LBL-33601
UC-350

Lawrence Berkeley Laboratory

UNIVERSITY OF CALIFORNIA

ENERGY & ENVIRONMENT DIVISION

The Human Dimension of Program Evaluation

E.L. Vine

May 1993

ENERGY & ENVIRONMENT DIVISION

REFERENCE COPY
Does Not
Circulate
Bldg. 50 Library.
LBL-33601
Copy 1

Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF00098

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. Neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or The Regents of the University of California and shall not be used for advertising or product endorsement purposes.

This report has been reproduced directly
from the best available copy.

Available to DOE and DOE Contractors
from the Office of Scientific and Technical Information
P.O. Box 62, Oak Ridge, TN 37831
Prices available from (615) 576-8401, FTS 626-8401

Available to the public from the
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road, Springfield, VA 22161

Lawrence Berkeley Laboratory is an equal opportunity employer.

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

LBL-33601
UC-350

THE HUMAN DIMENSION OF PROGRAM EVALUATION

Edward L. Vine

Energy Analysis Program
Energy and Environment Division
Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

May 1993

This work was supported by the Assistant Secretary for Conservation and Renewable Energy, Office of Utility Technologies of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

recycled paper

The Human Dimension of Program Evaluation

Edward L. Vine

Energy Analysis Program
Energy & Environment Division
Lawrence Berkeley Laboratory
One Cyclotron Road
Berkeley, CA 94720

ABSTRACT

Social science issues play an important role in the evaluation of demand-side management (DSM) programs. In the very early years of DSM program evaluation in the United States, there was a fair amount of social science research applied to the behavioral aspects of energy efficiency. Since the mid-1980s, however, there has been a heavy emphasis on impact evaluation, technical measurement, and engineering methodologies. Although some have articulated the need to integrate behavioral research into energy evaluation, most emphasis has tended to center on the technical/engineering aspects.

Increasingly, however, the realization is growing that it is necessary to integrate important behavioral variables into impact evaluation techniques. In addition, it is being further recognized that behavioral research questions are central to a number of critical evaluation issues: e.g., design of samples for evaluation studies, net energy savings, self-selection bias, free riders and free drivers, persistence of energy savings, process evaluation, and market impact evaluation. Finally, it is increasingly being realized that the utilization of evaluation results relies heavily on behavioral factors.

Social science researchers should be poised to expect a greatly expanded role of behavioral research in evaluation. As new techniques are developed and perfected, as the results of impact evaluations become more abundant, and as the gap between technical energy savings potential and realized savings becomes more visible, research regarding the "human dimension" of program evaluation will be crucial.

This paper provides an overview of the human dimension of program evaluation and focuses on key evaluation issues in demand-side management which will require the use of social science research for addressing these issues. For people actively involved in conducting evaluations, the above statements should not be too surprising, and only confirm what they experience daily. However, for individuals new to this field, this paper should serve as an introduction to the critical behavioral issues affecting evaluation and, perhaps more importantly, as a stimulus to motivate more people to become involved in the human dimension of program evaluation.

INTRODUCTION

Historical Context

In the 1970s and 1980s, some research was carried out in the United States on analyzing the behavioral dimension of energy use, particularly in residential buildings and, to a lesser extent, in commercial and industrial buildings.[1] Most of the research focused on explaining the variation in energy use and savings among households, and improving predictive models of energy use and savings in buildings. The studies demonstrated the importance of behavior in affecting energy use and savings and clearly showed that technical factors alone were neither sufficient nor adequate for explaining energy use in buildings.

During this time, federal and state governments in the United States, along with a few utilities, began promoting energy efficiency through energy conservation programs: primarily, information and audit programs (e.g., Project Conserve, the Residential Conservation Service, the Energy Extension Service, and the energy-efficiency labeling of cars, homes, and appliances) (Clinton et al. 1986). Some evaluations of these programs applied social science research to the behavioral aspects of energy efficiency (e.g., Stern and Gardner 1981). The primary focus of these evaluations was on process evaluation: e.g., analyzing program participation rates, reasons for participation or non-participation, and program satisfaction. The impacts of these programs were rarely examined; one exception was calculating the number of participants and measures installed.[2] Furthermore, the evaluations of these program primarily focused on individual decision makers (especially, residential consumers), rarely households (i.e., two or more individuals living in a dwelling unit), and often excluded decision makers within large organizations (e.g., commercial customers, utility companies, and manufacturers).

Since the mid-1980s, several changes have occurred. First, utility companies have been very active in implementing demand-side management (DSM) programs in response to regulatory direction from state public utility commissions to acquire DSM as a resource (NARUC 1988). Second, the range and diversity of programs offered has expanded: complementing consumer information and marketing programs are a greater variety of financial incentive programs (e.g., rebates, conservation rates, and reduced loan interest rates and loan-qualifying criteria), technical information programs (e.g., design tools, design assistance, and standards-related training, compliance, and quality control), pilot demonstration programs, and programs aimed at the commercial and industrial sectors (Vine and Harris 1990; Nadel 1992). Third, several utilities in the United States now receive financial incentives for promoting energy efficiency (Reid and Brown 1992). Fourth, the emphasis on cost-effectiveness tests by regulatory and utility organizations for examining utility program performance has led to more evaluation activities. In conclusion, DSM programs have proliferated at an exponential rate, complemented to a somewhat lesser degree by an interest in program evaluation.

