere.energy.gov/ba/pdfs/2004peerreviewguide.pdf

Notice