Energy Efficiency (2008) 1:79–104
DOI 10.1007/s12053-008-9009-7

Feedback on household electricity consumption: a tool for saving energy?

Corinna Fischer

Received: 3 December 2007 / Accepted: 28 March 2008 / Published online: 6 May 2008

Abstract Improved feedback on electricity consumption may provide a tool for customers to better control their consumption and ultimately save energy. This paper asks which kind of feedback is most successful. For this purpose, a psychological model is presented that illustrates how and why feedback works. Relevant features of feedback are identified that may determine its effectiveness: frequency, duration, content, breakdown, medium and way of presentation, comparisons, and combination with other instruments. The paper continues with an analysis of international experience in order to find empirical evidence for which kinds of feedback work best. In spite of considerable data restraints and research gaps, there is some indication that the most successful feedback combines the following features: it is given frequently and over a long time, provides an appliance-specific breakdown, is presented in a clear and appealing way, and uses computerized and interactive tools.

C. Fischer (✉)
Fachbereich 2 - Bauen, Energie, Umwelt,
Verbraucherzentrale Bundesverband,
Markgrafenstr. 66,
10969 Berlin, Germany
e-mail: fischer@vzbv.de

Keywords Consumers · Electricity consumption · Energy conservation · Feedback · Electricity bill · Advanced metering · Literature review

Sustainable electricity consumption: a Herculean task?

Electricity seems a particularly difficult area within which to promote sustainable consumption; and households seem a particularly difficult target group. In Germany, for example, the household sector is the one with the fastest growing end energy consumption. Electricity consumption, especially, is rising even faster than total end energy consumption.

Sustainable electricity consumption, in this context, comprises different things. First, it may mean choosing electricity from renewable or other less environmentally detrimental sources (which will not be addressed in this article). Secondly, it means a conscious choice of appliances and of their duration and modes of use with the ultimate goal of curbing overall consumption[1]—in short, electricity conservation. Stimulating electricity

[1] All energy scenarios, e.g., for Germany, agree that a sustainable energy system is impossible without significant cuts in overall consumption (Enquête-Kommission 2002; DLR et al. 2004; DIW et al. 2005).

conservation is a difficult task, because electricity differs in significant ways from other consumer goods. It is abstract, invisible, and untouchable. It is not consumed directly but indirectly via various energy services. Electricity consumption is therefore not perceived as a coherent field of action. Rather, it involves activities as diverse as listening to music, cooking meals, working with the computer, or making a phone call. Moreover, electricity conservation is not limited to the act of using electricity but starts with choosing and purchasing energy-using appliances like a TV set, washing machine, computer equipment, or electric heater. In each of these activities, conservation means a different set of behavioral modifications. It is difficult for the consumer to link all these various activities and develop a coherent, comprehensible, and concise cognitive frame of what "electricity conservation" could mean in everyday life.

The invisibility of electricity also means that the consumer usually receives little feedback on her consumption—she does not experience the "diminishing stock" and does not find herself in control of her consumption. Also, electricity's qualities—including its ecological features—cannot be directly perceived, making it hard for the consumer to develop an emotional involvement. It is hardly a product to be proud of, to show around, or to worry about. Consumers regard electricity as a necessary, but unspectacular everyday product of which security of supply is important, but specific features do not matter much. In contrast to products like organic food or sustainable housing, sustainable electricity consumption can therefore not easily become an element of lifestyle (Birzle-Harder and Götz 2001). And neither do its costs usually make up for an important share of a household's budget. Thus, all in all, electricity turns out to be a "low interest" product.

Consumer feedback as a road to sustainable consumption?

One idea for supporting sustainable electricity consumption is to improve feedback on consumption, on its cost, and its environmental impacts. Today, such feedback is far from what it could be. Kempton and Layne (1994) equate consuming electricity to shopping in a grocery store in which no individual item has a price marking, and the consumer receives a monthly (or, in many countries, even annual) bill on an aggregate price for "food consumption". She has no idea how, when, or by which appliances electric current was used. Nor is she informed whether her consumption is relatively high or low (which could stimulate a search for reasons), or whether it has increased or decreased (and thus, whether her actions had any effect).

Feedback may be improved in various ways. Possibilities include increasing the frequency of feedback, providing a time-, room- or application-specific breakdown, improving the visual design, or adding further information, for example, time series, comparisons with an average, or information about environmental impact.

As shown by a number of international model projects and scientific studies, such improved feedback can help to repair the problems associated with electricity conservation. In an excellent review of experience, Darby (2006) has found that improved feedback may reduce consumption by up to 20%. Recently, EU policy has been taking on such encouraging experience: EU Directive (2006/32/EC) on energy end-use efficiency and energy services, dating from April 2006, calls for informative billing and other types of feedback, "where appropriate" (see the "Conclusions").

The present article builds on existing review work on feedback (Darby 2001; Roberts and Baker 2003; Abrahamse et al. 2005; IEA 2005). It re-analyzes relevant projects and studies reviewed by these articles as well as some additional literature not yet covered. Its aim is threefold. First, it wants to contribute to a more theoretically guided understanding of why and how feedback works. For this purpose, a psychological model of environmentally relevant behavior is presented and tentatively linked to the topic of feedback. Secondly, the article would like to shed some light on the question of why results of individual studies on feedback differ so much and what it is that causes feedback to succeed (or fail). To achieve that, relevant dimensions are identified that differ between the various studies. The differences relate to the context and method of the respective project, but more importantly, to design features of the feedback itself. Linking these features to the psychological model, some hypotheses are derived on how feedback needs to be designed in order to achieve optimum results. Empirical evidence is sought from the studies and research gaps are identified. In doing so, the paper pursues its third, methodological objective: to comment more critically and in more detail on the available database than existing articles do, allowing the reader to judge results more carefully.

Some theory

Environmental psychology has developed various models to explain environmentally relevant behavior and provide a basis for successful behavioral change. Matthies (2005) has reviewed theory and findings from all over the discipline and integrated them into a heuristic model of environmentally relevant behavior (see Fig. 1). This integrated model can be helpful for explaining why and how feedback on electricity consumption can reduce consumption.

The model distinguishes between two types of action: routinized or habitual behavior (here presented in the form of "environmentally detrimental habits" on the bottom of the figure) and conscious decisions (represented in the mid and upper part of the figure). Habitual behavior is behavior that is not reflected upon. It is performed regularly in the same way. Many of our everyday activities are habitual and this is also true for most electricity-consuming activities: We switch on the light, stove, or electric heating without thinking; we use the washing machine the way we have learned to use it, we have a routine of throwing clothing into the dryer after washing, or of switching on the radio or TV when we come home. Habitual behavior is functional because it spares us the time and effort of decision-making on issues that re-occur regularly and for which we have once learned or worked out a way of how to solve them. However, habitual behavior may also deliver sup-optimal results because we have never really thought about an optimum way to do it. Or if we have, the situation may have changed in the meantime and the behavior may not be appropriate any more. In the field of environmentally relevant behavior, many habits are environmentally detrimental because the environment was not a relevant issue to consider at the time the habit was formed, or because beliefs about environmental effects held at that time have been shown to be wrong, or because the situation has changed so that a once beneficial behavior is not useful any more.

For new norms and considerations to enter the decision-making process, a *conscious decision* needs to be taken. This means that habits must be broken up somehow. A person must realize that there are various options to choose from, and that norms and criteria are needed for evaluating those options. This process

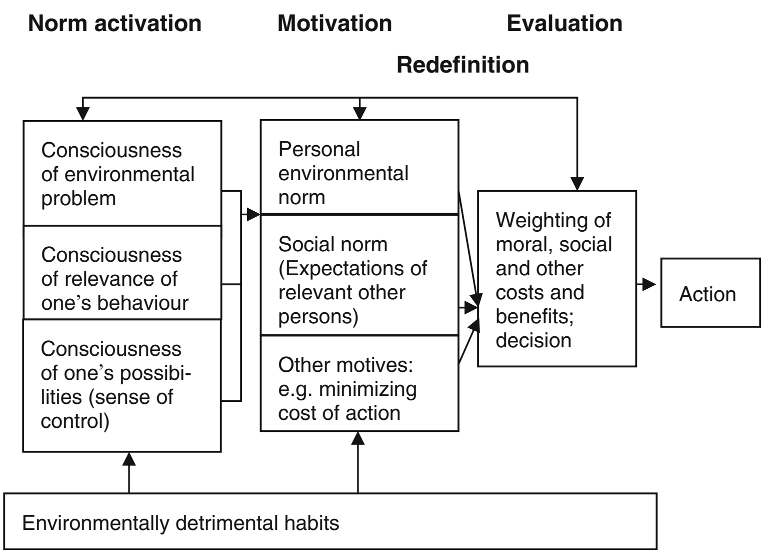

Fig. 1 Heuristic model of environmentally relevant behavior. Source: Matthies (2005; own translation)

is called *norm activation*. According to the model, norm activation is made up of three building blocks. First, the person must realize that there is a *problem*[2]. We realize that there is a problem if a habitual solution is not available, or does not work out any more, or is being questioned.

After realizing that there is a problem, two further steps are necessary to complete norm activation: A person must realize that his or her *behavior is relevant* to the problem, and s/he must become conscious that s/he has *possibilities to influence* his or her behavior and its outcomes (in the literature, this is usually called *sense of control*). Only then will s/he reflect upon changing his or her behavior in order to solve the problem. For example, if a person realizes that her annual electricity bill is high, but attributes this fact to electricity prices and not to her consumption behavior, she will not make a decision on this behavior. The same is true if she attributes the cost to her behavior, but does not know how to control it, e.g., because she has no idea which appliances caused the high consumption.

