Ecological Economics 148 (2018) 178–210

Contents lists available at ScienceDirect

Ecological Economics

journal homepage: www.elsevier.com/locate/ecolecon

Behavioral Economics and Energy Conservation – A Systematic Review of Non-price Interventions and Their Causal Effects

Mark A. Andor*, Katja M. Fels

RWI, Hohenzollernstraße 1-3, 45128 Essen, Germany

ARTICLE INFO

JEL Classification:
D10
D12
L94
L95
Q41
Q48
Q58

Keywords:
Systematic review
Behavioral economics
Energy demand
Energy efficiency
Environmental certification
Social norms

ABSTRACT

Research from economics and psychology suggests that behavioral interventions can be a powerful climate policy instrument. This paper provides a systematic review of the existing empirical evidence on non-price interventions targeting energy conservation behavior of private households. Specifically, we analyze four nudge-like interventions referred to as social comparison, commitment devices, goal setting, and labeling in 44 international studies comprising 105 treatments. This paper differs from previous systematic reviews by solely focusing on studies that permit the identification of causal effects. We find that all four interventions have the potential to significantly reduce energy consumption of private households, yet effect sizes vary immensely. We conclude by emphasizing the importance of impact evaluations before rolling out behavioral policy interventions at scale.

1. Introduction

Climate change mitigation programs are on the political agenda worldwide. As a result of ambitious CO2-reduction goals, policymakers are increasingly interested in non-price interventions targeting private household energy consumption. Research from both economics and psychology has shown that behavioral interventions – also referred to as nudges – can be powerful tools in shaping people's behavior in a variety of domains (see, among others, the influential publication by Thaler and Sunstein 2008).[1] Non-price measures are relatively inexpensive to implement and do not interfere with people's choice sets as strongly as, for example, taxes or bans on certain products. Consequently, policy makers are now exploring nudges as a cost-effective approach for reducing energy consumption (Allcott 2015). If proven effective, these interventions could be established as integral and complementary components of climate change policy (Allcott and Mullainathan 2010, Benartzi et al. 2017). This is why researchers are increasingly interested in understanding the effect of non-price measures on residential energy consumption.

This paper presents findings of a systematic review on the effectiveness of behavioral interventions to induce energy conservation. We study the following four interventions: social comparison, commitment devices, goal setting, and labeling. Furthermore, the review focuses on causal in contrast to correlational effects. To this end, we only include those studies that employ an empirical estimation strategy enabling the identification of a *causal* relationship between a policy intervention and consumption behavior. To our knowledge, this is the first study that systematically reviews all published results from behavioral economics and related areas of research that are based on a rigorous evaluation of causal effects.

Our study builds on a few earlier reviews that only focus on a subset of our interventions. Many of these point to potential problems of including effects from correlational studies in their sample, i.e. studies that are not able to disentangle causation from correlation. Abrahamse et al. (2005) evaluate the effectiveness of some interventions aiming to encourage households to reduce energy consumption. They conclude that information has an influence on knowledge, but does not necessarily result in behavioral changes or energy savings. Rewards have effects on energy conservation, but they are rather short-lived. Feedback, in particular when it is given frequently, can also be effective.

* Corresponding author.
E-mail address: andor@rwi-essen.de (M.A. Andor).

[1] The facts that the book "Nudge" (Thaler and Sunstein 2008) has already been cited more than 9000 times (Google Scholar, checked 11/15/2017) as well as that Richard Thaler won the Nobel prize "for his contributions to behavioral economics" in 2017 (Nobelprize.org 2017) can be seen as two indicators of a growing academic interest in behavioral interventions.

https://doi.org/10.1016/j.ecolecon.2018.01.018
Received 20 July 2017; Received in revised form 18 November 2017; Accepted 15 January 2018

More recently, Karlin et al. (2015) conducted a meta-analysis on the effect of feedback on energy usage. They conclude that feedback is effective but with significant variation in effects. Furthermore, Delmas et al. (2013) analyze the effect of information strategies on energy savings and find a substantial reduction effect on average. However, in a similar vein as Abrahamse et al. (2005), they conclude that the effect diminishes with the rigor of the study, indicating potential methodological issues in the considered literature. In particular, none of the existing reviews takes into account whether the considered studies apply a method that has the potential to identify the causal effect of the intervention, which is critical to the question of its policy relevance (Imbens and Wooldridge 2009).

Consequently, our systematic review differs from previous research by solely focusing on studies that have the potential to identify causal effects between the intervention and the outcome. Furthermore, we include articles published up to May 2017 in working paper series as well as peer-reviewed journals to provide the most comprehensive and up-to-date account of research in economics and psychology. This is particularly important because there has been a growing number of high-quality studies in the recent past. Hence, our review comprises several very recent large-scale randomized controlled field experiments. As an additional contribution, our systematic review is the first to account for labeling as a non-price intervention, which has been applied worldwide on a large scale and potentially affects millions of household decisions each year.

The paper proceeds as follows. In the subsequent section, we define and motivate the four considered interventions. Section 3 explains the methodology of the systematic review. In Section 4, we synthesize and discuss the results. Section 5 concludes with recommendations for researchers and policy makers.

2. Behavioral Interventions and Energy Conservation

A considerable percentage of annual emissions in industrial countries is induced by residential energy consumption. In addition, private households are a prime target for behavioral interventions (Karlin et al. 2015). Households may conserve energy in two ways: First, they can change their consumption of energy services, for example by reducing lighting use. Second, they can modify their purchasing behavior and invest in energy efficiency, for example by buying a highly efficient washing machine.[2] Both, the purchase decision and the consumption behavior, can be targeted by policy interventions. Non-price interventions are usually justified with so called internalities, i.e. externalities that the agent imposes on herself by making suboptimal choices, measured by her own experienced utility (Chetty 2015).

