

## ANSWERING BEHAVIORAL QUESTIONS ABOUT ENERGY EFFICIENCY IN BUILDINGS†

PAUL C. STERN

Commission on Behavioral and Social Sciences and Education, National Research Council, 2101  
Constitution Avenue, N.W., Washington, DC 20418, U.S.A.

ELLIOT ARONSON

Stevenson College, University of California, Santa Cruz, CA 95064, U.S.A.

JOHN M. DARLEY<sup>1</sup> and WILLETT KEMPTON<sup>2</sup>

<sup>1</sup>Department of Psychology and <sup>2</sup>Center for Energy and Environmental Studies,  
Princeton University, Princeton, NJ 08540, U.S.A.

DANIEL H. HILL

Institute for Social Research, University of Michigan, Ann Arbor, MI 48106, U.S.A.

ERIC HIRST and THOMAS J. WILBANKS

Oak Ridge National Laboratory, Oak Ridge, TN 37830, U.S.A.

(Received 2 May 1986)

**Abstract**—We identify behavioral questions that arise with 4 kinds of policy interventions for energy efficiency in buildings: information, incentives, standards, and technological research and development. A general strategy is described for answering such questions by using 6 analytical methods: formal models, analysis of existing data, surveys, ethnographic methods, small-scale experimentation, and evaluation research. We evaluate each method for addressing each behavioral question in policy analyses.

### 1. INTRODUCTION

Policies and programs intended to improve the energy efficiency of buildings in the United States have often had disappointing results. When the Residential Conservation Service program offered homeowners individualized energy information at low cost or even for free, the response was decidedly underwhelming.<sup>2-4</sup> When conservation programs involve offers of loan subsidies for home weatherization, few people take out loans, and the rate of participation has had much less to do with the size of the subsidy than with the way the programs are marketed and managed.<sup>5,6</sup> Even when weatherization programs seem successful and they lead to the insulation of walls, the caulking of doors and windows, and the installation of energy-efficient furnaces, energy savings have not matched predictions. On the average, savings are somewhat less than predicted, but the variability is a bigger surprise: while savings for some buildings are double what was predicted, other buildings yield substantial increases in energy use.<sup>7-10</sup>

The most plausible explanations for such policy disappointments have been behavioral. People may not notice, understand, or trust the information they are given. They may find it onerous to go through all the paces demanded of participants in financial incentive programs. They may fail to adopt new energy-saving technologies that promise high returns on investment, in part because energy bills contain confusing information that

†This article is based on our work as members and staff of the Committee on Behavioral and Social Aspects of Energy Consumption and Production of the National Research Council, and appears in similar form in Chapters 1 and 2 of the Committee's 1985 report.<sup>1</sup> The committee's work was supported by a grant from the U.S. Department of Energy to the National Academy of Sciences.

cannot readily be used to confirm energy savings. And when they do invest in energy-efficient equipment, people may take some of the benefits as increased comfort by altering thermostat settings.

Such explanations often occur to policy analysts only after their programs have disappointed expectations. In future policy planning, it will be important to anticipate problems that may arise because of the human dimension of energy use, to evaluate the seriousness of the problems, and to design programs and policies to take them into account. But few policy makers or policy analysts at any level of government, on utility company staffs, or in the other organizations that sponsor interventions for energy efficiency have the expertise needed to conduct the necessary studies, or even to judge wisely between contractors competing to do such studies.

This article is intended as a tool for conservation program decision makers and analysts. It builds on the past work of the Committee on Behavioral and Social Aspects of Energy Consumption of the National Research Council, which has documented the behavioral processes and principles that so often produce surprises for energy policy makers.<sup>1,11,12</sup> The article is organized around four types of conservation policy—information programs, financial incentives, standards, and technological research and development. We describe 6 methods for answering behavioral questions that arise with each kind of policy and discuss the appropriateness of each method, given present knowledge, for addressing each type of behavioral question.

## 2. SIX ANALYTIC METHODS

### *Traditional energy demand models*

Energy demand models are analytic tools in which mathematical equations are used to estimate how demand might respond to various policy choices. Such models can address a wide range of policy questions and reveal unanticipated effects of policies on other parts of the energy or economic system. They may yield quantitative answers, often quickly. When correctly formulated, models can provide necessary checks of consistency with physical and economic constraints that might otherwise be overlooked in policy analysis. But, along with their well-known strengths, social-system models in general have many limitations,<sup>13–16</sup> as do most energy policy models.<sup>17,18</sup>

In policy analysis, models are most appropriate for anticipating effects of quantitative interventions that operate by processes that are well understood or that have been successfully modeled in similar situations. Often, however, not enough is known to defensibly quantify the variables, or the path of implementation is not straightforward. In such cases, the use of existing models cannot be easily justified. For example, few data exist on how consumers receive or act on information. To estimate the effectiveness of an information program, a modeler might adjust the price elasticity or a lag coefficient as a proxy for the program's effect, but to do this is to assume the program's effect rather than to estimate it. Models based on the economic theory of information and consumer search may improve the situation once an empirical basis is developed for choosing among the search strategies consumers may plausibly use.<sup>19–21</sup>