Recently, the primary focus of evaluation activity has been on impact evaluation, technical measurement, and engineering methodologies. This type of evaluation has focused on gross program impacts, without considering the energy savings that can be attributed to the program itself ("net impacts") (see below). Although some have articulated the need to integrate behavioral research into energy evaluation (e.g., Lutzenhiser 1990 and 1993) and some very interesting behavioral research has continued, most emphasis has tended to center on the technical/engineering aspects. While there has been some interest in process evaluation and integrating the findings of process and impact evaluation (Bronfman and Peters 1991; Spinney and O'Rourke 1992), most recent evaluations in the U.S. have largely been accounting studies,

to verify that the estimated savings have occurred, often focusing on technological and engineering features.[3] Thus, for many, the first order of business is answering the basic question of how much savings a program achieved, and the second order of business is examining the reasons accounting for the realized savings (including penetration rates). We would assert that once people become more interested in the underlying determinants of energy savings, the design of experimental programs and approaches, and more effective utilization of evaluation results, interest in the human dimension of program evaluation will increase.

Social Science Research

For the purposes of this paper, the "human dimension" of program evaluation refers to: how people think about the "energy problem" and "energy efficiency" in general and about energy services (and their relationship to comfort, health and other environmental conditions), energy efficiency equipment. and energy efficiency programs in particular; how people decide on investing in energy efficiency equipment; how people operate and maintain energy efficiency equipment; and research design methodologies and techniques for addressing social problems such as the inefficient use of energy (e.g., development of formal models and field experimentation) (Stern 1985). In addition to looking at the attitudes and behaviors of individual decision makers, the "human dimension" also examines energy use and decision-making by households (i.e., two or more individuals living in a dwelling unit) and by large organizations (e.g., commercial customers, trade allies, utility companies, and manufacturers).

Social scientists can play a valuable role in program evaluation by using their expertise in the following areas: (1) survey questionnaires for measuring attitudes and behaviors; (2) designing sampling strategies for program evaluation; (3) creating experimental designs for studying, for example, different market delivery strategies or financial incentive mechanisms; (4) developing analytical models; (5) designing different measurement techniques for examining self-reported behavior; (6) utilizing innovative interviewing techniques and case studies; and (6) providing general insights on individual and program dynamics based on their work in the fields of sociology, psychology, anthropology, human ecology, and economics.

Importance of Behavior in Evaluation

Past reports of the National Research Council's Committee on Behavioral and Social Aspects of Energy Consumption and Production found that "conservation policies and programs have been built on an inadequate understanding of how people respond to prices, information, incentives, and other stimuli. Some of the needed knowledge exists, but much still has to be developed in the process of designing and implementing energy policies and programs" (Stern 1985). We would argue that the field of program evaluation also needs to develop this knowledge base.

Attitudinal and behavioral variables play an important role in both impact and process evaluation for improving (1) our understanding of key program design and evaluation issues, (2) the research design and analytical techniques used in analyzing energy use and savings, and (3) the utilization of program evaluation results.

Program design and evaluation

In designing and evaluating DSM programs, program designers, managers, and evaluators need to be aware of the following: (1) programs are conducted by individuals within organizations that have their own internal dynamics that may affect overall savings of a program (e.g.,

the relationship between top management and middle and lower management, or the capability and inclination of field staff regarding the use of new software tools in selecting energy efficiency measures and targeting customers) (Vine et al. 1992); (2) the adoption and diffusion of energy efficiency measures is dependent on key behavioral factors (e.g., customer aversion to risk, sociodemographics of participants, and maturity of the marketplace) (Braun and Williams 1979; Claxton et al. 1981; Sharma 1983); (3) the performance of energy efficiency measures is not only dependent on the proper design, installation and efficiency of a particular measure, but also, and perhaps more importantly, on how the measure is used and maintained (Vine 1992); and (4) the manner in which program participants respond to programs is diverse, ranging from receiving information to installing energy efficiency measures.

These concerns are central in addressing specific evaluation issues, as noted below:

1. How does an energy efficiency program affect people's decisions? (Important for calculating net energy savings, self-selection bias, free riders (program participants who would have adopted program recommended actions regardless of the existence of the program) and free drivers (customers that take energy efficiency actions recommended by DSM programs because of the program, but do not participate directly in the program - the "spill over" effect), and cream skimming (customers who install the lowest cost energy efficiency measures and do not install other cost-effective measures);

2. How do occupants/consumers operate energy efficiency measures? (Important for determining baselines, persistence of savings, and the rebound (takeback) effect (changes in customer behavior resulting in greater energy use stimulated by participation in a DSM program), and for comparing measured and estimated savings);

3. What is the experience and expertise of building managers and operators in operating and maintaining buildings and energy efficiency equipment? (Important for determining baselines and persistence of savings, and for comparing measured and estimated savings);

4. What is the experience and expertise of trade allies in purchasing and stocking energy efficiency equipment? (Important for determining baselines);

5. How reliable is the information supplied by consumers? (Important for determining baselines and comparing measured and estimated savings);

6. How heterogeneous is behavior? (Important for designing evaluation samples and determining baselines).

Additional evaluation issues will also require significant contributions from the social sciences (Bronfman and Peters 1991): (1) attitudes and level of satisfaction of field staff, contractors, and customers; (2) effectiveness and efficiency of program implementation delivery; (3) effectiveness of contractors and installers; (4) attitudes and level of satisfaction of building owners and occupants; (5) barriers to program participation; (6) differences between program participants and nonparticipants; and (7) assessment of market potential.

Evaluation techniques

In designing and using the techniques and methodologies to evaluate programs, evaluators need to be cognizant of the following: (1) the evaluation issues noted above; (2) the modeling of energy use in buildings relies on the use of key variables representing important behavioral elements (e.g., thermostat settings, hours of building occupancy, and equipment usage) (Vine et al. 1982; Gladhart and Weihl 1990); and (3) program evaluation is not simply a technical exercise that is objective and value-free, but can be quite subjective and value-laden (Spinney and O'Rourke 1992). For example, as discussed below, many assumptions need to be made regarding key evaluation estimates (e.g., number of free riders, free drivers, rebound effect, measure lifetime, and hours of operation) that may be derived from studies using sound and/or sophisticated research methodologies; however, for some utilities, these assumptions often rely on "expert opinion" at best and subjective impressions at worst.