According to Allcott (2016), six main internalities are responsible for consumer mistakes in the domain of energy conservation: present bias, bias toward concentration, biased beliefs, costly information acquisition, exogenous inattention, and endogenous inattention. Our study selected those non-price interventions that are most common and suitable to address each of these internalities (see Table 1).[3] The resulting four non-price interventions and the internalities they address are explained in more detail below.

2.1. Social Comparison

Social comparison refers to the process of giving households information about their energy consumption in relation to the consumption of comparable households. Such a comparison is closely connected to also receiving feedback about one's own behavior. The chosen reference group should be relevant for the treated household (Abrahamse et al. 2005) and can be, for instance, consumers of the same energy provider or households within the same postcode-level. Moreover, the choice of the reference level is important: the household's consumption can either be compared to the average consumption level of the reference group or to a more ambitious group, e.g. the most efficient 10%.

Social comparison addresses *biased beliefs* about one's own consumption behavior in comparison to others. For example, a person might consider herself an environmentally friendly energy consumer and underestimate her actual consumption level when compared to other consumers. This biased belief can be corrected by a social comparison.

The potential energy conservation effect of a social comparison might be triggered by three phenomena. First, many people exhibit reference dependent preferences (Kahneman 2003). Accordingly, social norms can constitute a reference point. Complying with these norms increases most individuals' utility whereas deviating from it typically leads to disutility caused by social disapproval (Schubert and Stadelmann 2015). Second, in situations of uncertainty, individuals may use other peoples' behavior as orientation by implicitly assuming that those others have more information about the socially desired behavior (Allcott and Mullainathan 2010, see also Delmas et al. 2013). Consequently, people tend to adjust their actions according to the prevalent group behavior. Third, social comparisons evoke feelings of competition (Abrahamse et al. 2005). This is especially important when the household's consumption level lies above the average or above some threshold that the household perceives as desirable (for example, belonging to the most efficient 10% of costumers).

Table 1
Internalities and chosen interventions.

Internality	Chosen intervention(s)
Present bias	Commitment devices and goal setting
Bias toward concentration	Labeling
Biased beliefs	Social comparison and labeling
Costly information acquisition	Labeling
Exogenous and endogenous inattention	Labeling

2.2. Commitment Devices and Goal Setting

Commitment devices are “a set of interventions that allow individuals to lock themselves today into the action that they want to take tomorrow” (Allcott and Mullainathan 2010, p. 2). Examples of commitment devices are oral or written pledges or promises to conserve energy (Abrahamse et al. 2005). The commitment can either be a promise to oneself, or alternatively be made public. Goal setting combines commitment with a concrete reference point. Instead of pledging to conserve energy, a household specifically promises, for instance, “to reduce energy consumption by 10 percent within the next year”. Not only setting a reduction level but also a deadline for achieving this goal facilitates an evaluation of success or failure. This increases pressure but also motivation by making satisfaction conditional on a desired level of performance (van Houwelingen and van Raaij 1989). A goal can be chosen by the household itself (being a form of commitment device) or be externally set (for example by institutions).

The idea behind voluntarily binding one's own future behavior is that some people are aware that they sometimes have time-inconsistent preferences (O'Donoghue and Rabin 1999). For instance, as O'Donoghue and Rabin (2008) point out, many people procrastinate, sometimes to the extent that the desired action is never taken. This distortion of one's own preferences is induced by *present bias*, the

[2] The purchase of an energy efficient appliance will ultimately result in reduced energy consumption when expected energy savings are not completely offset by an increase in the use of the appliance, which is known as the rebound effect (see, for instance, Frondel and Vance, 2013).

[3] A further important intervention in this regard is feedback. Yet, because the comprehensive study of Karlin et al. (2015) provides a recent account of the existing research on the intervention, we do not consider this intervention in our review.

tendency to discount more strongly between today and tomorrow than between two future dates (Laibson 1997).

A commitment device helps individuals to overcome such time-inconsistent preferences by providing a source of motivation: it compares the present situation with a desired future state (van Houwelingen and van Raaij, 1989). When the commitment is a pledge to oneself, it appeals to a personal norm (the individual wants to satisfy expectations toward itself). A public commitment creates expectations by others, thereby leading to social pressure (Abrahamse et al. 2005). In addition to that, a specific goal targets reference-dependent preferences: individuals aim at a pre-determined level and judge their performance according to this reference point. Achievement of the goal provides a feeling of accomplishment, whereas failure creates disutility even if the goal level was externally set (Bandura 1986). This is mainly due to loss aversion (losses loom larger than gains), one of the key concepts of prospect theory (Kahneman and Tversky, 2013; for an overview see Reisch and Zhao 2017).

2.3. Labeling

A label is a tag that summarizes information on a good in an easily accessible way. This can be achieved by presenting a selection of information about the product's attributes on the label or by visualizing the most relevant information in a graphical manner. In the domain of energy consumption, labels can, for example, comprise information on energy usage levels of appliances or energy efficiency standards of houses.

In principle, there are two different sorts of labels. Either the labeling program is voluntary and an appliance is awarded the label for satisfying certain criteria like the US Energy Star Label, or a label is mandatory for all appliances, as with the EU Energy Label. In the latter case, goods can be ranked according to their performance along the criteria specified by the label, for instance the EU energy efficiency classes.