When norm activation is completed, a person enters into a process of weighing and evaluating different motives in order to reach a decision on how to act. The motives, according to the model, comprise personal norms, social norms[3], and “other motives”. *Personal norms* are personal ideas about how one should act. *Social norms*, in contrast, are ideas about which norms relevant others might hold. They are important because persons value social relationships and therefore orient their behavior, among other things, along lines of what they think is socially desired. Finally, there are a host of *other motives* which are not specified in detail in the model, but may easily be so. In our case, they may comprise the need or desire for all the services associated with electricity consumption, such as lighting, cooking, or heating. These services may also be associated with other, more general motives, such as a desire for comfort, relaxation, or efficient work organization. Finally, there may be desires to receive these services in certain ways, such as conveniently, cheaply, or reliably.

Norms may conflict with each other or with other motives. Therefore, a person must enter in an *evaluation* process; during which, moral, environmental, personal, or social costs and benefits are weighed. During this process, norms and motives may also be redefined in the light of the available information. As a result, the decision for a certain—more or less environmentally beneficial—action emerges. Under specific conditions, such an action may be performed regularly and develop into a new habit or routine.

What is not explicitly mentioned in the model is the fact that considerable *information* is necessary in order to perform the decision process. Basically, a person must know about the nature of the problem, the existing options and their respective consequences, in order to judge them in terms of norms and motives.

Thinking in categories of such a model enables us to detect in which ways feedback can operate. First, it can direct attention towards electricity consumption, demonstrating to the consumer how much electricity everyday activities consume. This way, consumers are confronted with a *problem* that requires conscious decision. A door is opened for reflecting their behavior, questioning habits, and receiving arguments. Depending on its form and content (see “Types of feedback” section for details), feedback can frame this problem in different ways (e.g., as a problem of wasting money, or of damaging the environment); thus, influencing the reasoning process.

Feedback can also increase the consciousness of the *relevance of one’s own behavior*. The more closely electricity consumption can be linked to specific appliances and activities, the clearer the relevance of behavior becomes. Appliance-specific feedback can help the consumer to detect how a certain appliance or a certain way of using it affects the amount of electricity consumed and the money spent. This also increases the *sense of control* because the consumer can find out how changes in behavior or appliance stock affect the outcome.

[2] As her aim is to explain environmentally conscious behaviour, defined as solving environmental problems, Matthies describes the problem more specifically as an *environmental* problem. However, electricity consumption, although it has important environmental impacts, cannot solely or even predominantly be conceptualized as behaviour that is directed at solving an environmental problem. Therefore, I think it is more appropriate in our context to choose a more general approach by analyzing how conscious decisions *generally* come about, and only in a second step, how environmental considerations may enter the process.

[3] Here again, Matthies specifies the norms involved as “environmental norms”. For reasons explained above (footnote 4), I prefer a more general approach.

On the level of motivation, improved feedback can activate *other motives* conducive to electricity conservation. Depending on how it frames the problem, feedback can activate a desire for cost savings or for minimizing environmental impact. Comparative feedback (as described in more detail below) can stimulate a sense of competition. To improve the incentive character even more, feedback could be combined with other instruments, like price incentives, goal setting, or a contest.

From these considerations, one can deduct the hypotheses that feedback is most effective if it:

- successfully captures the consumer's attention
- draws a close link between specific actions and their effects
- activates various motives that may appeal to different consumer groups, such as cost savings, resource conservation, emissions reduction, competition, and others.

In the following section, I will review empirical evidence on the effects of feedback and interpret it along the lines of these theoretical considerations.

A review of international experience

Database

The review presented here covers five review studies (Darby 2001, 2006; Roberts and Baker 2003; Abrahamse et al. 2005; IEA 2005) and 21 original papers on the effects of feedback on electricity consumption and on consumers' reactions, attitudes, and wishes concerning such feedback. The criteria for the choice of papers were as follows: in order to retain some topicality, I restricted myself to papers dating from the last 20 years, that is, from 1987 onward. I also confined the analysis to projects that were explicitly designed for giving feedback (e.g., via the meter, displays, or the bill) and excluded broader approaches where feedback may come indirectly as a by-product (e.g., energy advice or community learning). Finally, feedback solely designed for the purpose of load shifting (usually as a complement to time-of-use pricing), was also excluded, focusing instead on feedback designed to have (also) an effect on overall consumption. Insofar as they fulfill these criteria and were available (which was a problem sometimes), the papers discussed in the five reviews were included[4]. They have been complemented by some additional papers not yet covered by those reviews, mainly from German-speaking or Nordic countries. A list of the papers analyzed is included in the references list.

All in all, the original papers cover 26 projects from ten countries: the USA (three), Japan (two), and many Northern and Western European countries (Denmark (four), Finland (two), Germany (two), the Netherlands (one), Norway (three), Sweden (six), Switzerland (one), UK (two)). Thus, there is a remarkable lack of knowledge from Southern European and Accession countries.

Project results depend on the project's goals, on methodological aspects (such as study design and sample), and on the different features of the feedback itself, such as frequency, content, breakdown, presentation, inclusion of comparisons, and combination with additional information and other instruments. The papers reviewed vary widely in all these respects. In the following sections, I will systematize the original papers[5] according to these aspects in order to give a better understanding of the information base.

Project goals

By providing feedback on electricity consumption, one may pursue different goals. Motivating and enabling households to lower overall consumption is the most prominent one, but feedback is also given with other goals. This must be kept in mind when evaluating results, as different methods of feedback may have different success with respect to the various goals.

In the projects reviewed, the main reasons for giving feedback were:

- to enable and motivate households to conserve energy, or to "stimulate ecological behaviour" (17 projects)
- to improve customer satisfaction or service (five projects, three of which in combination with energy conservation)

[4] Due to language constraints, only English and German papers could be considered. As the paper by Darby (2006) became available within short notice, some of the references cited there could not be considered.

[5] I do not include the review studies in this section. However, the insights gained from the review studies will enter the reasoning process when developing conclusions.

- to achieve load shifting or peak shaving (two projects, both in combination with energy conservation)
- to raise consumers' "consciousness" (one project)
- to explore consumer preferences, trying to detect what kind of feedback households would like to have on their electricity bills (two projects)
- or, less specifically, to test any "effects" of improved feedback (two projects, one of which in combination with energy conservation).

Study designs and samples

The *study design* determines the sort of questions that can be asked, the sort of answers that may be provided, and the ecological validity of the results. We can broadly distinguish between more research-oriented and more applied projects. Research-oriented projects try primarily to test the implication of a theory/theories or to fill knowledge gaps left open by earlier research. The design follows the research question and does not care too much about applicability in the real world, for example regarding cost efficiency, intensity of labor, or technical requirements for the solutions that are being tested. In these cases, the approach is usually systematic and results are reliable. But on the other hand, problems may arise with putting the knowledge into practice. On the other hand, more practically oriented projects tend to test one or two preferred and applicable solutions. The proven solutions are therefore easier to implement, but on the other hand, the information is often rather unsystematic and there may be a lack of scrutiny.

For a systematic evaluation of different sorts of feedback, designs with a control group and several experimental groups exposed to various types of feedback are ideal. With respect to ecological validity, model projects or field experiments tend to be superior to laboratory studies or surveys (although they, too, may use designs that are difficult to implement in the real world).

In our database, 15 projects have a more scientific and ten a more applied character. Two studies rely on surveys alone, focusing on preferences and attitudes (Egan 1999; Sernhed et al. 2003). One is a laboratory study (McCalley and Midden 2002). By far, the majority of the projects are *model projects* or *field experiments* (the boundary between both being blurred, though).[6] Fourteen of those employ a design with multiple groups, using several experimental groups and/or an experimental group and a control group.[7] Eight others are restricted to the implementation and evaluation of one specific feedback method. For purposes of data gathering, the model projects and field experiments use various combinations of different methods: electricity consumption measured by electronic meters, external meter readers, or customers themselves reported conservation activities, and all sorts of surveys, questionnaires and interviews dealing with issues such as satisfaction, attitudes, preferences, and conservation motivation.

Knowing the size and type of the *sample* is important for assessing the representativeness of the result. Households are not all the same—cultural, social, or political variations may make crucial differences. For example, there are indications that feedback works very differently in different social milieus (Nielsen 1993). The projects reviewed here cover quite a range of different household types in terms of household size, features of the building, appliance stock, ownership, income, and social status. In a number of projects, this mixing is deliberately done in order to achieve a representative sample. This broad array allows some assessment of the generalizability of results. On the other hand, with regard to project size, the situation is not as good. Many field experiments include no more than ten to 50 households. This leads to subgroups being very small (around ten households) and raises questions about the significance of results. Some studies (Dobson and Griffin 1992; Haakana et al. 1997; Brandon and Lewis 1999; McCalley and Midden 2002) include around 100–120 participants, but by splitting them into several subgroups, again, arrive at rather small subgroups. A number of big field experiments with over 1000 participants are not reported in great detail (Henryson et al. 2000). For one experiment (Karbo and Larsen 2005), there are only preliminary results

[6] By a "field experiment", I mean a project conducted with the core purpose of generating information and insight on the effects of feedback. It has a more scientific character. A "model project" means a project that is being conducted with the core purpose of testing a certain type of feedback in practice.

[7] Three of them have only one experimental group and a control group, and one has several experimental groups but lacks a control group.

available. This leaves us with nine to ten reasonably well-documented projects with big samples for analysis: three field experiments (Sexton et al. 1987; Nielsen 1993; Arvola et al. 1993), five model projects (Garay and Lindholm 1995; Wilhite and Ling 1995; Wilhite et al. 1999; and, with some restrictions, Dünnhoff and Duscha 2008) and two surveys (Egan 1999; Sernhed et al. 2003).