Integrating important behavioral variables into impact evaluation methodologies is very important. For example, in a recent comparison of engineering estimates of gross savings to impact evaluations of net savings, Nadel and Keating (1991) found large percentage differences between the two estimates in many kinds of DSM programs. Where engineering estimates and impact evaluations differed, a number of common explanations emerged, including the following: (1) wrong assumptions in the engineering estimates (e.g., over-estimating lighting system operating hours); (2) "take-back" effects for some measures (air-conditioners in moderate climates, low-flow showerheads, and residential compact fluorescents) were not adequately modeled in the engineering estimates; (3) quality control problems during measure installation, commissioning, and maintenance; and (4) greater than expected adoption of conservation measures by non-participating customers.

Evaluation utilization

The results of program evaluations will not be utilized without the effective communication of the major findings to key stakeholders within the utility as well as to outside parties (e.g., regulatory commissions, third-party intervenors, etc.). A clear identification of the objectives of each of the interested parties and their information needs is vital prior to the initiation of an evaluation. Once completed, the outcome of the evaluation must be effectively explained to these key parties.

KEY EVALUATION ISSUES

Table 1 lists some key evaluation issues that social science research needs to address. The issues are grouped into six categories: evaluation objectives, evaluation design, impact evaluation, process evaluation, market evaluation, and evaluation utilization. These categories are not mutually exclusive; in fact, the amount of overlap among the categories is very large.

Table 1. Key Program Evaluation Issues

Evaluation Design
Resources available for evaluation
Sample design
Selection of evaluation methodology
Impact Evaluation
Net and gross energy savings
Comparison (control) groups
Self-selection bias
Free riders
Free drivers
Persistence of energy savings
Process Evaluation
Barriers to program participation
Customer satisfaction with program
Market Evaluation
Assessing market transformation
Evaluation Utilization
Utilization of evaluation results

Evaluation Objectives

The first issue facing evaluators is both philosophical and pragmatic: what are the objectives of the evaluation? And whose objectives should the evaluation respond to? These are important questions that need to be addressed at the outset so that the evaluation is useful for the key stakeholders interested in the outcomes of the evaluation (see below). Focus groups and personal interviews are often used at the outset to help guide the evaluation in meeting its objectives.

Evaluation Design

The next issue confronting evaluators is methodological: how should the evaluation be designed? Illustrative examples include the following: What, and how many, people should be interviewed, if any? What kind of, and how many, measures (and buildings) should be monitored, if any? How many years should measurement last, if any? How often should data be collected on people, measures, and/or buildings, if any? What evaluation methodologies should be used, if any? Who should be involved in designing the evaluation study, if any? The proper design of a study of customer behavior, and particularly the design of the sampling and analysis plan to exclude bias (e.g., self-selection bias, survey response bias, recruitment bias, data quality

bias, and outlier bias), is considered by some to be the single most important element of program evaluation (Horowitz 1992). Thus, a sampling strategy, if required, needs to be implemented that considers, and corrects for, as many sources of potential bias as can be identified. Often, steps can be taken to ensure that a study design is unbiased at little or no cost.

Some guidance on evaluation design can be provided by management tools and professional publications. However, the answers to these questions will depend to a large degree on the attitudes of key people involved in the evaluation, affected by the evaluation, or simply interested in the evaluation: e.g., the sponsors of the evaluation (who are typically resource constrained), regulators, and the people responsible for conducting the evaluation. Thus, one needs to assess: (1) the availability of expertise for conducting the evaluation; (2) the level of precision, risk, and uncertainty the key stakeholders are willing to accept; (3) the amount of financial resources available for conducting the evaluation; and (4) the incremental value of information gained in obtaining additional evaluation information. There is no absolute standard for when to be confident in study findings; acceptable precision levels depend on the conventions of the field of study. The costs of error, the costs of measurement, what is technically achievable given the available measurement tools, and finally, the tolerable level of uncertainty, all play a role in establishing reasonable and prudent standards (Horowitz 1992). The field of DSM impact evaluation has yet to agree on standards for confidence or required precision levels, although people have their own opinions. Numerous tradeoffs need to be made. As regulators and third-parties to the evaluation become more involved at the design stage of the evaluation (as evident in recent attempts at "collaborative evaluations"), we believe the role of the evaluator will be that of a facilitator with a social science background.

Impact Evaluation

Net and gross energy savings

Gross savings are the changes in annual energy use and peak demand experienced by participants in DSM programs. In contrast, net savings are typically the total savings that can be directly attributed to a DSM program. In other words, net savings are the difference between gross savings and the change in consumption and demand that participants would have achieved had the program not existed. Nonprogram savings (or energy increases) reflect customer responses to changes in electricity and fossil-fuel prices, changes in economic activity or personal income, introduction of new electricity-using technologies, and other nonprogram factors. Table 2 lists some of the key adjustments that need to be made to determine net energy savings, including such behavioral factors as the thermostat schedule and settings and the hours of operation (of a building, or a particular measure).