### *Analysis of existing data*

The Energy Information Administration and the utility companies have extensive data on residential and commercial energy use. Such data are useful for relatively quick and low-cost analysis of relationships that are represented in the data set, such as responses to fuel price changes or to incentives offered in different conservation programs. But analyses of existing data are limited by the data available. For example, data can be found on appliance purchases and list prices, but information on costs of production is held by manufacturers as proprietary. Utility records accurately report energy use, but usually lack information on consumers' incomes, demographic variables, or behavior. And there have been problems getting access to existing data at the individual level because of concern

about privacy. Better data exist for analyzing energy use in the residential sector than in the commercial or industrial sectors; aggregate data are generally more available than disaggregated data; and energy use data are better than data on equipment stocks, with data on attitudinal factors even less adequate.

The value of existing data also depends on its level of aggregation in relation to the question at hand. Disaggregated data on residential consumption can be aggregated to compare utility service areas or states in which different programs, incentives, or regulations are in effect. Such comparisons can be valuable if interpretations are made with sufficient care.

### *Surveys*

Surveys of energy users and other relevant populations, such as manufacturers, lenders, architects, and building owners, can give information about responses to new technologies, programs, and policies, both before and after they are introduced. Surveys are particularly good for assessing qualitative variables such as awareness and trust of information or the attractiveness of a new technology or program. They are also valuable for interpreting observational data. Data on miles driven in the family car or money spent on a new energy-efficient home may reflect a variety of behaviors or decision processes, and surveys can help reduce ambiguity. And surveys can ask such questions of a sample that is representative of a population of interest.

But surveys suffer from some generic limitations. Respondents may give socially acceptable rather than accurate responses. Surveys may fail to predict behavior because respondents' answers are based on faulty memories of what they have done or because they are unable to predict what they will do. Unreliability increases when surveys are used to assess responses to a hypothetical situation (e.g. a planned information program) or to predict behaviors that involve many steps before completion (e.g. expensive investments in energy efficiency).

In the U.S. Government, surveys present a practical problem because of the difficulty and delay involved in getting approval from the Office of Management and Budget for survey instruments. The requirement for approval, which rests on the rationale of reducing the burden on respondents, has stimulated researchers to develop various alternatives to the usual survey approaches: respondents have been paid, which satisfies concerns about undue burden; surveys have been funded by the National Science Foundation, whose procedures for protecting human subjects satisfy the same concerns; and data have been collected by utility companies, state governments, or other independent groups.

### *Ethnographic methods*

Detailed, open-ended interviews such as anthropologists conduct when trying to understand foreign cultures sometimes offer insights into behavior when it is not yet clear which behaviors or beliefs are most important to understand. For example, ethnographic interviews revealed that many people think of energy in budget-based units, such as dollars per month, rather than in energy units.<sup>22</sup> This finding was a revelation to some analysts, who were designing information programs on the assumption that people would easily understand physical units. The ethnographic approach is also useful for getting a first approximation to the decision processes of individuals or organizations.<sup>23</sup> As understanding of the issues becomes clearer, research can move from ethnographic approaches to more quantitative methods, such as surveys or small-scale experiments.

Focused group discussion is a technique developed by marketers that combines features of both survey and ethnographic methods. A trained leader directs a discussion among 10 or so members of a population whose response to a program element or product design is of interest. The participants' comments are used as a rough gauge of the reactions of the group they are presumed to represent. Like ethnography, focused group discussion

does not involve representative sampling, and like ethnography, it can yield an early qualitative indication of people's reactions. Focused group discussion is not as systematic as survey research, and it is not always less expensive, but it usually allows faster data collection.

### *Small-scale controlled experiments*

The experimental approach has been generally neglected in energy policy analysis. The best-known exception has been the time-of-use pricing experiments conducted during the 1970s, some of which involved random assignment of households to experimental electricity rates. Experimentation was the method of choice in those studies because there was no empirical basis for modeling the effect of prices based on time of use and because the experimental rates were so far from most energy users' past experience that self-reported intentions could not be relied upon. The same rationale suggests that experiments could provide the most valid answers to many questions about the design of energy information programs and about the marketing and implementation of conservation programs.

The greatest advantage of experiments over other research techniques is that they allow control over large numbers of extraneous variables whose effects make the interpretation of non-experimental data difficult. This is the situation with most conservation programs. For example, most evaluations of the Residential Conservation Service (RCS) have treated it as a single, uniform program and have attempted to make summary judgements about the RCS concept. But the variation among nominally identical programs is more striking than the averages,<sup>5</sup> and policy success depends on understanding and replicating program success where it occurs. Many factors that may be responsible for success can easily be varied in low-cost experimental field trials, for example, telephone marketing or the use of community groups or private companies to provide program services.<sup>24,25</sup> Strong inferences can be drawn from trials comparing new and existing approaches to program management.

The experimental approach is inexpensive relative to full implementation of a program or policy. In the context of an already planned pilot program, an experiment requires only normal evaluation efforts and the addition of special care in assigning participants to programs and in making data on program participants comparable with data on suitable comparison groups.