All in all, there is a lack of projects that are both scientifically and ecologically valid. Such projects would use a representative sample, systematically vary the feedback given, and use control groups in order to trace back the effects of specific types of feedback. They would use feedback options that could be implemented in real life, and provide a detailed documentation of their methods and results. Unfortunately, most reviewed projects lack one or the other of these features (see for more details "Research gaps" section and the "Conclusions").

Types of feedback

On the basis of our model, we theorized that successful feedback has to capture the consumer's attention, to link specific actions to their effects and to activate various motives. If this is the case, then different characteristics of the feedback itself become relevant, among them, its frequency, content, breakdown, presentation, inclusion of comparisons, and combination with additional information and other instruments. The feedback described in the papers varies widely with respect to these features. The following section gives an overview.

Frequency and duration From the model, it would follow that feedback is more effective, the more directly after an action it is given. Quick feedback would improve the link between action and effect, and therefore, increase consciousness about the action's consequences. Furthermore, persistent effects would be more likely if feedback is given over a longer time, because new habits can form during that time. In the reviewed projects, the frequency of feedback ranges from continuous to bimonthly with eight projects giving feedback more often than monthly (six of which continuously), five projects giving it monthly and nine projects giving it less often (it is not reported for all projects). With respect to duration, there is a very clear-cut division: In nine projects, the feedback is given for less than 3 months (usually 4–6 weeks; in two billing projects only once).[8] In another eight (including the rest of the billing projects), it is given over at least 9 months (up to one or several years).

Content Feedback may be given on electricity consumption alone (e.g. kWh), on cost, or on environmental impacts of consumption. The model suggests that these different contents frame the problem in different terms and thus activate different motives and personal and social norms. It remains an open question which motives and norms would be strongest in which target groups. In the projects reviewed, all three kinds of information are used, though the emphasis is on consumption and cost. Seventeen projects feedback consumption and cost; four, consumption only (Haakana et al. 1997; McCalley and Midden 2002; Mack and Hallmann 2004; Mosler and Gutscher 2004). Only two projects (Jensen 2003; Brandon and Lewis 1999 in one experimental condition) feedback environmental information.

Breakdown Providing a breakdown, e.g., for specific rooms, appliances, or times of the day is provided, is almost the only way of providing a direct link between action and result and thus, establishing consciousness of the relevance of *individual* actions. However, only six of the reviewed projects provide some sort of breakdown while two restrict themselves to a single appliance type anyway (cooking appliances in Mansouri and Newborough 1999; Wood and Newborough 2003; and washing machines in McCalley and Midden 2002). Sexton et al. (1987) and Dobson and Griffin (1992) provide a breakdown for all major appliances. Wilhite et al. (1999) test a breakdown for typical uses (lighting, heating...), based on interview data. Karbo and Larsen (2005) use a daily load curve, based on measured data, and an appliance-specific breakdown, based on interview data, both upon request. And Ueno et al. (2005 and 2006) provide appliance- and time-specific breakdowns (daily and ten-daily load curve) upon request, based on real consumption data.

[8] The projects by Ueno et al. (2005, 2006) actually lasted longer, but have been evaluated only at one early point of time, namely after they had been running for 4 weeks (or 6 weeks, respectively).

Medium and mode of presentation Our model does not directly alert us to the relevance of the medium and way of presentation. However, it has long been clear from communication sciences and learning theory that the way information is presented is crucial for its adoption (Roberts and Baker 2003). The reason is of course that the information needs to capture attention and be understood before it can become effective. Two basic media may be used: electronic media and written material. Electronic media is used in eight studies, taking different forms. One relatively unique approach is to install an electronic display directly at an appliance, which can provide information about the consumption of this particular appliance (Mansouri and Newborough 1999; McCalley and Midden 2002; Wood and Newborough 2003). Also, an electronic, maybe interactive, meter may show the total consumption of a household, provide additional information such as time-specific breakdown or cost (Sexton et al. 1987; Jensen 2003). Another approach is to use computer and internet as interactive tools. A computer program is supplied with data that may stem from user input (e.g., on household size, appliance stock) and/or from metering of actual consumption data, and can provide the user upon request with a broad range of information, e.g., load curves, appliance-specific breakdown, comparisons, or energy-saving tips (Dobson and Griffin 1992; Brandon and Lewis 1999; Karbo and Larsen 2005; Ueno et al. 2005, 2006). Advantages of electronic feedback are its flexibility (being able to react to users' demands, and showing different kinds of information upon request), and its ability to quickly process and present actual consumption data. Interactive tools may also stimulate users' curiosity and experimenting. On the other hand, electronic feedback may be difficult to access for users not used to electronic media, and interactive tools require more user involvement.

Written material may come on its own in the form of direct mailings, brochures, etc. This is done in four projects (Haakana et al. 1997; Brandon and Lewis 1999; Jensen 2003; Mack and Hallmann 2004). Another possibility, used by nine projects, is to use the electricity bill as a carrier of feedback information. This approach seems promising because it can be expected that the bill is read more carefully and raises more interest than additional material. Such efforts are described in Arvola et al. (1993), Garay and Lindholm (1995) Wilhite and Ling (1995), Wilhite et al. (1999), Egan (1999), Henryson et al. (2000), Dünnhoff and Duscha (2008).

Equally important is the way of presentation. Much depends on the comprehensibility and appeal of text or graphics. The projects apply numerous variants of presentation, the most common being text, load curves, bar charts or pie charts (for an application-specific breakdown or comparisons in time and with other households), and horizontal lines or bell curves (for comparison with other households). Here, the devil is very often in the details. Most projects do not seem to reflect these problems: the choice of a specific design is usually not discussed at all nor are reasons given for a specific choice. Only two projects test design variations systematically (Egan 1999; Wilhite et al. 1999).

Comparisons There are two basic types of comparisons: historic comparison relates actual to prior consumption (often, temperature-corrected, with the same period in the previous year). Normative comparison compares consumption to that of other households (e.g., with a national or regional average, households in the neighborhood, or households that are in some way similar, e.g., in size, type of house, application stock). Comparisons may stimulate specific motives for energy conservation, for example, a sense of competition and ambition. They also make transparent whether consumption in a certain period or of a certain household is "out of the norm", thereby capturing the consumer's attention, alerting him to a potential problem and activating the search for reasons and redress. Almost all reviewed projects present, or deal with, historic comparison. Twelve studies also deal with normative comparisons.

Additional information and other instruments Feedback is very often combined with other instruments which makes a lot of sense from a theoretical point of view. Information on consumption will not work without a motivation to conserve, which may be provided by other instruments like financial incentives (Sexton et al. 1987; Nielsen 1993), goal setting (McCalley and Midden 2002; Mosler and Gutscher 2004), or personal commitment (Mack and Hallmann 2004; Ueno et al. 2005, 2006). On the other hand, feedback will not work if households have no idea on what they can do about their consumption. This problem may be remedied by additional information on how to save energy; ideally, closely connected to the appliance or situation on which feedback is given. Most projects use or explore such additional information (with the exception of Dobson and Griffin 1992; Egan 1999; Jensen 2003; Sernhed et al. 2003, and the Tibro project reported in Henryson et al. 2000).

Method and methodological problems

My aim was, first, to find out whether feedback works at all, and secondly, how it must be designed to work best. The latter is methodologically very challenging. Projects can only be with greatest care. As has become clear from the above, they differ markedly with respect to study design, sample, and method of data gathering, differences occurring both in substance and in scientific elaborateness. What is more, results are not always reported quantitatively or in sufficient detail to make a comparison. And if they are reported, studies use very diverse reporting schemes. They vary in baseline, in time and duration of measurement, and in the unit for which savings are reported. Table 1 summarizes the studies, giving an overview of the reporting schemes used.

To arrive at some conclusions, I first checked “best cases”, that is, the projects or experimental conditions which produced highest savings. For this purpose, I grouped studies according to their reporting schemes, so that studies with at least roughly comparable schemes fell into the same group (see Table 1). For a comparison, I used all studies that reported average savings of an experimental group as compared to a control group (the other groups of studies were too small and too heterogeneous to allow for a meaningful comparison). Within this selection of studies, I identified the projects or experimental conditions in which the difference between experimental group and control group was highest. In addition, I also identified the experimental conditions providing highest savings *within* each study that worked with several experimental groups. This way, I could identify which design features were present in those two kinds of “best cases” and could thus be regarded as supportive for success. However, results were not too clear because some of the same features were also present in less well-performing projects. Therefore, as a second approach, I took each design feature at a time, sorted the cases according to whether they include it or not, and compared the performance of the cases including that feature with that of the cases not including it. Taken together, these two steps provided some preliminary conclusions on the relationship between success and design. In the following section, I present those conclusions. Results of existing review articles are incorporated in the discussion.

Results

Does feedback work?

One result, at least, seems clear: feedback stimulates energy (and specifically, electricity) savings. Not all studies discuss actual savings; but those who do generally find savings ranging from 1.1% to over 20%. Usual savings are between 5 and 12%.[9]

However, in a few instances, no savings were found. To look carefully at these examples teaches us something about the preconditions for feedback to work.

In the project described by Dünnhoff and Duscha (2008), electricity customers received a one-time supplement to their annual bill that provided normative comparisons and energy-saving advice. Apparently, this kind of one-shot feedback was too unobtrusive to raise attention and too loosely linked to concrete actions to help consumers enhance control.