Table 2. Energy Savings Adjustments

Weather
Daylight/daylength
Building occupancy
Building function
Installation of additional equipment
Repair, replacement, removal, or retrofit of existing equipment
Thermostat schedule and settings
Hours of operation
Power outages and other supply disruption
Free riders
Free drivers
Industrial production
Agricultural production

Baseline

In order to measure changes in energy use (for both participants and nonparticipants), a baseline must be identified from which a change can be measured or estimated. Unless baseline construction practices can be accurately determined, a correct assessment of energy and demand savings is not possible. Determining this baseline is a critical step, but fraught with difficulties. Unlike retrofit programs where there is a clear before and after, the determination of the before (no DSM or non-participant case) is less straightforward in new construction. Furthermore, determining the correct baseline is even more complicated when incentives are provided to vendors or manufacturers: the baseline will shift as dealers stock equipment to meet rebate opportunities. The lack of data is the most important problem in determining baselines, especially for new construction. As discussed below, three additional problems are associated with the selection of an appropriate baseline: self-selection bias, free riders, and free drivers.

Approaches that can be used in establishing new construction baselines include the following (Reed 1991): (1) examination of state and local building codes; (2) review of national and state appliance standards; (3) interviews and focus groups with local design professionals, code officials, vendors, and manufacturers; (4) on-site inspections of new and existing construction; and (5) mail surveys of builders and contractors' current building practices.

Comparison groups, self-selection bias, free riders, and free drivers

Self-selection bias occurs when program participation is voluntary: systematic differences may occur between program nonparticipants, who either chose not to participate or were unaware of the program, and participants. A key analytical issue confronting the evaluator is the determination of an appropriate control group. In educational, psychological, and medical studies, it is often possible to select individuals randomly into control and treatment groups. However, randomization of observations is usually not possible in the evaluation of energy-efficiency programs. Accordingly, without strict control over selecting the different study groups, many researchers balk at the term "control" group, preferring "comparison" or "non-participant" group. Furthermore, in some regions of the country, energy efficiency programs have been implemented

over ten years or more, so that everyone who has access to a newspaper, radio, or television has probably read or heard about the program, although they may not have participated. Finally, certain types of customers (e.g., large industrial firms) may have unique facilities that have no comparative buildings. Accordingly, finding a proper control group may be very difficult, if not impossible; in this case, modeling of energy use before and after retrofit may be the only method to estimate energy savings. Thus, a challenge for social science researchers is defining and selecting an appropriate comparison group.

The comparison group may be similar to the group of program participants or may be more representative of the population (assuming that program participants are "not typical" - e.g., program participants may have higher energy use, higher incomes, higher levels of education, or higher process loads): the researcher may use a comparison group that is representative of the population in general when thinking about future program participants. Many comparisons of control and participant groups focus on income, education, dwelling size, geographic location, and pre-program energy use. However, in only a few cases are key attitudinal and behavioral characteristics compared, needed for examining such issues as self-selection bias. Accordingly, an important social science research need is to identify those attitudinal and behavioral factors that might affect the results of the evaluation and to design the evaluation so that these parameters can be incorporated into the evaluation design and analysis.

Free riders are defined as those participants in a conservation program who would have installed the same energy conservation measures even if there had been no program. Free ridership can be viewed as a subset of self-selection bias. Customer surveys of different types of utility programs report free rider fractions ranging from 3 to 89% (Saxonis 1991a). Utilities need to minimize the effect of free riders if they do not wish to pay for DSM actions that would have been installed without the program. Furthermore, utilities need to measure this effect if they wish to examine net energy savings (although free riders is only one issue in the net/gross energy problem). Because free riders represent a cost to the program but offer no direct benefits to the utility in return, DSM programs that are highly successful from a gross savings perspective may prove to be less attractive when free riders are considered. Accordingly, the accurate measurement of free ridership is a significant issue for evaluators and program managers of DSM programs.

Violette (1991) discusses the pros and cons of various data analysis methods for estimating the impact of DSM programs on kW or kWh consumption with regard to self-selection, and Saxonis (1991a) examines the various options for measuring free riders (surveys, marketplace data, comparison groups, and statistical models). The measurement of free riders can be enhanced with social science techniques in the following manner. Surveys can be improved by targeting questions to address (1) free ridership, (2) net program impacts, and (3) types of programs (e.g., by type of appliance or market delivery mechanism). In addition to developing a positive relationship with manufacturers and distributors, market data can be obtained by requiring dealers to provide sales data as a prerequisite for participating in a program. Monitoring activity of a non-participating region with similar characteristics to the test market is a practical and efficient option for analyzing free riders (see below). Also, a number of modeling techniques are available for evaluating self-estimation bias and free riders (see Violette 1991), but few studies have used these new techniques. Finally, program designs can be improved to reduce the impact of free ridership: (1) require preinstallation inspections to ensure that potential participants do not have the efficient equipment promoted in the program; (2) encourage the use of very efficient (advanced technology) equipment; and (3) target customer groups with, historically, low free

rider rates (e.g., low-income households).

A free driver contributes to the goals of the program (e.g., reduced energy consumption) but is not formally a program participant (Saxonis 1991a). A free driver is affected by the program either through a conscious awareness of the program or because of program-induced changes in the marketplace (e.g., a customer who purchases a product that qualifies for a rebate but does not claim the rebate, or a builder who constructs a home to program standards for competitive reasons (since all others are building to this level) but does not choose to participate in the program). Free drivers are more likely to be a significant problem for program evaluators when programs have been in existence for several years and have achieved high participation levels. Also, free drivers may become a more visible research area if utilities receive financial incentives for the savings achieved by free drivers.

Research on free drivers is limited. A study of one of the oldest residential conservation programs in the United States (i.e., the Davis Energy Conservation Building Code) revealed a 15% reduction in electricity consumption that was attributed to both the structural features of the code and to the energy conserving behavior of the residents of the city (free drivers) (Dietz and Vine 1982).