Labeling addresses several of the internalities that are relevant in the domain of energy conservation. First, consumers potentially respond to labels by ascribing more attention to certain features of the good and by this overcome *inattention*. As a label renders selected criteria (more) salient, it aims at the availability heuristic, i.e. a simplifying rule that gives highly accessible features a stronger influence on decisions while information of low accessibility will largely be ignored (Kahneman 2003). By this, a label also helps to reduce *costly information acquisition*. However, the mechanism of the availability heuristics remains true regardless of whether the costs of gathering relevant information, for example on average usage levels or life expectancy of appliances, is high or low (Schubert and Stadelmann 2015). Moreover, labels target *biased beliefs* about the value of different energy conservation actions. If, for example, a consumer underestimates the savings' potential of an appliance, she might underpurchase energy efficiency based on her own preferences (Allcott 2016). A label correcting her beliefs by pointing out the potential savings in energy costs addresses this internality and might trigger more energy efficient purchases. The same is true for the *bias toward concentration* that describes the phenomenon that consumers might downweight energy cost savings that accrue in small amounts compared to a bigger sum like the sales prices for an energy efficient appliance, which has to be paid immediately (cf. Kőszegi and Szeidl 2012).

3. Methodology

Following the guidelines for systematic reviews suggested by the Campbell Collaboration (2014), we take five successive steps to identify and analyze relevant studies: setting up criteria for including studies in the review, literature search, selection of studies, data extraction, and data analysis.

3.1. Inclusion Criteria (PICOS)

As a first step, we developed a set of criteria along the so-called PICOS – an acronym for participants, interventions, comparisons, outcomes, and study designs (Campbell Collaboration 2014). The PICOS guided the selection of studies in the further process. We include studies targeting private households or individuals living in an industrialized or emerging country where the living situation (especially with regard to energy consumption) is comparable to an industrialized country. Studies focusing on energy consumption in official buildings or enterprise settings are excluded.

Our focus is on the four interventions social comparison, commitment devices, goal setting, and labeling. We also consider combinations of these interventions with each other or with additional non-price measures (e.g. feedback, energy saving tips). In contrast, all studies combining these interventions with financial incentives (for instance, dynamic pricing) are excluded from the review, as long as they do not allow identifying the effect of the non-price intervention separately.

Regarding the outcome, we include all studies that report an individual's or household's actual or self-assessed usage level of energy, gas, or water. For labeling, we also consider secondary outcomes like the perception of and the willingness to pay (WTP) for energy efficiency. These secondary outcomes influence usage levels via different channels (as discussed in Section 2).

We include all studies that permit the identification of the causal relationship between the intervention and the outcome of interest. Hence, studies are considered if their methodology is based on a causal inference design like randomized controlled trials (RCT), matching, difference-in-differences, instrumental variable estimation, and regression discontinuity design, or if they control for self-selection with alternative methodologies.[4] Studies that do not employ methods suitable for identifying the causal effect of an intervention are excluded from the review. Overall, we apply no time restriction: all study results published in a journal or as working paper up until May 2017 are considered in the review.

3.2. Literature Search

We used two main avenues to identify relevant studies: first, a keyword search in databases, and second, a backward search on relevant review studies. Before the database search, we pre-determined a systematic combination of keywords for each of the four interventions (see Appendix B1). These keywords were employed on two databases: EconLit, consisting of more than 785,000 articles from peer-refereed journals and acknowledged working paper series in economics, and ScienceDirect, accessing more than 13 million articles from journals in different disciplines, from which our study chose the disciplines “Economics, Econometrics, Finance”, “Psychology”, “Social Sciences”, “Environment”, and “Energy”. This returned a total of 1,121 results.

Additionally, we conducted backward searches in the literature of the following four review studies: Abrahamse et al. (2005), Delmas et al. (2013), Karlin et al. (2015) and Lokhorst et al. (2013). Of the 147 articles identified, we included 71 studies after ruling out duplicates. This garnered a total of 1,192 studies for the screening process of this review (see Fig. 1).

3.3. Selection of Studies

All studies were screened by independently reading the title and abstract and assessing whether a study satisfied all inclusion criteria set up in the PICOS via an inclusion decision form (see Appendix B2). We excluded 1,040 studies due to this procedure, leaving 152 articles for

[4] For an excellent overview of causal inference designs see Angrist and Pischke (2009) or Imbens and Wooldrigde (2009).