In the study of Sexton et al. (1987), the main purpose was load shifting. Feedback accompanied the test of a tariff structure where peak and off-peak tariffs differed considerably (between 3:1 and 9:1). Feedback informed consumers about their current use and projected cost per hour, and a light signal alerted them to the switch between peak and off-peak hours. Apparently, the feedback showed to customers that electricity was unexpectedly cheap in off-peak hours and stimulated heavy load-shifting activities. Thus, the savings that occurred in peak periods were, all in all, cancelled out by increased off-peak consumption.

Nielsen (1993) found that almost no savings occurred in a working-class area with small flats,

[9] Information on statistical significance of the findings is often lacking, but the sheer number of studies which report savings is a good indicator for the general effectiveness of feedback.

Table 1 Overview of reviewed literature

Source (scientific/applied)	No. of projects	Sample	Method (type of feedback, data collection)	What is reported?	Reporting scheme: Unit for which savings are reported	Reporting scheme: Baseline	Reporting scheme: Measurement period and time(s)	Results
Reporting of average reduction compared to historic baseline								
Nielsen (1993) (S)	1	1,500 households in three Danish regions (Jütland and Kokkedal: middle class, detached houses; Odense: working class, flats) in three experimental groups	Field experiment. Various combinations of written advice, feedback on consumption and cost via monthly meter self-reading, personalized energy audit, financing audit & increased tariffs. Measuring of yearly consumption, questionnaires	Electricity consumption	Average reduction in each experimental group during each treatment year and total average	Baseline interval: 1 year before treatment	Monthly during the three treatment years	Reductions in Jütland and Kokkedal Group 1 (“advice, feedback, energy audit, financing audit, increased tariffs”): 8–10% savings Group 2 (“advice, feedback, financing audit, increased tariffs”): 7–9% savings Group 3: (“advice, feedback, energy audit, financing audit”): 6–8% savings Reductions in Odense: Group 2: 3–6%, Group 3: 0–4%
Ueno et al. (2006) (S)	1	9 households in Japanese neighborhood	Field experiment. Computerized interactive tool with daily feedback on consumption and cost; breakdown. Measurement of electricity consumption, monitoring of appliance usage and feedback tool usage	Electricity consumption	Average reduction of participants. Percentage that took conservation activities	Baseline interval: average of forty weekdays before begin of treatment	Average of forty weekdays after begin of treatment	Electricity consumption reduced by 9% (appliances with feedback 12%, appliances without feedback 5%)
Reporting of average reduction compared to control group								
Sexton et al. (1987) (S)	1	480 Californian households testing new tariff structure (of which 68 with feedback, of which 51 analyzed)	Model project on time of use tariffs. In feedback households: continuous consumption and cost monitoring. Continuous	Electricity consumption	Average consumption of participants (and control group)	Control group that received ToU pricing but no feedback	Monthly averages during the 10-month feedback period	During 9 of the 10 months, HH with feedback consumed more than HH without feedback This was due to heavy load shifting: The rise

			measurement of electricity consumption, telephone survey					took place in off-peak periods, while in peak periods, feedback HH consumed on average 1.2% less than control group (reduction was significant during 6 of the 10 months) Difference to control group: 0
Dobson and Griffin (1992) (S)	1	100 US households (25 in experimental group, 2 control groups)	Field experiment. Continuous feedback on consumption and cost, broken down to various appliances and time intervals. Measurement of consumption	Electricity consumption	Average consumption in each group	Control group	Average of 60 day treatment period	12.9% less consumption than control groups Difference to control group: 12.9 percentage points (pp)
Wilhite and Ling (1995) (A)	1	1,450 Oslo households (of which 1285 remained) in three experimental groups and one control group. Response rate to questionnaire not reported	Model project. Improved electricity bill with various combinations of increased frequency, historic comparison and advice. Measuring of consumption, questionnaire, interviews	Electricity consumption, customer satisfaction, understanding of bill	Average reduction in each experimental group. Reported conservation activities	Control group	Yearly average during each of the three treatment years	Experimental groups consume 7.6% less than control group in second year and 10% less in third year Experimental groups report various conservation activities more often than control group. Customer satisfaction and understanding of bill rose markedly Difference to control group: 7.6/10 pp
Reporting of average reduction compared to historic baseline and control group								
Arvola et al. (1993) (A)	1	696 Helsinki households in three experimental groups and one control group. 550 answered questionnaire or interview	Field experiment. Improved bills (actual consumption, historic comparison) and advice. Measuring of electricity consumption, interviews, questionnaire	Electricity consumption, satisfaction and comprehension of bill	Average reduction	Control group, Yearly average in year before treatment	Yearly average during 2 treatment years	Rising overall consumption compared to base year, but consumption in experimental groups rose less than in control group Exp. Group 1 increased 1.1 percentage points less than control group

Table 1 (continued)

Source (scientific/ applied)	No. of projects	Sample	Method (type of feedback, data collection)	What is reported?	Reporting scheme: Unit for which savings are reported	Reporting scheme: Baseline	Reporting scheme: Measurement period and time(s)	Results
								Exp. Group 2 increased 3.0 percentage points less Exp. Group 3 increased 4.7 percentage points less Satisfaction and energy awareness rose Difference to control group: 1.1/3.0/4.7 pp
Garay and Lindholm (1995) (A)	1	600 Helsingborg households in 4 treatment groups (differing in heating system); matching control group for each. Response rates to questionnaires not reported	Model project. Improved bill (consumption statistics, normative comparison, billing based on actual consumption). Meter readings, interviews, questionnaires	Electricity consumption, satisfaction with new bill, understanding of bill		Control group. Historic baseline mentioned but not described	Not reported	Not reported in detail. No significant differences to control group in electricity consumption (attributed to methodological problems). Satisfaction with bill rose markedly; respondents reported that the bill was helpful for saving energy Difference to control group: not significant
Haakana et al. (1997) (S)	1	105 single-family houses in five Southern Finnish regions. Three experimental groups, one control group	Field experiment. Various combinations of monthly feedback (written or video) and advice. Questionnaire on satisfaction and conservation activities, calculation of savings from the activities	Electricity consumption, conservation activities, satisfaction	Average and median reduction in each experimental group. Percentage of households that took certain activities. Percentage who are satisfied	Baseline interval: 3 months before beginning of feedback (temperature corrected). Control group	Monthly averages during the 17-month feedback period	Average reduction in first half of feedback period: “feedback plus video information” group: 21%, “feedback plus written information” group: 19%, “feedback only” group: 17%, control group: 14% Second half (also compared to baseline): 7%, 5%, 5%, 1%, respectively. 81–84% in exp. groups and 77% in

								control group took conservation activities. Ninety-eight per cent were satisfied with the feedback, 83% desire normative comparisons Difference to control group: first half 7/5/3 pp; second half: 6/4/4 pp
McCalley and Midden (2002) (S)	1	100 Dutch individuals in three experimental groups and one control group	Lab study. Computerized machine washing simulation. Feedback on consumption after each wash, in combination with self-chosen or assigned goal. Calculation of consumption per wash, questionnaire.	Electricity consumption per wash	Average reduction over 20 washes in each experimental group	Baseline "interval": Average of 6 pre-treatment washes. Control group	Twenty times (each wash)	Average reduction Group 1 "Feedback and self-chosen goal": 21.9% Group 2 "Feedback and assigned goal": 19.5% Group 3 "Feedback only" and "control group": about 10% each Difference to control group: 11.9/9.5/0 pp
Mack and Hallmann (2004) (S)	1	30 households in German neighborhood (19 experimental group, 10 control group, one excluded)	Field experiment. Weekly written feedback. Meter readings, interviews	Electricity consumption. Conservation activities	Average reduction of participants (and control group. Average number of conservation activities in participating households)	Baseline interval: average of 6 measurements during 3 months before treatment (temperature corrected). Control group	4 weekly measurements during intervention, 30 measurements over 10 months after treatment, divided in 5 intervals with 6 measurements each	Reduction during treatment period that can be attributed to the treatment: 2.9%. Reduction during five posttreatment intervals (altogether 10 months): between 1.5 and 3.6%. No reduction in control group. In average 4.8 conservation activities during treatment, 4.4 after treatment Difference to control group: 1.5–3.6 pp
Mosler and Gutscher (2004) (S)	1	48 individuals in four experimental groups and one control group in Swiss canton Zurich self-recruited	Field experiment. Feedback as daily meter self-readings, in various combinations with advice and goal setting. Evaluation via weekly	Electricity consumption	Average reduction in each experimental group	Baseline interval: 4 weeks before intervention. Control group	1 (or 4?) measurements during each phase: 4-week baseline, 4-week treatment	Treatment period: Group 1 ("advice, feedback and commitment"), gr. 2 ("advice and feedback") and gr. 3 ("advice only")

Table 1 (continued)