The experimental method has had difficulties as a policy tool. Experimental studies often meet practical opposition from program managers who are eager to get on with their programs and who feel they know enough to act without awaiting the results of formal research. Experiments also face political opposition on the ground that if the policy is a good one, it should be made available to all, not just a small experimental group. Moreover, if experimental subjects believe an experiment to be temporary rather than a permanent change in policy, this may affect their behavior.

An ethical question is sometimes raised about experimenting with human populations because participants in some experiments will benefit relative to participants in others. There are often ways to avoid such problems. For some policies (such as utility rate reforms), it is possible to use crossover designs in which participants take turns living with each experimental rate so that all receive the same set of incentives. Or a program can be offered to the control group after a delay to minimize the differential benefit. When it is not possible to equalize incentives, it becomes necessary to judge what the public and prospective participants will consider fair. A procedure for doing this was demonstrated in the Wisconsin time-of-use electricity pricing experiment. To see if it was possible to randomly assign households to different electric rates in a way that was ethically acceptable to the public, the research team convened random samples of people to judge the fairness of alternative rate structures for the experiment. The juries, and eventually the participants, agreed that it would be fair to set rates so that the average household in each group would experience no change in bills if it did not change its times of using electricity.<sup>26</sup> This jury approach may be applicable to determining the fairness of potentially controversial experimental approaches before conducting the experiment.

### *Evaluation research*

Evaluations of past and present energy programs are a great untapped source of knowledge, not only about what works but about the reasons for successes and failures. Outcome evaluations quantitatively estimate overall program effects and can be used for cost-benefit analysis. For example, they may measure rates of participation in a program, sales of a new technology, improvement in the energy efficiency of building shells, or the net energy savings from a policy or program. Process evaluations examine the way a policy or program is implemented rather than focusing on its final effect. They usually involve surveys, close observations, and interviews of program staff and clients and can offer insight into why a program succeeded or failed that cannot come from an outcome evaluation.

Evaluations of conservation programs have usually focused on outcomes rather than on the processes of marketing and implementation that can mean the difference between success and failure. But process and outcome evaluations can be used together to tell which features of a program were responsible for its outcome and to suggest promising revisions for programs.

Evaluation research can use any of the methods outlined above. The most reliable information comes from explicitly treating programs and policies as experiments from their beginning. To do this, an evaluation plan would include creation of a suitable comparison group, randomly assigned if possible, and careful measurement of effects in all groups. When random assignment is not feasible, some quasi-experimental research designs retain many of the advantages of controlled experiments. A full discussion of issues in evaluation research design can be found in the text by Cook and Campbell.<sup>27</sup> More can be learned from the experience of a program if an evaluation plan is developed as a program is developed. An evaluation plan tacked on after a program has been operated inevitably produces weaker research because of the inability to measure preprogram conditions and because important questions must be answered from memory or by reference to incomplete archives rather than by observation.

Evaluation studies can often be strengthened by using several research methods in concert. Surveys are ideal for getting participants' reactions to a program, which can only be inferred from data on energy savings or participation rates. In process evaluations, open-ended interviews can identify critical features of a program that both researchers and program operators have failed to anticipate. And small-scale experiments with program elements can be very informative as part of an evaluation even if the overall evaluation does not use an experimental design.

### 3. A STRATEGY FOR ASSESSING BEHAVIORAL ISSUES

Because energy efficiency depends on the decisions of numerous individuals and organizations to produce, market, and adopt energy-efficient technologies, policy interventions need to take into account all the relevant actors and their choices. The best way to do this is to mold the intervention to increase its acceptability to those actors. This strategy requires repeated and structured interaction from the start between the developers of the program or policy and those who are its targets. It is best described by an example, a home energy rating system. One would begin designing such a system by asking potential users what they would like to learn from a rating. One might conduct open-ended discussions or ethnographic interviews with bankers, builders, realtors, homeowners, and so forth, to understand how they evaluate and compare homes and to generate ideas for types of ratings. Then the potential users could be asked to respond to proposed ratings in a focused group discussion or survey format, to rule out some rating systems as unacceptable. Ratings that pass the screening could be tried in a more realistic setting on a few houses with user reactions assessed by open-ended interviews or surveys. The most attractive ratings can be tried in the field with experimental controls, using surveys to assess the reactions of the relevant populations.<sup>28</sup>

At each stage of this procedure, the list of options is narrowed and their presentation is made more realistic. Data collection moves from open-ended to more tightly controlled

research methods—from ethnographic interviews and discussions to surveys and then to experiments. At each stage, however, more than one method of research may be appropriate. Surveys and ethnographic methods are useful for learning what issues concern people, and surveys can help assess reactions to alternative versions of a program in pilot stages. However, the experimental method offers the most definitive knowledge of what specific versions of a policy or program work best.