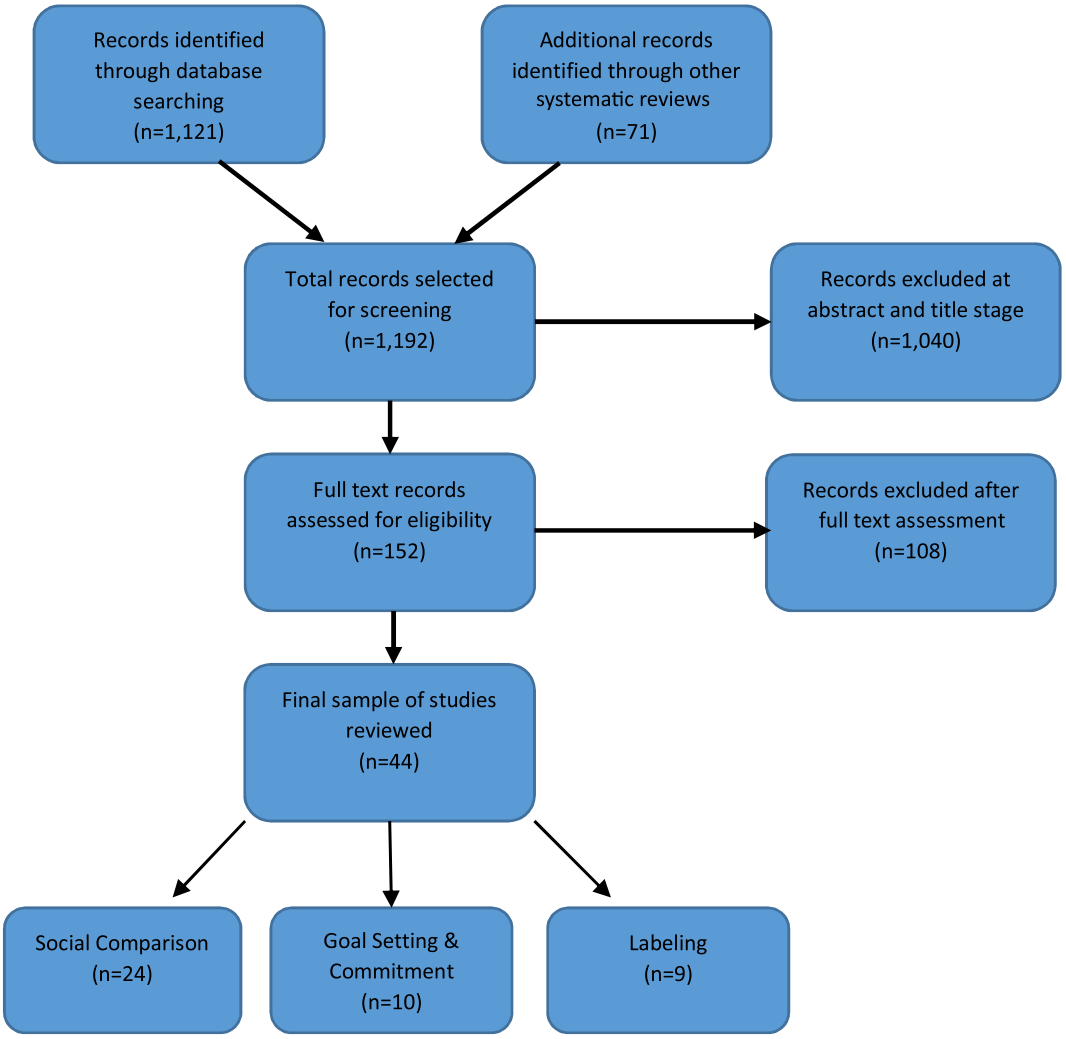

Fig. 1. Overview of results in the searching, screening and selection process.
Note: The articles by Ferraro et al. (2011) and Ferraro and Price (2013) document short-term and long-term effects of the same study. They are thus summarized in the final results table and depicted as one study in the last stage of this figure. The study by Kurz et al. (2005) investigates a social comparison treatment as well as labeling. Here it is categorized as a social comparison study. The study by Jaeger and Schultz (2017) evaluates treatments that include social comparison and commitment. As they do not provide a statistical analysis for the social comparison treatment effects, we list them as a commitment study.

full text screening (see Fig. 1), which again was conducted independently. In cases of differing assessments, we consulted a third (independent) scientist and solved discrepancies by discussion. A common reason for exclusion was that studies discussed strategies for energy conservation behavior theoretically or approached it with a qualitative design without actually measuring the effect of an intervention. When screening the selected full text articles, the main reason for exclusion was that studies neither applied a causal inference design nor controlled for self-selection. In contrast, at the beginning of the screening process a majority of papers was completely off-topic. Finally, 44 articles satisfied the criteria of this systematic review.

3.4. Data Extraction

We developed a detailed coding sheet (see Appendix B3) based on the guidelines of the Campbell Collaboration (2014). Two reviewers coded each study independently of each other according to the same criteria. In cases in which the extracted information was ambiguous (for instance regarding the method of causal inference design), we consulted a third (independent) scientist and discrepancies were solved by discussion. The results from the data extraction are summarized in a results table (see Appendix A1).

4. Results

4.1. Brief Characterization of Included Studies

4.1.1. Research Interest

Among the considered interventions, social comparison has attracted the most research attention. While 24[5] independent studies tested at least one social comparison treatment, only ten evaluate the effects of goal setting and commitment. Additionally, nine studies assess the impact of labeling. In terms of available treatment effects, the analysis of our systematic review can rely on 47 documented effects for social comparison, 30 for goal setting and commitment, and 28 for labeling. The difference in numbers between articles and treatment effects is due to the fact that many studies consider more than one relevant intervention or evaluate multiple versions of the intervention in question.

Regarding the studies' publication dates an interesting pattern emerges. While the peak of research interest in goal setting and commitment was in the 1980s, followed by a longer neglect and a recent

[5] As Jaeger and Schultz (2017) do not report any statistical analysis for their social comparison interventions, we do not include them in the social comparison analysis.

rediscovery,[6] labeling has been largely neglected up till 2011. Since then, we observe a growing academic interest in analyzing the impact of the intervention. Similarly, research interest in the causal effects of social comparison on energy consumption behavior is fairly recent: the overwhelming majority of studies was published after the year 2004, with 80% of studies in this sub-sample (19 articles) being more recent than 2010.

4.1.2. Methods

For social comparison and goal setting/commitment, one dominant method is applied, namely the evaluation via randomized controlled trial. This methodology not only produces a high internal validity but also has advantages in terms of external validity as it (in most studies) observes real-life behavior. In contrast, most of the studies on labeling are conducted as laboratory experiments or choice-experiments within online-surveys, which have low external validity. Specifically, researchers ask subjects, for instance, which appliance they would choose in a set of choices and attribute differences in choice patterns to the way the choice set is presented (which differs between control and treatment group). Obviously, such choices are not real purchases but rather decisions within a hypothetical and limited choice set. It is thus not certain whether the reported effects would also occur in a real world setting, and results should be treated with caution despite their high internal validity. However, a few labeling studies also exploit a policy change as a natural experiment or conduct an RCT with a labeling intervention.

4.1.3. Geographical Location

Most of the studies identified by the systematic review are located in the US. This is especially true for goal setting/commitment, which up to date only has been evaluated in the US and the Netherlands. For social comparison and labeling, studies are conducted in several countries, for instance Japan, Australia, Finland and Great Britain. The majority of published research (social comparison) or at least a great share (labeling), however, also targets a study population in the US.