Source (scientific/applied)	No. of projects	Sample	Method (type of feedback, data collection)	What is reported?	Reporting scheme: Unit for which savings are reported	Baseline	Measurement period and time(s)	Results
			meter self-readings				and 4-week posttreatment	reduced by 10.2–10.9% each. Gr. 4 ("advice and commitment") reduced by 1.1%. Control group reduced by 4.8%. Posttreatment period: Gr. 1: 18.1%; gr. 2: 21.7%, gr. 3: 21.9%, gr. 4: 13.3%, control group:7.5% Difference to control group: treatment 5.4–6..7 pp; posttreatment: 10.6/14.2/14.4/5.8 pp
Ueno et al. (2005) (S)	1	19 households in Japanese neighborhood (10 experimental group, 9 control group)	Field experiment. Computerized interactive tool with daily feedback on consumption and cost; breakdown. Electricity consumption measurement, monitoring of feedback tool usage, questionnaire	End energy, gas and electricity consumption, conservation activities	Average reduction of participants (and control group). Percentage of participants that took conservation activities	Baseline interval: average of 28 weekdays before start of treatment. Control group	Average of 28 weekdays after begin of treatment	End energy consumption reduced by 12%, electricity consumption by 17.8% (control group 4.7.) Six of the ten households reduce standby consumption and seven reduce electrical heating Difference to control group: 13.1 pp
Dünnhoff and Duscha (2008) (A)	1	4,500 Heidelberg households in three experimental groups and one control group (249 usable questionnaires)	Model project. Supplement to electricity bill with normative comparison and written advice. Partly combined with personal consultancy. Postal survey before and after intervention, measuring of electricity consumption	Evaluation of the supplement, conservation motivation, general environmental motivation electricity consumption	Number of reported conservation activities, average reduction of each group	Average consumption in year before intervention. Control group	Average consumption in year of intervention and year after intervention (extrapolated from first 6 months)	More than three quarters of the respondents found the additional information on the bill interesting or useful About half of them felt motivated to obtain additional information about energy conservation No statistically significant group differences in number

								of reported conservation activities nor in electricity consumption. All groups reduced consumption by about 5% Difference to control group: not significant
Miscellaneous reporting schemes								
Brandon and Lewis (1999) (S)	1	120 UK households in seven experimental groups and one control group	Field experiment. Various combinations of media (paper, electronic) and content (cost, environment, comparisons). Meter readings, interviews, focus groups	Electricity consumption	No. of households in each experimental group who increased vs. decreased consumption. Each group's average change	Baseline interval: yearly average before begin of treatment	Average of 8 monthly readings during treatment and one final reading after treatment	Only in experimental group 6 (computerized feedback) there was a significant effect: 12 participants decreased consumption (by 31% on average), only 3 participants increased it (by 4% on average). In the other groups, about half of the participants increased and half decreased
Wilhite et al. (1999) (A)	2	(1) 2,000 Stavanger households; actual number of usable data sets not reported	(1) Model project. Improved bill with historical comparisons. Postal survey	(1) Customer satisfaction, attitudes, understanding of bill, activity: lowering room temperature	(1) Share of persons who lowered room temperature, share of persons who report increase/decrease in consumption, average electricity consumption	(1) One survey before treatment	(1) One survey immediately after treatment, another one two years later	Rising customer satisfaction and understanding of bill
		(2) 2,000 households in Oslo and Stavanger. About 25% overall return rate[a]	(2) Field experiment. Improved bill with normative comparisons and appliance-specific breakdown (averages, not individualized). Postal survey	(2) Customer satisfaction, understanding of bill, learning, motivation to conserve	(2) Share of persons who reported motivation to conserve	(2) None	(2) Survey sent out together with the feedback diagrams	(1) Immediately after treatment 6–8% rise in the number of people who report to lower room temperature at night or when they leave home 2 years after, recipients of historical feedback seem to have lowered overall consumption by 4%

Table 1 (continued)

Source (scientific/ applied)	No. of projects	Sample	Method (type of feedback, data collection)	What is reported?	Reporting scheme: Unit for which savings are reported	Reporting scheme: Baseline	Reporting scheme: Measurement period and time(s)	Results
								while total population increased consumption by 4%. 27% report that electricity consumption has decreased (11% report an increase) (2) High Interest in normative feedback, it is seen as motivating for conservation. High interest in appliance-specific breakdown. 72–77% "agree" that it would motivate them to conserve if their consumption was comparatively high
Mansouri and Newborough (1999), Wood and Newborough (2003) (S)	1	36 UK households (results only reported for 31) in three experimental groups and one control group[b]	Field experiment: Direct display attached to stoves feeding back consumption and cost. Measuring of electricity savings, some sort of survey/interview[c]	Electricity consumption, satisfaction	No. of households that saved more than a certain amount (total and within experimental groups)	Baseline interval: average of 56–89 days immediately before treatment (temperature corrected)	Average of treatment period (56–84 days)	Households appreciated feedback. Fourteen of 31 households saved more than 10%, six of those saved more than 20% Group 1 "Information only": Three of twelve saved >10% Group 2 "Feedback only": Seven of ten saved >10% Group 3 "Information and feedback": Four of ten saved >10%
Jensen (2003) (A)	1	Copenhagen working-class neighborhood.[c] Results reported for selected cases (=flats and blocks)	Field experiment. Improved meters and written material showing consumption and environmental impact. Meter readings	Electricity consumption[d]	Amount of reduction in selected cases; number of cases that follow the tendency	Baseline interval: yearly average before begin of treatment	Yearly average during the first year of treatment	Best case: 22% electricity reduction, two others in that size, five others (of nine total) follow the tendency

Reporting scheme unavailable								
Henryson et al. (2000)[d] (A?)	3	Households in Nordic countries (1 Danish, 2 Swedish), sample size between 1,000 and 1,400. Actual number of usable data sets not reported	Improved electricity bill with varying combinations of consumption statistics, historic comparison, advice, graphics. 6–12 bills annually. Meter readings	Electricity consumption	Average reduction of participants	Not reported	Not reported	2–12% savings
Karbo and Larsen (2005) (A)	1	3000 Danish households	Model project. Online tool giving comparative feedback on consumption and cost, load profile and appliance-specific breakdown. Measurement of electricity consumption planned	Electricity consumption	Results unavailable due to missing evaluation[e]	Results unavailable	Results unavailable	Results unavailable
Studies on attitudes								
Egan (1999) (S)	1	600 Delaware residents (survey), 580 customers of a cooperating utility (model project)	Research and model project. Various graphic designs for normative feedback were tested. Survey and interviews on consumers' comprehension and attitudes	Attitudes, comprehension of bill	Inapplicable			A bell curve in which data points were represented as little houses was judged as most attractive, comprehensible and motivating with respect to electricity conservation. However, positive judgment does not necessarily mean comprehension or motivation to conserve
Sernhed et al. (2003) (S)	2	(1) Customers of 3 utilities, 1,000 households each, 35% response rate. (2) National representative	Surveys	Preferences with respect to design of electricity bill	Inapplicable			Customers find bill difficult to understand, wish to receive bill based upon actual consumption. They are interested in a high

Table 1 (continued)

Source (scientific/ applied)	No. of projects	Sample	Method (type of feedback, data collection)	What is reported?	Reporting scheme: Unit for which savings are reported	Reporting scheme: Baseline	Reporting scheme: Measurement period and time(s)	Results
		telephone survey						consumption alert, graphic presentation and historic comparison. Only 50% wish normative comparison. Generally satisfied with the current (monthly to quarterly) billing frequency
NUTEK (1996), as described in Henryson et al. (2000) (S)	1	290 Swedish households	Quantitative survey (270 households) and qualitative interviews (20 households)	Preferences with respect to design of electricity bill	Inapplicable			Consumers suggest to have their bill based on real consumption, historic comparison and normative comparison (both in graphic form) and practical information about savings potential or different activities. Information overload should be avoided

[a] 50% return rate to a first questionnaire determining key figures for comparison, 50% to second—evaluation—questionnaire

[b] The reported size of the groups varies across the text

[c] It is not reported in detail how households' responses were obtained

[d] Sample size is not reported

[e] Water and gas consumption were also monitored. However, this paper deals only with electricity consumption

[f] All in all, the paper describes seven projects, but there are duplications with Nielsen (1993); Arvola et al. (1993); Wilhite and Ling (1995), and Garay and Lindholm (1995)

[e] Troels Fjordbak Larsen, personal communication

low income and low consumption. Other studies also show that households with a previously low consumption do not feel encouraged to conserve if they receive feedback—they might even increase their consumption (Bittle et al. 1979–1980; Brandon and Lewis 1999). On the one hand, there might just be no saving potential. On the other, the findings point to a relevant precondition for feedback to work: There must be a—implicit or explicit—motivation. Without a motivation to conserve, information about how well you perform in this discipline is useless. It may even be counterproductive, for example, when comparative or historical feedback shows that your consumption is relatively low (or has been dropping), signaling that there is space for improvement on comfort.[10]

Which types of feedback work best?

In the first evaluation step, I studied the design features of the "best cases". "Best cases" include best groups *within* studies, and also best projects *across* studies (see for details of the sample "Method and methodological problems" section). Table 2 and Fig. 2 give an overview.

The results must be qualified, though. Some of the studies work with very small samples. All of the "best cases" across studies must be taken with care because of the small sample size or the laboratory situation.[11] To deal with this situation, I generously split the sample of "best cases across studies" in half, so that best cases also include studies by Wilhite and Ling (1995) and Haakana et al. (1997). With all due care, some first deductions can be made that make theoretical sense.

- All designs[12] that provide *computerized feedback*, offering multiple feedback options at the user's choice (e.g., consumption over various time periods, comparisons, additional information like environmental impact or energy-saving tips) turn out to be "best cases".
- All designs that used an *interactive element* that engages households—through computerized feedback or through required activities like self-feedback or self-meter reading—made it to the "best case" group.
- All designs that provide *detailed, appliance-specific* breakdown are "best cases".
- Three of the four designs that give feedback *very often* (daily or more) are in the "best case" group.[13]

In the second step, I compared, for each design variable, the performance of the cases with different values of that variable.

Frequency I grouped projects into those that provided feedback less than monthly, monthly to weekly, and daily or more. It emerges that none of the "less than monthly", and all but one of the "daily or more" projects are among the best performing (as far as they can be compared). One project with continuous feedback (Mansouri and Newborough 1999) cannot be directly compared to the others, but suggests high savings in some cases. In the "weekly to monthly" group, there are some well-performing but also a number of quite low-performing projects. This indicates that immediate feedback could be very helpful while weekly to monthly feedback may be helpful, but is not sufficient for best performance on its own.