The above strategy is also appropriate for incentive and regulatory programs and for the development of new technologies. Manufacturing firms rely on market research to assess the reactions of distributors and customers. Government, however, has sometimes failed to look carefully enough at what is acceptable before promoting policies and technologies. The failure of federal building energy performance standards in the United States is traceable to insufficient communication between government and the building industry, with the result that many in the industry felt the standards did not address legitimate concerns. Similarly, a screw-in fluorescent bulb developed with government funds in the 1970s met market resistance because the developers had not given enough attention to the problems of introducing a \$7 product to a 50¢ market.

Designing programs and technologies by involving the potential users has added advantages: It gives the users early information about the existence of the innovation, simplifying the marketing task later on, and it tends to commit the users to the version that they helped choose. It follows that it is important to involve individuals or groups that are influential with other members of the target population for the new program, policy, or technology.

The remaining sections discuss methods for answering the behavioral questions that arise in policy analyses of energy information, incentives, standards, and technologies. This four-fold classification is somewhat artificial because many programs have aspects of more than one policy type. Thus, the important behavioral questions for any one policy or program may be found in more than one of the following sections. In each section we offer judgments about the usefulness of different methods for answering the behavioral questions. The judgments, which are summarized in Tables 1–4, are conditioned on the present state of knowledge and the current adequacy of analytic tools.

#### 4. INFORMATION AND INFORMATION PROGRAMS

Some prominent examples of information programs are the Residential Conservation Service and other energy audit programs, home energy ratings, and appliance labeling. The success of these programs depends on the effect of new or improved information on major expenditures by energy users. To design and implement such programs effectively, five types of behavioral questions must be addressed (see Table 1).

*How can a program be designed so that the information it offers is used?*

The effects of energy information depend not only on completeness but on credibility, specificity, comprehensibility, vividness, and other qualities.<sup>11</sup> For analyzing the effects of such factors, existing data sets are irrelevant and existing quantitative models are almost useless. Currently available models tend to assume information to be complete or at least constant or to subsume its effects under other explanatory concepts, such as elasticity, discount rate, or time lag. To gain understanding for the purpose of designing information, it makes more sense to address the behavioral questions directly, using non-modeling approaches.

Surveys and ethnographic methods are promising. Ethnographic interviews can yield fruitful hypotheses about the way people understand energy use that can be refined and tested for generality with survey data.<sup>22</sup> For example, surveys can be used to identify householders' misconceptions about energy used in their homes and estimate the prevalence and magnitude of the misconceptions.<sup>29</sup>

Experiments can yield even more definitive knowledge about qualitative factors in energy information. For example, experiments varying the source from which people receive

Table 1. Appropriateness of six analytic methods for addressing behavioral issues related to the effects of energy information and information programs

Method	Issue					
	What Qualities of Information Affect its Use?	What Spreads Information?	Forecasting Effects of Programs	Assessing Effects of Programs	Understanding Causes of Program Effects	
Demand models	Not appropriate	Potentially of some value	Potentially valuable	Not appropriate	Not appropriate	
Analysis of existing data	Not useful	Not useful	Valuable	Not appropriate	Not appropriate	
Surveys	Especially valuable	Especially valuable	Not useful	Somewhat valuable	Especially valuable	
Ethnographic methods	Especially valuable	Valuable	Not useful	Somewhat valuable	Especially valuable	
Small-scale experimentation	Especially valuable for specific comparisons	Valuable	Valuable in pilot phases	Especially valuable	Valuable for validation	
Evaluation research	Not appropriate	Not appropriate	Valuable	Especially valuable	Somewhat valuable	
Outcome evaluation	Valuable	Especially valuable	Somewhat valuable	Somewhat valuable	Especially valuable	
Process evaluation						

information quantify the effect of a choice of program sponsors.<sup>24,30</sup> Such knowledge cannot now come from models and would not be as convincing if it came from surveys.

Evaluations of information programs can offer uniquely valuable knowledge if interviews or surveys are used to determine how information about a program reached people and how they responded to that information. Even more convincing results can come from evaluations in which experimental controls are used to study some aspect of the information offered.

#### *How can a program be designed to spread information widely?*

Community groups are being used increasingly to implement energy programs because of a belief, supported by case studies, that such groups spread information more successfully.<sup>5,31</sup> The key to their success seems to be their credibility and their reliance on word-of-mouth communication. Two strategies are useful for learning more about the spread of information in a particular context. One is to ask people, using surveys or ethnographic methods, how and from what sources they get their information. The other is to use different methods of spreading information in a field setting and measure the results. The second strategy gives more reliable results but can involve much more effort. It is relatively easy to implement in the context of a program evaluation. For example, an evaluation of the Minnesota Residential Conservation Service program compared rates of requests for audits for utilities that had the audits performed either by utility personnel, private contractors, or community groups.<sup>31</sup>

#### *How can the effects of a program be forecast?*

There is no good theoretical basis for forecasting the effects of information; the only reasonable approach uses data from past programs and judgments about differences and similarities between those programs and the one whose effects are to be predicted. As implemented, most government energy information programs have had small effects or none; stronger effects may be expected only by adopting some of the more effective implementation techniques that have been demonstrated in various studies.<sup>11</sup>

#### *How can the effects of a program be assessed accurately?*

The effects of information on energy use are best measured directly; self-reported behavior is less reliable. But metering energy use gives an incomplete picture of the effects of a program. Certain non-energy outcomes (e.g. comfort) that can alter the energy effects of programs or influence subsequent participation should be measured along with energy use.