Three approaches are available to address the issue of free drivers (Saxonis 1991a). First, use a historical baseline from the early years of the program. Second, use survey methods to determine whether nonparticipants have changed their energy use as a result of the program and to determine (via trade ally interviewing) if the market for the DSM actions promoted by the program has moved in such a way that nonparticipants are impacted. And third, use a community (or set of communities) outside the area in which the program is offered, and compare the distribution of efficiencies for program-sponsored measures in the participating area with those in the comparison area (see below).

Persistence of energy savings

The persistence of energy and demand savings is an important issue to all stakeholders: e.g., building owners, architects and engineers, utility program managers and evaluators, regulators, utility shareholders, resource planners, forecasters, researchers, and consumers. Until recently, persistence was assumed to be relatively constant, and most analyses of persistence relied on engineering estimates of measure life and judgement (with adequate justification). For example, most planners assumed that knowing the physical life of a measure installed was sufficient to determine persistence: i.e., first-year savings continued for the life of the measure (e.g., 20 years). Recently, this assumption has been challenged as the issue of persistence has gained more prominence in the evaluation of energy efficiency programs. In fact, the limited empirical research conducted so far raises questions about the validity of using manufacturer's claims for physical measure lives as a basis for projecting persistence.

Two dimensions of persistence exist: measure persistence and program persistence (Keating 1991). The former focuses on measure lifetime and operation, while the latter emphasizes gross and net program impacts (Vine 1992). The two dimensions are interrelated in that program persistence includes measure persistence variables as well as other factors. Key parameters affecting the different dimensions of persistence are noted in Table 3. Both technology and behavior affect persistence, often interact, and are difficult to separate, particularly for "behaviorally dependent" measures, such as cleaning of refrigerator condenser coils and the resetting of HVAC time clocks to accommodate occasional, unplanned occupancies (Keating 1991).

Table 3. Factors Influencing the Persistence of Energy Savings
(Program persistence factors include measure persistence factors plus those listed in the table.)

Measure Persistence	Program Persistence
Technical lifetime	Rebound (snap-back, take-back) effects
Measure installation	Surge effect (additional measures added by customer after initial program participation)
Measure performance or efficiency decay	Replacement effect (replacing efficiency measures with less or more efficient measures)
Measure operation (behavior)	Changes in energy use by control group
Measure maintenance, repair, commissioning	
Measure failure	
Measure removal	
Changes in the building stock (renovations, remodels, alterations, additions)	
Occupancy changes (turnover in occupants; changes in occupancy hours and number of occupants)	

One key finding of persistence research is that the reliance on technical or average service lifetime may overestimate DSM measure savings, particularly in the commercial and industrial sectors where renovations and remodels occur frequently and where removal or deactivation of energy conservation measures occurs often. In addition, certain building types appear to be more susceptible to frequent remodeling and turnover: e.g., office, retail, restaurant, and warehouse sectors. Another finding is that typically, certain measures in the residential sector are prone to removal by the occupants: e.g., low-flow showerheads, compact fluorescent bulbs, and door weatherstrips.

Process Evaluation

Process evaluations examine program operation during implementation and can provide: (1) prompt feedback to help improve program design and implementation; (2) important insights that can be used to interpret findings from impact evaluations and for changes in the design of future programs; and (3) a stimulus for collecting appropriate monitoring and tracking data. Process evaluations typically focus on program goals, history, and activities, and are often based on interviews with utility program staff, program managers, participants, and trade allies. The rewarding of financial incentives for utility investment in energy efficiency measures has elevated impact evaluation to a higher level of importance than before (which had already been

considered by many to be more important than process evaluation). Unfortunately, process evaluation has been relegated to a second order of importance when, in fact, process evaluation can serve many important objectives that can complement (not just supplement) impact evaluation (Schueler and Quaid 1992).

The goal of process evaluation is program optimization through: improvement in implementation efficiency, assessment of market segments and targeting of specific segments, improvement in quality of measure installation, identification of program-design issues, interim accounting of program progress through reviews of the program database, and examination of special issues (e.g., measure life and program comprehensiveness) (Bronfman and Peters 1991).

As DSM-program performance becomes more closely tied to integrated resource planning and DSM incentive regulation, process evaluation will assume more importance (Bronfman and Peters 1991). First, because many programs are designed and implemented quickly, early process evaluations will be critical to fine-tune programs, assess the market potential, and identify market segments not reached. Second, because process evaluations are in the field early and because impact evaluations often do not produce results for a year or two, process evaluations will be instrumental in producing early reports on program effectiveness: e.g., assistance in designing the program database, assessing the reliability of program data, and periodic reporting of implementation data (e.g., the number of installations, estimated and actual costs, and projected savings). And third, the results of process evaluations will need to be reported in the context of other evaluation findings within the same utility, so that the results of different evaluation components will provide a complete, integrated evaluation, and so that impact data are interpretable in light of the implementation of the program.

The analysis of barriers to program participation and the level of satisfaction (by customers and trade allies) with the program are two key process evaluation issues requiring the attention of social scientists. Barriers to program participation often include the following (BEEPRG 1992): (1) information gaps; (2) split incentives (plant engineer versus financial officer, owner versus tenant, etc); (3) the large number of decision makers involved; (4) difficulty of computing actual energy savings (see above); (5) lack of trust in vendors; (6) lack of trained service and maintenance personnel; and (7) misconceptions about energy efficiency. Program satisfaction is critical for determining the willingness of consumers to participate in future programs and for assisting the marketing of a program to others. Reasons for customer dissatisfaction include the following: (1) actual energy savings were lower than estimated energy savings; (2) actual payback levels were longer than estimated; (3) thermal comfort, aesthetics, and other indoor environmental conditions were less than desired; (4) rebate application process was burdensome and slow; and (5) interactions with utility program staff were unfriendly and disruptive to work and home life.