4.1.4. Combination of Interventions

For social comparison and goal setting/commitment, almost all studies combine the considered interventions with other treatments that we do not study. This poses a challenge for identifying the actual effect of the intervention. It moreover makes it difficult to compare effects from different studies, as they do not evaluate the same treatment. Fig. 2 shows that energy saving tips and a comparison of usage history are most popular among treatment combinations. Within the sub-sample of studies that allow identification of the pure effect by not combining the intervention with additional treatments, other limitations apply. For instance, except for the study by Dolan and Metcalfe (2013), the relevant social comparison studies suffer from methodological shortcomings (e.g. no reported effect sizes). For goal setting and commitment devices, there are substantial differences regarding the intensity of the treatment, because the pre-determined level of the goal varies as well as the criterion whether the level was self-selected or externally set. All in all this prohibits arriving at a coherent picture regarding the existing empirical evidence on the pure effects of the considered interventions.

4.1.5. Outcomes

The vast majority of studies on social comparison and goal setting/commitment focuses on actual or self-assessed energy consumption as their outcome of interest. For labeling, the case is different. Most studies evaluate the effect of labels on the WTP for energy efficient appliances, followed by studies on the estimation of energy saving potential or hypothetical purchase decisions regarding energy efficient appliances (see Fig. 3). Only one labeling study looks at actual energy consumption. Consequently, the quality of empirical evidence differs between the interventions: while the majority of labeling studies are based on stated preferences approaches (with their well-known limitation of hypothetical nature compared to revealed preferences), for the other interventions most studies analyze real behavior.

4.2. Synthesis of the Evidence

4.2.1. Social Comparison

Overall, the empirical evidence suggests that social comparison triggers energy conservation. All but two of the 24 studies present statistically significant results in at least one of their analyzed treatment groups.[7] A social comparison intervention results in a reduced energy consumption of private households ranging from 1.2% to 30% compared to the control group (see Fig. 4). The only study showing an increase in energy usage level is Schultz et al. (2007) in one of their treatment groups. Yet, the respective treatment group was deliberately restricted to low users only, demonstrating the so called boomerang effect, i.e. a social comparison can lead to adverse effects for the group that performs well in the comparison. They additionally show that the boomerang effect can be eliminated by adding an injunctive message. With regard to other outcomes than actual energy consumption, Komatsu and Nishio (2015) find a significant (positive) effect of the social comparison treatment on the motivation to conserve energy.

The heterogeneity of effects may be attributed to the medium via which the household receives the social comparison. As a tendency, we observe that online/e-mail and In-Home-Display (IHD) treatments seem to result in a higher effect than letters. However, social comparisons using IHDs are still underresearched. Another noteworthy fact is that studies with higher sample sizes (above 80,000) find smaller effects (around 2%). Both observations are compatible with existing evidence from previous reviews (Karlin et al. 2015).

One specific intervention design, the social comparison based “home energy reports” (HER) by the private company Opower, has attracted major research attention. Opower cooperates with numerous US-energy suppliers and mails the HER to more than ten million households in the United States with the aim of reducing electricity consumption. The HER is a two-page letter with a bar graph comparing the household's energy consumption to its geographically nearest neighbors in similar house sizes on the first page and energy saving tips on the second page. Several high quality studies in this review (Allcott 2011; Allcott and Rogers 2014; Ayres et al. 2013; Costa and Kahn 2013) report significant, yet modest reduction effects. The empirical evidence for this specific intervention is outstanding as the internal and external validity for the US is high, long-term effects are documented and cost-benefit analyses are conducted.[8] In a very recent study, Andor et al. (2017a) show that social comparison based HERs are most likely not a cost effective climate policy instrument in many countries, in particular in Europe, due to lower electricity consumption levels and carbon intensities. For a study population in Germany, they additionally measure a substantially smaller treatment effect of social comparison based HERs. The numerous high quality studies can be seen as best practice for the evaluation of an intervention at large scale.

In all studies, in which a comparison of two similar combinations of

[6] Almost two thirds of studies were published between 1978 and 1989, the remaining share since the year 2002.

[7] The study of Hakaana et al. (1997) is not considered in this regard, as the article does not report any significance levels at all.

[8] Yet, on the basis of 111 RCTs that evaluate the HER, Allcott (2015) shows that even for the exceptional case that there exists many replications, program evaluations can still give systematically biased out-of-sample predictions due to a “site selection bias”. In this example, predictions from the first 10 replications substantially overstate efficacy because, among others, utilities in more environmentalist areas are more likely to adopt the program, their customers are more responsive to the treatment, and utilities initially target their treatment at higher-usage consumer subpopulations.