The most effective outcome evaluation compares participants in a program both with non-participants in the program and with similar consumers who are not served by the program. Comparison with eligible non-participants gives an index of direct effects of the program, although the possibility of self-selection complicates interpretation of the results in most research designs; comparison with consumers not served allows a researcher to identify contagion effects in which a program effects non-participants through their indirect knowledge of it. Although each of these comparisons offers valuable information, such quasi-experimental studies are not definitive.<sup>27</sup> It is useful to build some experimental control into a program, for example, by offering information to different clients in different forms. If control is not built in from the beginning, evaluation researchers usually arrive on the scene too late to use the experimental approach.

#### *To what can program effects be attributed?*

To answer this question adequately requires a process evaluation in combination with an outcome evaluation. Process evaluations can help explain outcomes, especially when there is variation within a program (e.g. Bonneville Power Administration's residential incentive programs, which have been administered in different ways by the participating utilities). After-the-fact questions to participants can give valuable insight into the reasons for program success or failure but, because participation can change the ways people make

sense of their experience, self-reports must be interpreted cautiously. The way to be sure of conclusions from a process evaluation is to alter the program based on those conclusions and observe the effects.

## 5. INCENTIVE PROGRAMS

Federal and state governments have offered tax incentives for energy efficiency, and utility companies have offered energy loan and rebate programs. Such incentives are effective if they encourage investments that would not otherwise have been made, and their success in turn affects the need for other conservation programs. There are five behavioral questions about incentives that should be considered (see Table 2).

### *How does investment change as a function of the size of an incentive?*

Existing models can be useful for estimating the effect of any given size of incentive, but there is some reason to question the usual assumption that the price-demand relationship is log-linear.<sup>5,32</sup> In addition, size may be a less important factor than awareness that an incentive exists.<sup>5,33</sup> Careful evaluation of these possibilities with existing data should be done to make models more reliable. Surveys offer only weak data on the effect of incentive sizes because people can only compare incentives in hypothetical situations. Experimental methods are a better alternative.

### *How does investment depend on the type of incentive offered?*

Energy models usually involve equating different types of incentives (e.g. loan, rebate, tax credit) on net present value criteria, implicitly assuming that only the size of an incentive matters. But consumers may respond differently as a function of other financial features of incentives: a grant reduces first costs while a long-term loan can prevent negative cash flow. Also, different kinds of consumers probably have different preferences between types of incentive.<sup>5</sup> Surveys of preferences among incentive types involve hypotheticals, so responses can only be suggestive. A more effective way to assess preferences is by comparing consumer responses to programs offering different types of incentives.<sup>3,5</sup> The most reliable knowledge can come from experiments that offer consumers a choice of incentives of different types but of equal value. This could be readily done in the context of an evaluation study of an ongoing incentive program.

### *What programmatic factors affect consumers' use of incentives?*

Non-financial features of incentive programs, such as the availability of technical assistance, consumer protection features, the credibility of a program's sponsor, or the quality of interaction between clients and program personnel may be critically important to a program's success.<sup>5,24</sup> After an incentive has been offered, surveys or open-ended interviews of users and non-users can help illuminate the roles of such factors. Valuable insights about non-financial features of programs have also come from evaluation studies that analyze programs offering a single incentive but administering it in different ways.<sup>34,35</sup>

As noted in section 4, experiments can often yield quite precise assessments of non-financial factors by manipulating them in the context of a program. For example, a program can give special training to some energy auditors and not others, offer additional promotional services on a random basis, or experiment with other marketing or implementation innovations.

### *How much investment would have occurred without the stimulus of an incentive?*

Surveys reporting what investments people believe they would have made in the absence of an incentive must be interpreted with extreme caution. A more reliable approach is to compare people to whom an incentive was offered with people who did not have this incentive but who were otherwise similar. This can be done by adding a comparison group to a program evaluation design. Because of self-selection of program participants, a comparison of eligible non-participants is less than satisfactory. A comparison group of

Table 2. Appropriateness of six analytic methods for addressing behavioral issues related to incentives for energy efficiency

Method	Issue				
	Effect of Size of Incentive	Effect of Type of Incentive	Effect of Nonfinancial Factors	Incremental Effect of Incentive	Effect on Pace of Investment
Demand models	Valuable	Potentially of some value	Potentially of some value	Potentially valuable	Potentially valuable
Analysis of existing data	Especially valuable	Especially valuable	Valuable	Potentially of some value	Potentially of some value
Surveys	Not useful	Somewhat valuable	Valuable	Somewhat valuable	Somewhat valuable
Ethnographic methods	Not useful	Somewhat valuable	Valuable	Not useful	Not useful
Small-scale experimentation	Especially valuable	Especially valuable	Especially valuable	Especially valuable	Valuable
Evaluation research	Somewhat valuable	Especially valuable	Somewhat valuable	Especially valuable	Especially valuable
Outcome evaluation	Not appropriate	Somewhat valuable	Especially valuable	Somewhat valuable	Somewhat valuable
Process evaluation					

people who took advantage of the incentive later is an improvement,<sup>36</sup> but there remain problems of comparability. Realistically, triangulation on the answer through different methods is probably the best approach.