Market Impact Evaluation

Very little research has been devoted to examining the impact of DSM programs on the marketplace ("market transformation"). Compared to DSM resource acquisition, the impacts resulting from market transformation are likely to be larger over the long run, but also more gradual, less documentable, and harder to predict and control (Schlegel et al. 1993). However, as more research is devoted to the analysis of net energy savings and free riders and free drivers, the impact of programs on the market place will receive increasing attention.

Schlegel et al. (1993) have developed a useful conceptual framework for understanding market transformation and for systematically classifying the ways in which DSM programs may transform energy efficiency markets. The framework is based on three perspectives: economic, social psychological, and sociological. For this paper, the latter two perspectives are of import. The social psychological perspective views market transformation as a basic change in attitudes or awareness of consumers sufficient to change their behavior with regard to energy conservation measures. The sociological perspective sees market transformation as a complex mixture of changes in social norms, institutional relationships, and social sanctions which act together to encourage energy efficiency.

As part of the evaluation of market transformation, the distribution infrastructure and market intermediaries for efficient technologies need to be studied. The decisions made by manufacturers, suppliers, and other distributors of efficiency can clearly affect the choices of end users, for example, by restricting their range of options, or by promoting very efficient equipment. For example, attitudinal and behavioral studies on the distribution infrastructure of energy-efficient residential refrigerators and water heaters have been recently completed in the Pacific Northwest (Pope et al. 1992; Gordon et al. 1991a, 1991b).

Utilization of Evaluation Findings

Effectively communicating the results of evaluations is crucial for ensuring that evaluation findings are utilized (rather than "sitting on a shelf") (McRae et al. 1992). Because of the diverse group of potential users of evaluations, it is important to identify those stakeholders who will use, will be affected by, or will have an interest in the program or the evaluation results. These people include utility staff (e.g., from load forecasting, resource planning, demand-side program planning, rate design, supply system operation, corporate communications, finance, and policy making/corporate strategic planning), customers, trade allies, and regulators. Each stakeholder has different information needs that evaluations can potentially meet, such as consumption data (e.g., energy use, peak demand, load profile, and coincidence), end use data (e.g., appliance and equipment stocks, locations, vintage, and time of use), operational data (e.g., how the customer uses the equipment and its controls), and customer data (e.g., attitudes, level of satisfaction, perceptions, and sources of information). The likelihood that evaluation results are used can be increased if this need is addressed in the entire evaluation process: during the scoping period of the evaluation, while conducting the analysis, in describing the findings, and in disseminating the findings. This is particularly important as attribution of changes and reconciliation of results is unlikely to occur without the stakeholders' agreement on methodologies and techniques used in the early phases of the project.

Spinney et al. (1991) note that the organizational structure, management orientation, and traditional culture of an electric utility (a large-scale organization) can create formidable obstacles for the effective utilization of evaluation results. For example, electric utilities tend to be large, hierarchical organizations with many departments and highly specialized functions, leading to different perspectives on the problems surrounding program evaluation and to time-consuming and resource-intensive resolution processes. Also, the cultural context of these utilities often reflects an "engineering orientation," preferring hands-on technical know how and relatively precise and predictable technical systems. This type of culture can often leave little room for an appreciation of human behavior and the key program evaluation issues noted previously.

Because the effective utilization of evaluation results requires good planning, good data, and a positive working relationship with program managers, executive management, regulators, and other concerned parties (Saxonis 1991b), research in energy use decision making in organizations is critical for the program evaluation field. Effective communication is essential to eliminate any potential misunderstandings prior to the beginning of the evaluation, because it is difficult to change course once the evaluation process has begun and impossible after the report is released. Because the lines of communication in organizations are often weak or sometimes non-existent, the evaluator must be a negotiator and a diplomat, and evaluations must be flexible.

As an example, B.C. Hydro has developed a multi-department evaluation team, representing all the stakeholders of evaluation information (El-Ramly 1993). The responsibility of the team is to jointly agree on an evaluation approach, validate assumptions, and identify the type and format of information required. B.C Hydro asserts that integrated information gathering and analysis is necessary for maximizing benefits and minimizing costs and this type of integrated planning represents a "corporate culture," not simply a planning activity.

Finally, we need to evaluate the evaluators to see what difference specific program evaluations have made on energy efficiency policies and on other energy efficiency programs. We may find that there are real limits to evaluation utilization, reflecting the lack of unambiguous and clear results from program evaluations.

CONCLUSIONS AND RECOMMENDATIONS

Although behavioral research still remains relatively underemphasized in the DSM evaluation arena, the situation is changing. Social science researchers should be poised to expect a greatly expanded role of behavioral research in evaluation for the following reasons. As increasingly sophisticated modeling techniques are developed and perfected on non-technical aspects, accurate and reliable behavioral inputs will be of increasing value. Even more importantly, as the flood of impact evaluations which have been initiated in the last couple of years begin to produce results--results that will undoubtedly be lower than predictions/expectations--there will be a clamor to provide information to "explain" those results. Research regarding the "human dimension" will be a crucial aspect of the necessary response. Finally, if market transformation is accepted as an appropriate policy goal, then a large amount of research will need to be conducted to see if this occurs and this will entail a fair amount of behavioral research.

We believe the role of the evaluator will be that of a facilitator and negotiator. Communication will be a key activity throughout all phases of the evaluation process. And social science skills will be at a premium for management type of activities as well as for analysis. In fact, as evaluation becomes more of a team effort and more multidisciplinary, and as the evaluation field grows, the value of social scientists is expected to increase. Similarly, the need for educating and training evaluation professionals is likely to grow, with courses and material in the social sciences playing a prominent role.