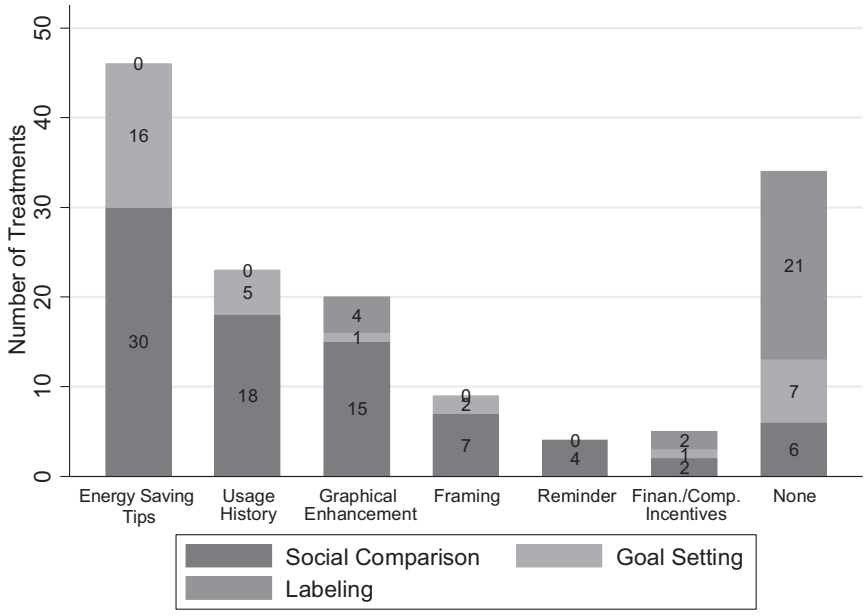

Fig. 2. Additional interventions combined with social comparison, goal setting/commitment, and with labeling treatment.
Note: For Jensen et al. (2016) we refer to model 1.

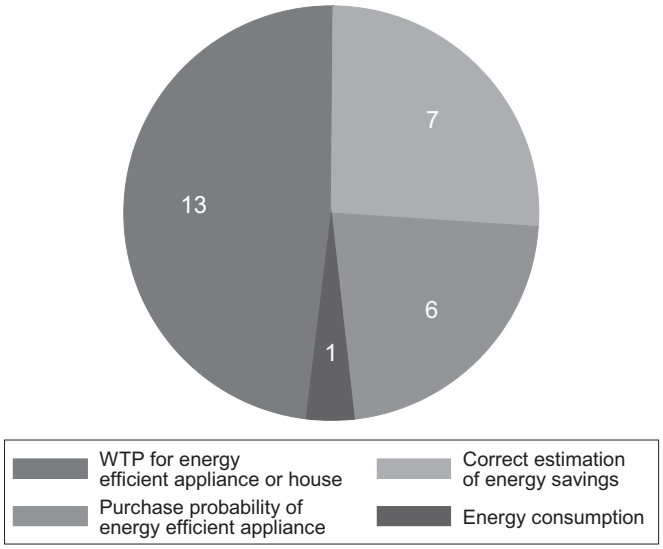

Fig. 3. Analyzed outcomes in studies with a labeling treatment.

interventions is possible, the effect increases when social comparison is added to the treatment (Ferraro and Price 2013; Mizobuchi and Takeuchi 2013; Tiefenbeck et al. 2016). Moreover, Schultz et al. (2015) find a significant reduction in energy consumption of 9% compared to an insignificant treatment effect, when social comparison is added to the initial treatment intervention. In addition, the reference group seems to matter: while Shen et al. (2015) do neither find a significant effect of a social comparison with the entire neighborhood nor with next-door-neighbors, the comparison with street neighbors results in a reduced consumption by 5.4%.

Two out of three studies that aim to identify the pure effect of social comparison have methodological limitations. Haakana et al. (1997) find that a pure social comparison treatment shows similar or smaller effects as a social comparison with energy saving tips. The study, however, provides no information about significance levels for any of the effects. Komatsu and Nishio (2015) analyze the self-assessment of one's own consumption in regard to the neighbors' consumption and the own motivation to conserve energy. They find ambiguous effects of four social comparison treatments, but do not report effect sizes. Moreover, the analysis is not based on actual behavior but solely relies on self-assessments. In contrast, Dolan and Metcalfe (2013) show that biannual feedback with a comparison to the neighbors' consumption triggers energy conservation of around 4%, while the combination of this intervention with tips to save energy increases the effect to almost 11%.

Some studies suggest that social comparison might also result in adverse effects. Schultz et al. (2007) show that low-consumption households significantly increased their energy consumption after having learned that they are below-average users. This raises important questions regarding a tailored application of social comparison treatments. Furthermore, Tiefenbeck et al. (2013) investigate side-effects of a social comparison treatment. They find a significant reduction in water consumption of 6% after they provided weekly feedback about water consumption levels per capita, accompanied by a social comparison with the most efficient 10% of users, and tips about how to conserve water. At the same time, the electricity consumption of the treatment group increased by 5.6% compared to the control group. The authors argue that their "findings are consistent with the concept of moral licensing" (Tiefenbeck et al. 2013, p. 160), the phenomenon that past good deeds favor a positive self-perception which in turn creates licensing effects, leading people to engage in behavior that is less likely to be in line with their moral values (Nisan and Horenczyk 1990). Yet, other phenomena could also explain this, for example, attention shifts from conserving electricity to conserving water (as noted by one reviewer). However, if such side-effects occur, it is not clear whether an intervention induces sustainable behavior in a broader sense, even if the estimated treatment effect suggests a reduction of the "direct" outcome (in this example: water consumption). Therefore, we agree with Tiefenbeck et al. (2013, p. 160) who advocate the "adoption of a more comprehensive view in environmental program […] evaluation in order to quantify and mitigate these unintended effects".