*To what extent does an incentive increase the pace of investment?*

The limited evidence suggests that incentives speed investment, but more needs to be known to see if the effect is large enough to justify particular incentive programs. Quantitative models do not yet have a strong enough empirical base to be reliable; other research methods must also be used. The best approach is probably evaluation research that follows carefully chosen comparison groups on a yearly basis to infer when program participants might have made the changes they made if the program had not been available.

## 6. STANDARDS

Policy analysts may be called on to evaluate energy-efficiency standards for buildings or appliances, or energy-efficient building codes. Table 3 summarizes the appropriateness of the six analytic methods for addressing four major behavioral questions related to energy efficiency standards.

*Under what conditions does energy efficiency influence consumers' purchases?*

If energy efficiency is an explicit consideration when consumers choose buildings or appliances, better information will make their decisions more economically rational in terms of energy. If energy is not being considered, however, standards will be a more effective way to increase efficiency.

Surveys or interviews of purchasers about the factors they consider are not definitive, but they can give useful information. Surveys of salespersons, dealers, and manufacturers may also give useful information. The role of energy efficiency considerations can also be estimated by calculating implicit discount rates from data on purchases of equipment for which standards might be set. High implicit discount rates indicate that energy efficiency is not a major influence on purchases. This approach does not, however, provide information on the conditions under which efficiency may become more influential.

*How might alternatives to standards, such as appliance labels or energy ratings, make energy efficiency a prominent consideration in purchase decisions?*

This question essentially concerns energy information. In addition to the methods noted in Section 4, laboratory experimentation can be used to assess whether a label or rating might make energy efficiency a more prominent part of the decision process. Consumers could be confronted with a hypothetical purchase decision and asked to request information in order of importance until they have enough to make a decision. Being hypothetical, this approach has limits: it is better for ruling out unattractive alternatives than for choosing a final label or rating. The effects of ratings and labels are most accurately assessed by field trials that use experimental methods in realistic situations and by program evaluation.

*How do the circumstances and purposes surrounding a purchase affect the importance of energy efficiency to a purchaser?*

Standards may be needed more for some appliances than others: People probably do not consider energy efficiency when replacing a furnace or water heater that suddenly breaks down, but there may be more time for comparison shopping and more chances to influence purchases with information with refrigerators or dishwashers purchased in non-emergency situations. Useful knowledge can be gained by simple surveys asking homeowners, builders, building owners, or other purchasers what factors they consider in purchasing particular appliances or other technologies. The implicit discount rate approach can also be used. If the implicit discount rate for air conditioners is about 20% and that for water heaters is about 150%,<sup>37</sup> the difference may be due to circumstances of the

Table 3. Appropriateness of six analytic methods for addressing behavioral issues related to standards for energy efficiency

Method	Issue				Forecasting Purchase Decisions
	When Does Efficiency Affect Purchases?	Effect of Labels and Ratings	Effect of Circumstances of Purchase		
Demand models	Not appropriate	Not appropriate	Potentially valuable		Especially valuable
Analysis of existing data	Valuable	Not useful	Especially valuable		Valuable
Surveys	Especially valuable	Valuable	Especially valuable		Potentially valuable
Ethnographic methods	Somewhat valuable	Especially valuable in pilot phases	Valuable		Not appropriate
Small-scale experimentation	Somewhat valuable	Especially valuable in pilot phases	Somewhat valuable		Potentially of some value
Evaluation research	Not appropriate	Valuable	Not appropriate		Not appropriate
Outcome evaluation	Not appropriate	Somewhat valuable	Not appropriate		Not appropriate
Process evaluation	Not appropriate	Somewhat valuable	Not appropriate		Not appropriate

purchase: one appliance may be purchased mainly by homeowners for their use and the other mainly by contractors for resale. Combining data from surveys with analysis of existing data provides a check on the results of each method.

*In the absence of standards, how do manufacturers, builders, and others make choices?*

To evaluate the need for standards one must know about the production and purchase of equipment in the absence of standards. There is a special need to know more about how appliance manufacturers use information about competition, expected energy prices, and market characteristics in deciding whether to develop new product lines. There is also a need to know more about how the choices of builders, distributors, contractors, and retailers affect purchase decisions.

For aggregate forecasts, quantitative modeling is the method of choice. However, existing models need a stronger empirical basis for their assumptions about behavior, particularly the behavior of purchasers: it is clear for appliance purchases that a simple assumption of cost-minimization does not do justice to the complexity of the phenomena.<sup>11,37</sup> The needed knowledge can come from research on the three previous questions.