Duration There is no clear indication that long-term projects provide higher (initial) savings than short-term ones. However, it seems sensible to assume that long-term projects contribute to habit formation and can therefore engender more persistent savings (during, but possibly also, after treatment).

Content As almost all projects combine consumption and cost information, there is no basis for separating the effects of both kinds of information. However, one may look separately at the two projects that test the effects of environmental feedback. Jensen (2003) delivers eco-information to nine housing blocks in a Copenhagen working-class quarter. He reports elec-

[10] A fourth project that could not produce measurable savings is Garay and Lindholm (1995). The authors attribute this to methodological problems with the composition of the groups, though.

[11] Mosler and Gutscher (2004) themselves report that their findings are not statistically significant because the groups were too small.

[12] "All designs" in this section always refers to those designs that could be included in the comparison.

[13] The exception is Sexton et al. (1987) who, for reasons discussed above, differ a bit from the rest of the studies.

Table 2 Best cases within studies

Study	Group	Treatment	Treatment of the less successful group(s)
Nielsen (1993)	Group 1 in Jütland and Kokkedal (middle class)	Received a combination of advice, feedback, energy audit, financing audit, and increased tariffs.	Received no change in tariffs
	Group 2 in Odense (working class)	Received a combination of advice, feedback, financing audit, and increased tariffs (but no energy audit)	Received no change in tariffs
Arvola et al. (1993)	Group 3	Received a combination of billing for actual use, feedback, and saving tips.	Received no saving tips
Haakana et al. (1997)	Group 1	Received feedback plus video advice	Received feedback only, or feedback and written advice.
McCalley and Midden (2002)	Group 1	Received feedback and had a self-chosen conservation goal	Received feedback, but had an assigned goal or no goal.
Mosler and Gutscher (2004)	Groups 1–3	Received advice only, advice and feedback, or advice, feedback and commitment (goal)	Received advice and commitment (goal) only
Brandon and Lewis (1999)	Group 6	Computerized feedback	Various combinations of different media and content of feedback (but no computerized)
Mansouri and Newborough (1999); Wood and Newborough (2003)	Group 2	Received feedback only	Received advice and feedback, or advice only.

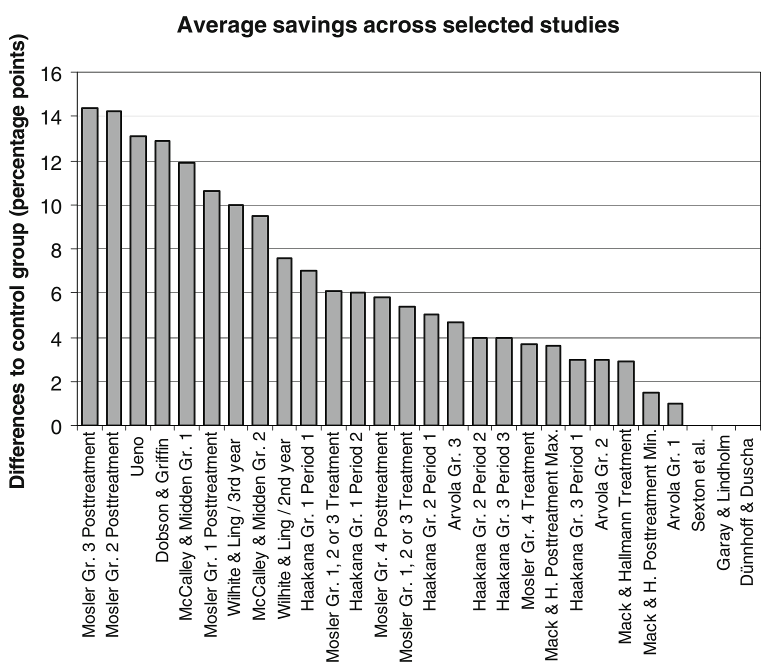

Fig. 2 Average savings (as compared to control group) across studies

tricity savings in the order of 20% against baseline for three cases; but unfortunately, no figures for the remaining ones. In Brandon and Lewis (1999), there is no significant difference between the "environmental information" group and other experimental groups. The findings, at least, suggest that environmental information may be as effective as other kinds of information. Our model would suggest tailoring the kind of information given to the potential motives and norms of the target group.

Breakdown Reliable data for the effectiveness of appliance-specific breakdown, again, is difficult to find. Of the seven breakdown projects, three (Mansouri and Newborough 1999; Wilhite et al. 1999; Karbo and Larsen 2005) provide no or no comparable data on savings. One (Sexton et al. 1987) is unsuccessful in promoting conservation due to its focus on load shifting. However, of the three remaining ones, two (McCalley and Midden 2002; Ueno et al. 2005) are among the most successful ones—a good indication of the potential usefulness of detailed, appliance-specific data.[14]

Medium and mode of presentation We have already seen that interactive, computerized feedback is very stimulating. Interactivity and the possibility of choice involve customers, raise their attention and allow for tailored solutions. It is less clear, however, what exactly the presentation must look like. Surprisingly, very few studies have considered the relevance of graphic design or formulation of text at all. Roberts and Baker (2003) suggest that the presentation should be simple but not simplistic, that it should not involve additional paper, and that a combination of text, diagrams, and tables is more effective than single-format presentations. This is a start, but there is not enough detail yet. The only two comparative studies show convincingly that households' reactions to graphical designs depend very much on the exact choice of diagram or chart type, labels, scale, symbols, and wording of the explanation. Designs may range from the completely unintelligible to the highly motivating (Egan 1999; Wilhite et al. 1999).

[14] It remains unclear, though, why the project by Ueno et al. (2006), which is a very similar project to Ueno et al. (2005), resulted in much lesser savings. Uncertainties due to the very small sample surely play a part.

A special case is the use of the bill as a medium. Only one of the billing projects (Wilhite and Ling 1995) is among those yielding the highest scores. Billing projects show quite a range of savings, from 0% (only one case) to 12%. However, they have other advantages. They can typically be implemented with comparatively little additional effort, and are therefore worth exploring for practical reasons (see, for political implementation, the discussion of the EU energy end-use directive in the "Conclusions"). Furthermore, they can be designed as long-term approaches, forming energy-conscious habits over time. One indication that such a long-term perspective could work is the success described by Wilhite and Ling (1995). It is the most long-running project, having been operative for 3 years.

Comparisons As almost all projects use some form of historical comparison, it is only worthwhile to look separately at normative comparison. It shows that none of the twelve studies dealing with normative comparison could demonstrate an effect on consumption so far. A simple reason presents itself: while it stimulates high users to conserve, it suggests low users that things are going not so bad and they may upgrade a little. These effects probably tend to cancel out each other. A similar argument may hold for historical feedback: it stimulates conservation only when consumption has risen.

Additional information and other instruments The theory postulates that motivating instruments (like goal setting, commitment, or financial incentives) and information on "how to" conserve must be present in order to make feedback work. The empirical evidence, though, is less clear (see also Table 2). With regard to motivation: on the one hand, Katzev and Johnson as early as 1987 highlighted the role of a commitment to save when they analyzed successful and unsuccessful examples of feedback. McCalley and Midden (2002) confirm in a laboratory experiment that feedback alone does not induce savings if it is not combined with a savings goal. However, in many studies, feedback alone seems to work. One project involving commitment delivers very small savings (Mack and Hallmann 2004), and one field experiment that explicitly tests the additional effect of commitment can find no such effect (Mosler and Gutscher 2004). In Nielsen (1993), financial incentives have very little effect. With regard to additional information, results are also very mixed. There are a number of studies in which it is of no use or even counterproductive (Wilhite and Ling 1995; Mansouri and Newborough 1999; Brandon and Lewis 1999; Mosler and Gutscher 2004) and only two in which it was explicitly helpful (Arvola et al. 1993; Haakana et al. 1997 (only the video advice)).

One methodological reason may be the small size of experimental groups. A possible substantial explication is that motivation and knowledge about energy-saving possibilities is already present to some degree in participating households, and can be activated by giving feedback. In this situation, additional information or tools may rather complicate the situation for participants and cause an "information overload". Other reasons lie in the design of specific studies (for example, a too unambitious goal rather discourages households from making further efforts, see Mosler and Gutscher 2004) Finally, as already reported, the usefulness of information depends strongly on how it is presented, and whether it is specific to the needs of the target group.

How would households prefer their feedback?

Most of the projects do not only study the quantitative effects of feedback, but also households' understanding, preferences, and needs concerning feedback. Some (Egan 1999, one study described in Wilhite et al. 1999; the NUTEK 1996 study reported in Henryson et al. 2000; Sernhed et al. 2003) focus exclusively on these aspects. Such aspects are important for building up customer satisfaction, but also for laying a fertile ground for motivating households to conserve electricity.

One unanimous finding is that households in all countries approve feedback that is more detailed and more closely linked to consumption actions. It gives them a sense of control and, if delivered with the bill, of being valued and well informed by their utility. A first important step is billing based on actual consumption (while electricity bills in many countries come in the form of estimates). Other valued aspects are a higher frequency of the feedback and an appliance-specific breakdown.