In regard to long-term effects, no clear picture emerges. Several studies document a reinforcement of effects in the long-run for some of their treatment groups (Alberts et al. 2016; Allcott and Rogers 2014; Delmas and Lessem 2014; Dolan and Metcalfe 2013; Schultz et al. 2007). In Staats et al. (2004), two initially insignificant short-term effects increase and become significantly different from the control group when measured two years after the intervention stopped. In other studies, treatment effects decrease (Ferraro et al. 2011; Schultz et al. 2015;

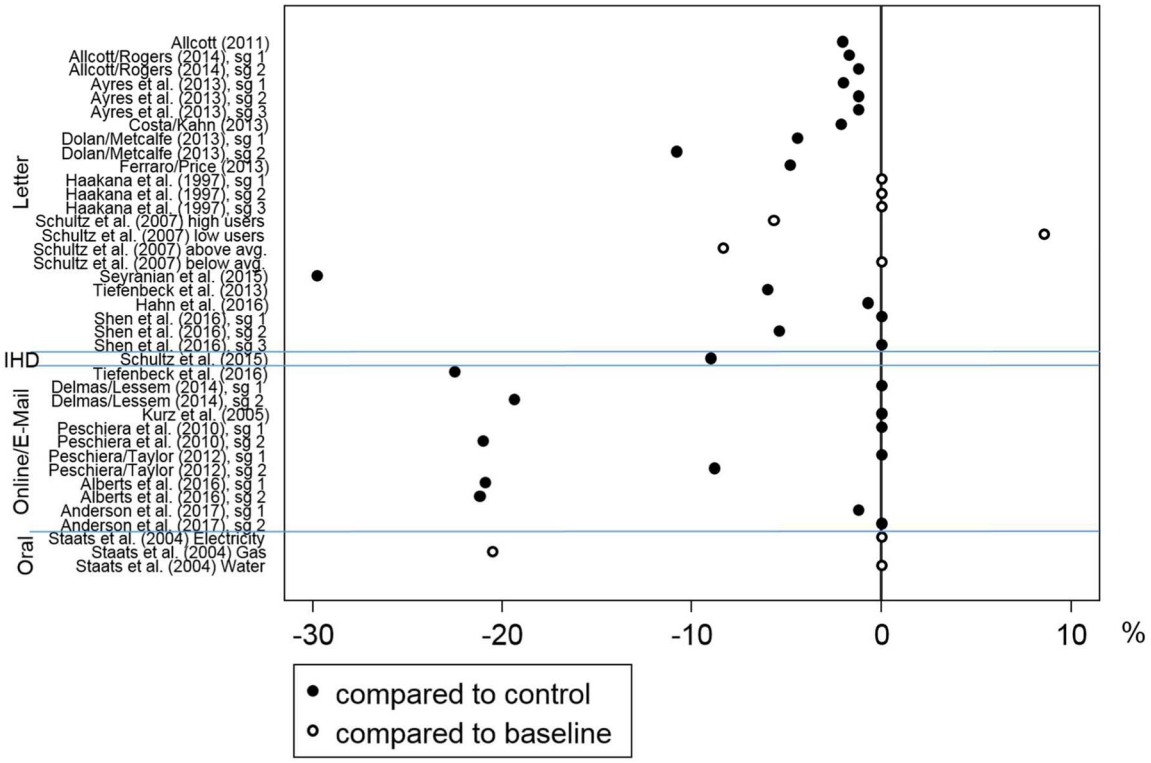

Fig. 4. Estimated effects of social comparison by medium.

Note: Insignificant effects and effects of studies that do not report significance levels are depicted as nil-effects in this graph. Studies are sorted by treatment-medium and in alphabetical order. "sg" and "avg" stand for "study group" and "average", respectively. Comment: The studies of Komatsu and Nishio (2015) and Kantola et al. (1984) are not included in the graph as they do not mention the magnitude of the measured significant conservation effects.

Schultz et al. 2007; Staats et al. 2004), remain or become insignificant (Anderson et al. 2017; Alberts et al. 2016; Delmas and Lessem 2014; Dolan and Metcalfe 2013; Tiefenbeck et al. 2013; Schultz et al. 2007). It would be desirable to identify the driving factors for these heterogeneous effects. Yet, based on the existing empirical evidence, we do not observe any obvious indications.[9] As long-term effects are crucial for the cost-effectiveness of the interventions, their analysis should be one focus of future research.

In sum, social comparison presents an effective treatment. Yet, researchers and policy makers should closely monitor potential adverse effects on certain subgroups of households (below-average users) and on consumption patterns regarding other goods (moral licensing effect). Furthermore, the analysis of long-term effects and cost-benefit analysis should be an inherent part of the impact evaluation.

4.2.2. Goal Setting and Commitment Devices

Many of the identified studies on goal setting and commitment suffer from methodological shortcomings, despite satisfying the criterion of applying a methodology suitable to identify a causal effect. For instance, they test several treatments in their RCT despite relying on very small study samples, or they do not report significance levels of results. Thus, the picture regarding an assessment of treatment effects is not quite clear: exactly one half of the documented effects are not significantly different from zero or cannot be depicted as such because no significance level is reported (see Fig. 5). This may either be because the treatment actually has no effect, or that the study's features (in particular a not sufficiently powered sample) contribute to the result.

Interpreting the results of the identified studies, self-set goals seem to result in a significant reduction effect when they are chosen realistically (Harding and Hsiaw 2014; Jaeger and Schultz 2017; McCalley and Midden 2002; Winett et al. 1979). Harding and Hsiaw (2014) report based on a field study with more than 12,000 households that the sub-sample selecting the achievable goal of energy conservation between 0% and 15% reduced their consumption by 11%. The authors still measure a significant reduction effect after 18 months. At the same time, the treatment group that selected a more optimistic goal (15%–50%) also saved energy shortly after the start of the program. Yet, their effort vanished after they received feedback about their usage development, presumably because they realized that it would be impossible to achieve the ambitious goal. In contrast, both sub-samples choosing either a zero-goal or a more-than-50% goal showed no behavior change at all. As we deem it important for policy advice, we would like to clarify the interpretation of the results: while it is a valid conclusion to state that people who set themselves realistic goals exhibit a conservation effect of 11%, the results of Harding and Hsiaw (2014) do not allow the conclusion that setting people realistic goals will end up in substantial energy savings (of around 11%) because this effect is only valid for the self-selected group. People who are in general more motivated to conserve energy could have selected themselves into more realistic goals, while people who are generally unwilling to change their behavior chose zero. Similarly, Jaeger and Schultz (2017) only document (significant) effects for compliers in their study. However, in order to identify an average conservation effect of the treatment goal setting, it would be necessary to estimate the average treatment effect for the full treatment group.