## 7. TECHNOLOGICAL RESEARCH AND DEVELOPMENT

New technologies are constantly being developed for building construction and retrofits and for use in appliances. Behavioral questions arise because the practical effect of any new technology depends on choices about its purchase and use. Adoption decisions, in turn, depend on whether estimates of energy savings from the new technology are reliable. It is hard to make such estimates when the energy savings depend not only on the operation of the technology but on the behavior of the people who use it. For example, super-insulated houses save energy, but if people open windows frequently to freshen the air, savings will be much less than expected. Table 4 summarizes the appropriateness of the six analytic methods for addressing two behavioral questions about energy-efficient technologies.

*Which energy-efficient building technologies are most likely to be readily accepted in the market?*

This is a market research question. For example, a heat reclaimer for flue gases may be easier to build as a new product than to include in a redesigned furnace or water heater, but the market for heat reclaimers to retrofit on flues may be very small compared with the market for energy-efficient furnaces or water heaters with built-in heat reclaimers.

Energy models are appropriate for estimating the economic costs of producing technologies and the energy saved by adopting them. But acceptance is also influenced by many factors those models do not address: the prices manufacturers charge for a piece of equipment with a given production cost; the rates of adoption of the new technology as a function of its consumer features; the marketing efforts of manufacturers and dealers; and so forth.

Surveys and ethnographic methods are especially useful for identifying design features that would be attractive to potential manufacturers or purchasers. Reactions of those groups to designs or prototypes can help guide choices of design modifications, which can be market tested while still in the prototype phase. As a new technology moves toward implementation, surveys and small-scale experiments become more useful for refining the design, just as they do for policies and programs. Design options can be subjected to experimental trial by users to assess public acceptance in the same way they are subjected to engineering tests of their costs and efficiency of operation. When new technologies are being introduced in conjunction with specific technology transfer efforts, evaluation research is appropriate for assessing those efforts.

Table 4. Appropriateness of six analytic methods for addressing behavioral issues related to research and development of energy efficient technology

Method	Issue	
	What Factors Enhance Acceptance?	Estimating Behavioral Component of Energy Savings
Demand models	Somewhat valuable	Potentially valuable
Analysis of existing data	Not useful	Not useful
Surveys	Especially valuable	Somewhat valuable
Ethnographic methods	Especially valuable	Valuable
Small-scale experimentation	Especially valuable	Especially valuable
Evaluation research		
Outcome evaluation	Not appropriate	Not appropriate
Process evaluation	Valuable in technology transfer programs	Not appropriate

*How can reliable estimates of energy savings be developed for new technologies?*

Energy use in a building can change by 100% when the occupant changes.<sup>38</sup> This fact is a warning against estimating energy savings from a new technology without observing how it works in field conditions when operated by people like the intended users. Even holding the user constant, engineering estimates are imperfect because when new energy-efficient technologies consume less energy, they also free income that can be spent on other things, some of which also use energy. This issue is amenable to modeling,<sup>39,40</sup> though data needs sometimes are serious limitations.<sup>41</sup>

To assess the effect of a new technology on behavior, it is useful to study it in normal use. Since only a few consumers can be involved in trying prototypes, ethnographic approaches, which gain the deepest insight from the fewest consumers, may be the method of choice for understanding reactions to prototypes. An experimental approach, comparing relevant behaviors before and after adoption of a new technology with behavior of comparable energy users without the technology, becomes useful as more prototypes become available for trial. Data collected in a few small experiments may be enough to validate or refine the assumptions of models, which may then become fully appropriate for forecasting the effects of new technology on behavior.

## 8. CONCLUSION

No conservation policy or program can be evaluated realistically without examining a broad range of behavioral issues. However, policy analysts have not yet given high priority to studying the processes of choice among consumers, manufacturers, builders, and other important actors. Moreover, few have the necessary expertise. Despite these problems and the reality of resource limitations, useful studies of the behavioral issues can be done. The framework outlined above can guide such research on the human dimension of energy efficiency in buildings and may thus help policy makers avoid some of the unpleasant surprises and disappointments that have been so frequent in conservation policy in the past.