Furthermore, there is usually an interest in comparisons with one's own previous consumption. It is equally clear that households prefer information that is easy to understand, while they find their current electricity bills often hard to understand. Easy-to-understand information includes (the list is not exhaustive)

- feedback based on actual consumption in a given period (instead of offsetting with previous periods, prepayments, or estimates)
- clear labeling and explanation of labels, acronyms and technical terms
- clear indication of the various components of the electricity price
- support by graphic presentations which are also clearly labeled. For purposes of breakdown, pie charts are preferred. For comparisons with previous periods, households like vertical bar charts. And for comparison with other households, horizontal bars or lines ranging from lowest to highest consumption are the design of choice, with the various levels of household consumption indicated as data points on the line.

Studies on billing report over and over again that improved bills, be it with respect to frequency, graphic design, inclusion of comparisons, or additional advice, lead to markedly rising customer satisfaction.

Some preferences, however, vary highly between nations and, probably, cultures. One instructive example is a comparison between Egan (1999) and Wilhite et al. (1999) which have tested the same four graphic designs for presenting a between-household comparison in Delaware, USA, and in Norway. The design that ranked highest in the USA was a distribution graph with the horizontal axis spanning from lowest to highest consumption, and the vertical axis showing the number of households on each level of consumption. The individual data points were represented by little houses (see Fig. 3). The same design bombed completely in Norway, being characterized as childish on the one hand, and difficult to interpret on the other, because it remained unclear whether the houses represented individual households or aggregate data.

For the UK (IEA 2005, p.10) and for Sweden (Sernhed et al. 2003), it is reported that citizens exhibit an interest in comparison with their own previous consumption, but are much less interested in comparisons with other households. On the contrary, the Finnish customers in the study conducted by

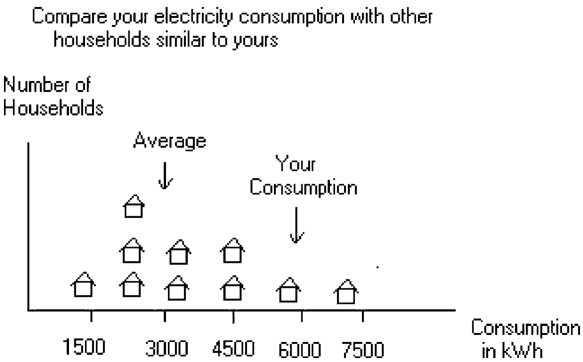

Fig. 3 Distribution graph with little houses, as tested in Wilhite et al. (1999) and Egan (1999)

Haakana et al. (1997) desired normative comparisons, and the Japanese respondents in Ueno et al. (2005, p.1,293) were much more interested in comparisons with others than with own previous consumption.

Research gaps

From the current state of affairs, a number of gaps can be identified that should be explored for useful consumer feedback to be implemented widely. First, many studies and projects use rather small samples. There is a lack of well-documented large-N studies which could provide reliable data on which kind of feedback will stimulate electricity conservation the most. Such studies should cover a representative sample of households, and vary systematically the kind of feedback given; ideally, only one feature of feedback at a time. Actual consumption should be measured during and some time after the feedback phase, and data should be provided on average savings within the experimental groups, on the range of savings that occurred and on differences between different target groups (e.g., “high” and “low” users). Consumption data should be complemented with survey data on motivation, preferences regarding the feedback, and types of action taken.

Another research gap is the lack of international comparative studies. As this short review already shows, there may be wide cultural and national differences not only in preferences, but also in the kind of information that is effective in stimulating conservation. As long as comparative studies are not available, one must be careful about applying results from other countries to a specific national situation.

Furthermore, specific information on some countries is completely missing. Especially for EU accession countries and for Southern Europe, the effects and preferred types of feedback still remain to be investigated.

Conclusions: chances and challenges for implementing consumer feedback

We have seen that though many details remain to be resolved, a relatively sound body of evidence indicates the usefulness of feedback for promoting electricity conservation in households. With all due care because of data restraints, there are reasons to identify some likely features for successful feedback (meaning, both effective in stimulating conservation and satisfying to households).

Such feedback

- is based on actual consumption
- is given frequently (ideally, daily or more)
- involves interaction and choice for households
- involves appliance-specific breakdown
- is given over a longer period
- may involve historical or normative comparisons (although these are appreciated by households, the effects are less clear)
- is presented in an understandable and appealing way (designs should be based on sound consumer research, as has been done in Wilhite et al. 1999 and Egan 1999 and recommended by Roberts and Baker 2003).

These findings go well with our hypotheses that successful feedback has to capture the consumer’s attention, to draw a close link between specific actions and their effects and to activate various motives that may appeal to different consumer groups. Interesting results are that interaction and choice seem to be an important motivating factor, and that long-term feedback is helpful for forming habits, which is also consistent with the theory. The aspects of actual consumption, frequency, interaction, and appliance-specific breakdown suggest that electronic “smart” metering, electronic data procession and communication is an especially useful tool.

However, it is important to check whether the recommendations hold for all target groups. There is probably not “the” perfect feedback for everybody. As we have seen, much-consuming customers react differently from little-consuming ones, and middle class groups from working-class groups. Similar considerations hold for computerized and interactive feedback: an overly complex tool requiring much understanding and initiative from users may not be the tool of choice for households with lower education, lower technical interest (e.g., many elderly people) or less spare time.

Sadly, implementation of useful feedback is lagging way behind knowledge. Implementation usually is not governed by scientific findings but by political interest, power constellations, opportunities, and incentives. Firstly, many variants of improved feedback hinge on technical preconditions that are not always met. For example, continuous electronic feedback requires "smart", two-way metering technology. A similar argument applies to more frequent (e.g., monthly) feedback, if meter reading should not become overly expensive (however, there could be ways out of the dilemma, like self-reading of the meter). Appliance-specific breakdowns would need even more sophisticated technology which is at the moment unlikely to be installed widely. Comparisons with similar households rely on adequate data bases which need to be built up.

Other forms of feedback, however, are less demanding. Comparisons to a previous period, presented in a graphic form, for example, should be feasible as well as the inclusion of environmental impact information or energy-saving tips. In some countries, advanced metering technologies are currently being introduced, providing a better basis for improved feedback (e.g., in Denmark, 25% of all meters will be replaced by remote metering and two-way communication technology by 2010. Norway is conducting pilots with smart meters. Italy has decided to implement them widely). In general, it would be advisable to rely on a little less effective form of feedback that can be more easily implemented. This points, for example, to the potential of improved electricity bills.

The biggest hurdle, of course, is energy utilities' motivation. In situations of overcapacities, cheap electricity available on the market, or oligopolies with little competition, there is little interest in demand-side management. And if conservation is not very important to customers, feedback is not the tool of choice for customer retention.

Here, EU legislation will provide a window of opportunity. Directive (2003/54/EC; concerning common rules for the internal market in electricity) obliges suppliers to disclose certain product features (fuel mix, carbon content, nuclear waste) in the bill. Therefore, utilities need to reconsider their bill format anyway. Even more important, Directive (2006/32/EC) on energy end-use efficiency and energy services (Energy Services Directive) requires Member States to introduce informative billing and other types of feedback, including more frequent billing, historic and normative comparisons, and contact details for obtaining further information on energy efficiency (Art.13). Several Member States have already started acting on the Directive. Denmark has a legal obligation to provide an "informative electricity bill" showing environmental impact as well as historic and normative comparisons. Companies are free to include further information and to choose the mode of presentation (IEA 2005, p.15). In Sweden, legislation foresees that by July 2009, all consumers will have monthly reading of their consumption based on actual use.

The Directive will also provide a favorable framework for systematic comparative large-N studies of various forms of feedback (which are expensive and technically challenging, so normally, there would be few actors with an interest in conducting them). For example, in the UK, the government has earmarked 9.75 million for a pilot study containing various trials of different sorts of feedback. A tender has been held by OFGEM and first trials have started in 2007 (OFGEM 2006).

However, the Energy Services Directive leaves ample space for Member States to define which measures they deem "appropriate", and how stringently they will implement the measures. Therefore, it is up to national actors to push for changes, promote interest in sustainable energy consumption, and introduce experiments with feedback. National energy agencies could be such actors. Where they are lacking, weak or disinterested, NGOs, research institutions, consumer advocacy groups, and innovative utilities could take up the same role. Without them doing so, widespread implementation of helpful feedback will probably not stand any chance.

Acknowledgements The research has been conducted within the project "TIPS—Transformation and Innovation in Power

Systems." I thank the German Ministry for Education and Research (BMBF) for funding TIPS within their program "Socio-Ecological Research". Furthermore, I thank Ms. Anita Eide, two anonymous reviewers of the ECEEE 2007 Summer Study, and two anonymous reviewers of Energy Efficiency for their detailed and important comments on earlier versions of this paper.

References

General references

Birzle-Harder, B., & Götz, K. (2001). *Grüner Strom—eine sozialwissenschaftliche Marktanalyse. [Green Power—a sociological market analysis]*. Frankfurt/M: Institut für sozial-ökologische Forschung.

Bittle, R. G., Valesano, R. M., & Thaler, G. M. (1979–1980). The effects of daily feedback on residential electricity usage as a function of usage level and type of feedback information. *Journal of Environmental Systems*, *9*, 275–287.

Directive 2003/54/EC of the European Parliament and the Council of 26 June 2003, concerning common rules for the internal market in electricity and repealing Directive 96/92/EC. http://europa.eu.int/eur-lex/pri/en/oj/dat/2003/l_176/l_17620030715en00370055.pdf.

Directive 2006/32/EC of the European Parliament and of the Council of 5 April 2006 on energy end-use efficiency and energy services and repealing Council Directive 93/76/EEC. http://eur-lex.europa.eu/LexUriServ/site/en/oj/2006/l_114/l_11420060427en00640085.pdf.