In order to highlight the importance of the human dimension of program evaluation, certain barriers to social science research need to be overcome: e.g., fear of experimentation, fear of making mistakes, a predominant engineering orientation in evaluation, diminished academic interest in energy and energy program evaluation, fewer social scientists familiar with these areas, and fewer graduate students trained in these areas (Lutzenhiser 1993). We need to provide better financial support for students and faculty in academia and better training of utility

and government staff involved in managing, designing, implementing, and evaluating DSM programs. The development of "centers of excellence" is needed to: (1) encourage more behavioral research, teaching, and graduate training; (2) improve contact between academic researchers, utilities, government, and manufacturers; and (3) facilitate cooperative research ventures that combine the resources and talents of academic, government, and private sector researchers (Lutzenhiser 1993).

ACKNOWLEDGMENTS

I would like to thank Marty Kushler, Ken Keating, and Jeff Schlegel for their comments on some of this material previously presented at the American Council for an Energy-Efficient Economy's 1990 Summer Study on Energy Efficiency in Buildings (Kushler et al. 1990). I would like to thank the following people who reviewed an earlier version of this paper: Peter Cebon, Tom Dietz, Joe Eto, Rick Fleming, Fred Gordon, Ken Keating, Loren Lutzenhiser, Alan Meier, Peter Miller, Gil Peach, Jane Peters, Mary Ann Piette, Ralph Prahl, Bill Saxonis, Linda Schuck, Vince Schueler, and Michael Togeby.

END NOTES

1. Key references on behavioral work include the following: the proceedings from the Behavior (Human Dimensions) panel of the American Council for an Energy-Efficient Economy's Summer Study on Energy Efficiency in Buildings (published every even year, from 1980-1992); Kempton and Neiman 1987; Lutzenhiser 1993; Stern 1985; and Gladhart et al. 1986. A recent collection of articles on the behavioral aspect of residential cooling is found in *Energy and Buildings*, Vol. 18, 1992.

2. During this time period, many states, the U.S. Department of Energy (DOE), and the Bonneville Power Adminstration did examine impacts of their programs, as part of the first wave of impact evaluation. For example, DOE required (and still does) an energy-savings report reflecting energy savings for all State Energy Conservation Program-funded programs (Collins et al. 1985).

3. Some utilities are noteworthy in being the exception to this trend and have spent significant resources in examining customer attitudes, satisfaction, and market segmentation.

REFERENCES

Braun, B. and S. Williams. 1979. "Changing Household Energy Uses." *Changing Energy Use Futures*, Volume 2, pp. 857-864, Pergoman Press.

Bronfman, B. and J. Peters. 1991. "Process Evaluation of DSM Programs." *Handbook of Evaluation of Utility DSM Programs*. E. Hirst and J. Reed, ed., pp. 135-148. Oak Ridge National Laboratory, Oak Ridge, TN.

Building Energy Efficiency Program Review Group (BEEPRG). 1992. *Achieving Greater*

Energy Efficiency in Buildings: The Role of DOE's Office of Building Technologies. American Council for an Energy Efficient Economy, Washington, D.C.

Claxton, J., C. Anderson, J. Ritchie, and G. McDougall. 1981. "Diffusion of Energy Conservation and Technologies." *Consumers and Energy Conservation*. D. Leonard-Barton, ed., pp. 97-107.

Clinton, J., H. Geller, and E. Hirst. 1986. "Review of Government and Utility Energy Conservation Programs." *Annual Review of Energy*, Vol. 11:95-142.

Collins, N., L. Berry, R. Braid, D. Jones, C. Kerley, M. Schweitzer, and J. Sorensen. 1985. *Past Efforts and Future Directions for Evaluating State Energy Conservation Programs*. ORNL-6113. Oak Ridge National Laboratory, Oak Ridge, TN.

Dietz, T. and E. Vine. 1982. "Energy Impacts of a Municipal Conservation Policy." *Energy* 7(9):755-758.

El-Ramly, Z. 1993. "Corporate Need for Monitoring and Evaluation." *Proceedings from the 1993 Integrated Resource Planning Conference*, pp. 307-317. Electric Utility Consultants, Inc., Englewood, CO.

Gladhart, P. and J. Weihl. 1990. "The Effects of Low-Income Weatherization on Internal Temperature, Occupant Comfort, and Household Management Behavior." *Proceedings from the ACEEE 1990 Summer Study on Energy Efficiency in Buildings*, Volume 2, pp. 2.43-2.52. American Council for an Energy Efficient Economy, Washington, D.C.

Gladhart, P., B. Morrison, and J. Zuiches. 1986. Eds. *Energy and Families: An Analysis of Lifestyles and Energy Consumption*. College of Human Ecology, Michigan State University, East Lansing, MI.

Gordon, L., L. Heifetz, B. Lagerberg, S. Rudeen, and V. Schueler. 1991a. *Energy Efficient Water Heaters and the Northwest Market*. Washington State Energy Office, Olympia, WA.

Gordon, L., B. Lagerberg, C. Murray, and S. Rudeen. 1991b. *Water Heater Market Analysis*. Washington State Energy Office, Olympia, WA.

Horowitz, M. 1992. "Savemetrics: The Science of Measuring DSM Impacts." *Proceedings from the 1992 International Energy Efficiency and DSM Conference*, pp. 189-199, Synergic Resources Corporation, Bala Cynwyd, PA.

Keating, K. 1991. "Persistence of Energy Savings." *Handbook of Evaluation of Utility DSM Programs*. E. Hirst and J. Reed, eds., pp. 89-99. Oak Ridge National Laboratory, Oak Ridge, TN.

Kempton, W. and M. Neiman. 1987. Eds. *Energy Efficiency: Perspectives on Individual Behavior*. American Council for an Energy Efficient Economy, Washington, D.C.