## REFERENCES

1. P. C. Stern, *Energy Efficiency in Buildings: Behavioral Issues*. National Academy Press, Washington, D.C. (1985).
2. U.S. Department of Energy, Residential Conservation Service Evaluation Report, U.S. Department of Energy, Washington, D.C. (1984).
3. E. Hirst, *Pub. Admin. Rev.* **44**, 421 (1984)
4. E. Hirst, L. G. Berry, and J. Soderstrom, *Energy* **6**, 621 (1981)
5. P. C. Stern, E. Aronson, J. M. Darley, D. H. Hill, E. Hirst, W. Kempton, and T. J. Wilbanks, *Eval. Rev.* **10**, 147 (1986).
6. P. C. Stern, L. G. Berry, and E. Hirst, *Energy Policy* **13**, 133 (1985).
7. C. Goldman, in *Doing Better: Setting the Agenda for the Second Decade*, Vol. B., pp.107–121. American Council for an Energy-Efficient Economy, Washington, D.C. (1984).
8. C. Goldman and B. S. Wagner, in *Families and Energy: Coping with Uncertainty*. (Edited by B. M. Morrison and W. Kempton), pp.657–672. Institute for Family and Child Study, Michigan State University, East Lansing, Mich. (1984).
9. E. Hirst, D. White and R. Goeltz, Comparison of Actual Electricity Savings with Audit Predictions in the Bonneville Power Administration Residential Weatherization Pilot Program. Oak Ridge National Laboratory Report ORNL/CON-142, Oak Ridge, Tenn. U.S.A. (1983).
10. E. Hirst and R. Goeltz, Comparison of Actual and Predicted Energy Savings in Minnesota Gas-Heated Single-Family Homes. Oak Ridge National Laboratory Report ORNL/CON-147, Oak Ridge, Tenn. U.S.A. (1983).
11. P. C. Stern and E. Aronson, *Energy Use: The Human Dimension*. Freeman, New York (1984).
12. P. C. Stern, *Improving Energy Demand Analysis*. National Academy Press, Washington, D.C. (1984).
13. W. Ascher, *Forecasting: An Appraisal for Policy-Makers and Planners*. Johns Hopkins Univ. Press, Baltimore, Md. (1978).
14. G. D. Brewer, *Behav. Sci.* **28**, 166 (1983).
15. D. Freedman, *J. Bus.* **54**, 479 (1981).
16. M. Greenberger, M. A. Crenson and B. L. Crissey, *Models in the Policy Process: Public Decision Making in the Computer Age*. Russell Sage Foundation, New York (1976).
17. D. Freedman, T. Rothenberg and R. Sutch, *J. Bus. Econ. Stat.* **1**, 24 (1983).
18. P. C. Stern, *J. Pol. Anal. Mgmt.* **5**, 200 (1986).
19. J. Hirshleifer and J. G. Riley, *J. Econ. Lit.* **17**, 1375 (1979).
20. S. Salop and J. Stiglitz, *Rev. Econ. Stud.* **44**, 493 (1977).
21. L. L. Wilde and A. Schwartz, *Rev. Econ. Stud.* **46**, 543 (1979).
22. W. Kempton and L. Montgomery, *Energy* **7**, 817 (1982).
23. R. R. Wilk and H. L. Wilhite, *Energy* **10**, 621 (1985).
24. R. D. Miller and J. M. Ford, Shared Savings in the Residential Market: A Public/Private Partnership for Energy Conservation. Energy Task Force, Urban Consortium for Technology Initiatives, Baltimore, Md. (1985).
25. M. D. Polich, in *Doing Better: Setting the Agenda for the Second Decade*. Vol. G., pp.141–151. American Council for an Energy-Efficient Economy, Washington, D.C. (1984).
26. J. S. Black, The Role of Social Scientists in Field Experiments on Energy. Paper presented at the *Meeting of the American Psychological Association*, New York (September 1979).
27. T. D. Cook and D. T. Campbell, *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Houghton-Mifflin, Boston, Mass. (1979).
28. A. D. Ackerman, M. K. Cox, L. J. Schuck and E. J. Tarini, The Massachusetts Home Energy Rating System. Pacific Northwest Laboratory Report PNL-4763. Richland, Wash. (1983).
29. W. Kempton, C. K. Harris, J. G. Keith and J. S. Weihl, *Marr. Fam. Rev.* **9**, 115 (1985).
30. C. S. Craig and J. M. McCann, *J. Consum. Res.* **5**, 82 (1978).
31. G. Gaskell and R. Pike, Consumer Energy Conservation Policies and Programmes in Britain. International Institute for Environment and Society, Science Center Berlin, CECF Technical Reports Vol. VI. Berlin, Federal Republic of Germany.
32. D. H. Hill and P. C. Stern, Effects of Price Changes on Automobile Driving: An Exploration of Alternative Behavioral Assumptions. Commission on Behavioral and Social Sciences and Education, National Research Council, Washington, D.C. (1986).
33. T. A. Heberlein and G. K. Warriner, *J. Econ. Psychol.* **4**, 107 (1983).
34. D. I. Lerman and B. H. Bronfman, Process Evaluation of the Bonneville Power Administration Interim Residential Weatherization Program. Oak Ridge National Laboratory Report ORNL/CON-158. Oak Ridge, Tenn. (1984).
35. D. I. Lerman, B. H. Bronfman and B. Tonn, Process Evaluation of the Bonneville Power Administration Residential Weatherization Pilot Program. Oak Ridge National Laboratory Report ORNL/CON-138. Oak Ridge, Tenn. (1983).
36. T. M. Newcomb, *Eval. Rev.* **8**, 425 (1984).
37. H. Ruderman, M. D. Levine and J. E. McMahon, The Behavior of the Market for Energy Efficiency in Residential Appliances Including Heating and Cooling Equipment. Lawrence Berkeley Laboratory Report, Berkeley, Calif. (1985).
38. R. C. Sonderegger, in *Saving Energy in the Home*, (Edited by R. H. Socolow). Ballinger, Cambridge, Mass. (1978).
39. J. A. Dubin, *Consumer Durable Choice and the Demand for Electricity*. North-Holland, Amsterdam (1985).
40. J. A. Dubin and D. L. McFadden, *Econometrica* **52**, 345 (1984).
41. A discussion in the context of energy-efficient appliances can be found in ref. 12, Chap. 5.