DIW (Deutsches Institut für Wirtschaftsforschung) Berlin, Forschungszentrum Jülich, Fraunhofer ISI, Öko-Institut (2005). Klimaschutz in Deutschland bis 2030. Endbericht zum Forschungsvorhaben Politikszenarien III. [Climate Protection in Germany up to the Year 2030. Final Report for Research Project "Policy Scenarios III".] Berlin: Umweltbundesamt. http://www.umweltdaten.de/publikationen/fpdf-l/2822.pdf.

DLR (Deutsches Institut für Luft und Raumfahrt), IFEU (Institut für Energie und Umwelt), WI (Wuppertal Institut für Klima, Umwelt, Energie) (2004). Ökologisch optimierter Ausbau der Nutzung erneuerbarer Energien in Deutschland. [Environmentally Optimized Expansion of Renewable Energy Use in Germany.] Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit. http://www.dlr.de/tt/institut/abteilungen/system/publications/Oekologisch_optimierter_Ausbau_Langfassung.pdf.

Enquête-Kommission (2002). Bericht der Enquête-Kommission des 14. Bundestags "Nachhaltige Energieversorgung unter den Bedingungen der Globalisierung und Liberalisierung". [Report of the 14th Bundestag's Commission of Inquiry "Sustainable Energy Supply Under Conditions of Globalization and Liberalization".] Bundestags-Drucksache 14/9400, Berlin. http://www.bundestag.de/parlament/gremien/kommissionen/archiv14/ener/schlussbericht/index.htm.

Matthies, E. (2005). Wie können PsychologInnen ihr Wissen besser an die PraktikerIn bringen? Vorschlag eines neuen, integrativen Einflussschemas umweltgerechten Alltagshandelns. [How can psychologists improve their outreach towards practitioners? A suggestion for a new, integrative model of environmentally sound everyday practice]. *Umweltpsychologie*, *9*(1), 62–81.

OFGEM (2006). Energy demand reduction pilot—invitation to bid. http://www.ofgem.gov.uk/temp/ofgem/cache/cmsattach/16285_155_06.pdf?wtfrom=/ofgem/work/index.jsp&section=/areasofwork/metering.

Reviewed literature

Abrahamse, W., Wokje, A., Steg, L., Vlek, C., & Rothengatter, T. (2005). A review of intervention studies aimed at household energy conservation. *Journal of Environmental Psychology*, *25*(3), 273–291.

Arvola, A., Uutela, A., & Anttila, U. (1993). *Billing feedback as a means to encourage household electricity conservation: A field experiment in Helsinki. In Proceedings of the 1993 summer study of the European Council for an energy efficient economy* pp. 11–21. Stockholm: ECEEE Retrieved Feb 28th, 2008, from http://www.eceee.org/conference_proceedings/eceee/1993/Panel_3/p3_2/Paper/.

Brandon, G., & Lewis, A. (1999). Reducing household energy consumption: a qualitative and quantitative field study. *Journal of Environmental Psychology*, *19*, 75–85.

Darby, S. (2001). Making it obvious: Designing feedback into energy consumption. In P. Bertoldi, A. Ricci, & A. de Almeida (Eds.), *Energy efficiency in household appliances and lighting* (pp. 685–696). Berlin: Springer.

Darby, S. (2006). The effectiveness of feedback on energy consumption. A review for DEFRA of the literature on metering, billing, and direct displays. Retrieved March 22nd, 2007, from: http://www.defra.gov.uk/environment/energy/research/pdf/energyconsump-feedback.pdf.

Dobson, J. K., & Griffin, J. D. A. (1992). Conservation effect of immediate electricity cost feedback on residential consumption behavior. In Proceedings of the ACEEE 1992 Study on Energy Efficiency in Buildings 10 (pp.33–35).

Dünnhoff, E., & Duscha, M. (2008). Effiziente Beratungsbausteine zur Minderung des Stromverbrauchs in privaten Haushalten. Endbericht [Efficient building blocks for energy counseling aimed at reducing electricity consumption in private households]. Study by the ifeu (Institut für Energie-und Umweltforschung), Heidelberg (in press).

Egan, C. (1999). Graphical displays and comparative energy information: What do people understand and prefer? Paper presented at the Summer Study of the European Council for an Energy Efficient Economy, 1999, paper no. 2–12. Retrieved December 3rd, 2007, from: http://www.eceee.org/conference_proceedings/eceee/1999/Panel_2/p2_12/.

Garay, J., & Lindholm, P. (1995). Statistics on the energy bill. Better control for the customer. In Proceedings of the seventh international energy program evaluation conference: Energy program evaluation: Uses, methods, and results, Aug 22–25, 1995, Chicago (pp. 499–504).

Haakana, M., Sillanpää, L., & Talsi, M. (1997). The effect of feedback and focused advice on household energy consumption. Paper presented at the Summer Study of the European Council for an Energy Efficient Economy, 1997. Retrieved March 22nd, 2007, from: http://proceedings.

eceee.org/library_links/proceedings/1997/pdf97/97p4–38.pdf.

Henryson, J., Håkansson, T., & Pyrko, J. (2000). Energy efficiency in buildings through information—Swedish perspective. *Energy Policy*, *28*, 169–180.

IEA (2005). International Energy Agency demand-side management programme, Task XI: Time of use pricing and energy use for demand management delivery, Subtask 1 Report: Smaller customer energy saving by end use monitoring and feedback. Retrieved Dec 3rd, 2007, from: http://dsm.iea.org/Files/Tasks/Task%20XI%20-%20Time%20of%20Use%20Pricing%20and%20Energy%20Use%20for%20Demand%20Management%20Delivery/Reports/Subtask1Report12May05.pdf.

Jensen, O. M. (2003). *Visualisation turns down energy demand. In Proceedings of the 2003 summer study of the European Council for an energy efficient economy* (pp. 451–454). Stockholm: ECEEE.

Karbo, P., & Larsen, T. F. (2005). *Use of online measurement data for electricity savings in Denmark. In Proceedings of the 2005 summer study of the European Council for an energy efficient economy* (pp. 161–164). Stockholm: ECEEE.

Kempton, W., & Layne, L. L. (1994). The consumer's energy analysis environment. *Energy Policy*, *22*(10), 857–866.

Mack, B., & Hallmann, S. (2004). Strom sparen in Lummerlund—eine Interventionsstudie in einer Passiv- und Niedrigenergiehaussiedlung. [Conserving electricity in Lummerlund. An intervention study in a passive and low energy house residential area]. *Umweltpsychologie*, *8*(1), 12–29.

Mansouri, I., & Newborough, M. (1999). Dynamics of energy use in UK households: End-use monitoring of electric cookers. Paper presented at the Summer Study of the European Council for an Energy Efficient Economy, 1999, paper no.3–8. Retrieved March 22nd, 2007, from: http://proceedings.eceee.org/library_links/proceedings/1999/pdf99/Panel3/3-08.pdf.

McCalley, L. T., & Midden, C. J. H. (2002). Energy conservation through product-integrated feedback: The roles of goal-setting and social orientation. *Journal of Economic Psychology*, *23*, 589–603.

Mosler, H.-J., & Gutscher, H. (2004). Die Förderung von Energiesparverhalten durch Kombination von instruierter Selbstverbreitung mit Interventionsinstrumenten. [Promoting Energy Conserving Behaviour by Combining Instructed Self-Diffusion with Intervention Instruments]. *Umweltpsychologie*, *8*(1), 50–65.

Nielsen, L. (1993). How to get the birds in the bush into your hand. Results from a Danish research project on electricity savings. *Energy Policy*, *21*, 1133–1144.

NUTEK (1996). *Hushållens krav påelräkningen och annan energiinformation*. Stockholm: NUTEK Report R1996:7 (in Swedish).

Roberts, S., & Baker, W. (2003). Towards effective energy information. Improving consumer feedback on energy consumption. A report to OFGEM. Retrieved March 24th, 2007, from: http://www.cse.org.uk/pdf/pub1014.pdf.

Sernhed, K., Pyrko, J., & Abaravicius, J. (2003). *Bill me this way!—customer preferences regarding electricity bills in Sweden. In Proceedings of the 2003 summer study of the European Council for an energy efficient economy* (pp. 1147–1150). Stockholm: ECEEE.

Sexton, R. J., Brown Johnson, N., & Konakayama, A. (1987). Consumer response to continuous-display electricity-use monitors in a time-of-use pricing experiment. *Journal of Consumer Research*, *14*, 55–62.

Ueno, T., Inada, R., Saeki, O., & Tsuji, K. (2005). *Effectiveness of displaying energy consumption data in residential houses. Analysis on how the residents respond. In Proceedings of the 2005 summer study of the European Council for an energy efficient economy* (pp. 1289–1299). Stockholm: ECEEE.

Ueno, T., Sano, F., Saeki, O., & Tsuji, K. (2006). Effectiveness of an energy-consumption information system on energy savings in residential houses based on monitored data. *Applied Energy*, *83*(2), 166–183.

Wilhite, H., Høivik, A., & Olsen, J.-G. (1999). Advances in the use of consumption feedback information in energy billing: The experiences of a Norwegian energy utility. Paper presented at the Summer Study of the European Council for an Energy Efficient Economy, 1999, paper no. 3-2. Retrieved March 24th, 2007, from: http://proceedings.eceee.org/library_links/proceedings/1999/pdf99/Panel3/3-02.pdf.

Wilhite, H., & Ling, R. (1995). Measured energy savings from a more informative energy bill. *Energy and Buildings*, *22*, 145–155.

Wood, G., & Newborough, M. (2003). Dynamic energy-consumption indicators for domestic appliances: Environment, behaviour and design. *Energy and Buildings*, *35*(8), 821–841.