Kushler, M., K. Keating, J. Schlegel, and E. Vine. 1990. "The Purpose, Practice, and Profession of DSM Evaluation: Current Trends, Future Challenges." *Proceedings from the ACEEE 1990 Summer Study on Energy Efficiency in Buildings*, Volume 7, pp. 7.1-7.25. American Council for an Energy Efficient Economy, Washington, D.C.

Lutzenhiser, L. 1990. "Explaining Consumption: The Promises and Limitations of Energy and Behavior Research." *Proceedings from the ACEEE 1990 Summer Study on Energy Efficiency in Buildings*, Volume 2, pp. 2.101-2.110. American Council for an Energy Efficient Economy, Washington, D.C.

Lutzenhiser, L. 1993. "Social and Behavioral Aspects of Energy Use." *1993 Annual Review of Energy and the Environment*, forthcoming.

McRae, M., T. Henneberger, and P. Hanser. 1992. *Proceedings from the ACEEE 1990 Summer Study on Energy Efficiency in Buildings*, Volume 7, pp. 7.133-7.142. American Council for an Energy Efficient Economy, Washington, D.C.

Nadel, S. 1992. "Utility Demand-Side Management Experience and Potential: A Critical Review." *1992 Annual Review of Energy and the Environment*, Vol. 17:507-535.

Nadel, S. and K. Keating. 1991. "Engineering estimates vs. impact evaluation results: How do they compare and why?" *Proceedings from the National Energy Program Evaluation Conference*, p. 24-33. National Energy Program Evaluation Conference, Rochester, IL.

National Association of Regulatory Utility Commissioners (NARUC). 1988. *Least-Cost Utility Planning Handbook for Public Utility Commissioners, Vol. 1.* National Association of Regulatory Utility Commissioners, Washington, D.C.

Pope, J., M. Slavin, and S. Rudeen. 1992. *Energy Efficient Home Refrigerators: A Market Analysis.* Washington State Energy Office, Olympia, WA.

Reed, G. 1991. "No Before, Only After: The Importance of Establishing a Baseline for New Construction." *Proceedings from the National Energy Program Evaluation Conference*, p. 574-579. National Energy Program Evaluation Conference, Rochester, IL.

Reid, M. and J. Brown. 1992. *Incentives for Demand-Side Management.* National Association of Regulatory Utility Commissioners, Washington, D.C.

Saxonis, W. 1991a. "Free Riders and Other Factors that Affect Net Program Impacts." *Handbook of Evaluation of Utility DSM Programs.* E. Hirst and J. Reed, eds., pp. 119-134. Oak Ridge National Laboratory, Oak Ridge, TN.

Saxonis, W. 1991b. "Evaluation Protocols: Real Problems, Real Answers." *Proceedings from the National Energy Program Evaluation Conference*, p. 2-7. National Energy Program Evaluation Conference, Rochester, IL.

Schlegel, J., G. Edgar, R. Prahl, M. Kushler, and D. Narum. 1993. *Evaluation of DSM*

Shareholder Incentive Mechanisms. Final Report. Wisconsin Energy Conservation Corporation, Madison, WI.

Schueler, V. and M. Quaid. 1992. "Rethinking the Role of Process Evaluation in Demand-Side Management (DSM) Resource Acquisition: An Assessment of Current Practice and Strategies for the Future." *Proceedings from the ACEEE 1992 Summer Study on Energy Efficiency in Buildings*, Volume 7, pp. 7.207-7.214. American Council for an Energy Efficient Economy, Washington, D.C.

Sharma, A. 1983. "Energy Conservation in U.S. Buildings: Solving the High Potential/Low Adoption Paradox from a Behavioral Perspective." *Energy Policy*, 11(2):148-167.

Spinney, P., L. Dethman, D. Jacobson, and D. Bowles. 1991. "Organizational Barriers to the Implementation of Commercial and Industrial DSM Programs." *Proceedings from the National Energy Program Evaluation Conference*, p. 470-475. National Energy Program Evaluation Conference, Rochester, IL.

Spinney, P. and P. O'Rourke. 1992. *DSM Process Evaluation: A Guidebook to Current Practice*. Report TR-100647. Electric Power Research Institute, Palo Alto, CA.

Stern, P. ed. 1985. *Energy Efficiency in Buildings: Behavioral Issues*. Committee on Behavioral and Social Aspects of Energy Consumption and Production, National Research Council. National Academy Press, Washington, D.C.

Stern, P. and G. Gardner. 1981. "Psychological Research and Energy Policy." *American Psychologist*, 36(4):329-342.

Vine, E. 1992. "The Persistence of Energy Savings: What Do We Know and How Can It Be Ensured?" *Energy*, 17(11):1073-1084.

Vine, E. and J. Harris. 1990. "Implementing Energy Conservation Programs for New Residential and Commercial Buildings." *Energy Systems and Policy*, 13:115-139.

Vine, E., O. De Buen, C. Goldman, and R. Prahl. 1992. "Mandating Utility Competition: One Option for Promoting Energy Efficiency." *Utilities Policy*, 2(1):51-61.

Vine, E., P. Craig, J. Cramer, T. Dietz, B. Hackett, D. Kowalczyk, and M. Levine. 1982. "The Applicability of Energy Models to Occupied Houses: Summer Electric Use in Davis." *Energy*, 7(11):909-925.

Violette, D. 1991. "Analyzing Data." *Handbook of Evaluation of Utility DSM Programs*. E. Hirst and J. Reed, eds., pp. 51-72. Oak Ridge National Laboratory, Oak Ridge, TN.

LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
TECHNICAL INFORMATION DEPARTMENT
BERKELEY, CALIFORNIA 94720