Goals that are externally set by the experimenters have resulted in either insignificant effects or energy savings up to 22% compared to baseline consumption (Fig. 5). Many of the reported significant effects stem from Winett et al. (1982), who conducted two RCT studies, but

[9] In a recent study, Brandon et al. (2017) explore the underlying mechanisms of long-run reductions in energy consumption induced by a social comparison (specifically the HER by Opower). Using data of 38 natural field experiments, they find that the physical capital channel (e.g. insulating the house or buying a more energy efficient washing machine) is more important for the persistence of effects than habit formation within the household.

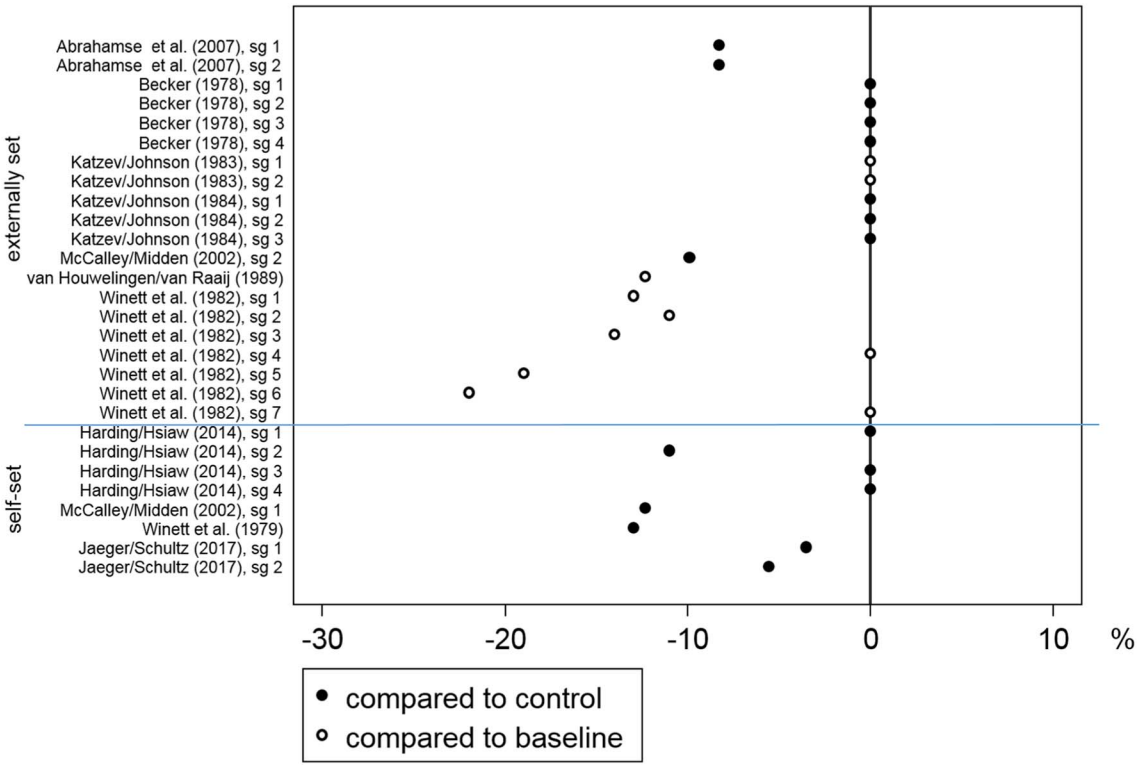

Fig. 5. Goal setting/commitment effects by self-set or externally set goal.
Note: Insignificant effects and effects of studies that do not report significance levels are depicted as nil-effects in this graph. Studies are sorted by externally set and self-set goals/commitments and in alphabetical order. 'sg' stands for 'study group'. Comment: The study by Jaeger and Schultz (2017) solely documents the effect for those individuals that comply with the commitment device and not for the full treatment group.

with a relatively small sample size of in total 83 households for five treatment groups. After having taken part in a meeting, researchers asked participants to sign a commitment to reduce their energy consumption by 15% within the next 35 days. In addition to frequent feedback about their progress, households received an audio tape with tips on how to save energy. One study was conducted during winter, the other one during summer. All but one of six treatment groups showed significant and relatively high reduction effects ranging from 11% to 22%.

In another study, an externally set goal of 10% resulted in a significant energy conservation effect of about 12.3% (van Houwelingen and van Raaj 1989). The effect remained significant one year after the intervention. Abrahamse et al. (2007) measure a smaller, but still significant effect of a 5%-goal in their RCT. In addition to these findings, a laboratory study by McCalley and Midden (2002) points to energy saving potential of externally set goals of around 20%.

On the other hand, Becker (1978) as well as Katzev and Johnson (1984) do not detect significant effects of external goal setting in the field. Yet, in another study, Katzev and Johnson (1983) find that insignificant effects in the short-run become significant in the long-run.

In sum, goal setting and commitment still require substantial research efforts. The significant results point to a relatively high energy conservation potential of around 10%, while at the same time we cannot rule out that many of the insignificant results are due to underpowered samples. It seems to be a promising avenue for further research to evaluate the effect of energy conservation goals with a well-powered study set up, especially goals that are externally set.

4.2.3. Labeling

All but two of the identified studies on labeling report significant results in at least a subsample of their study population. Regarding effect sizes, they also point to potentially pronounced effects of labeling.

Six studies evaluate the effect of labels under real-world conditions. Brounen and Kok (2011) document significantly higher sales prices for houses with a "green" energy efficiency EU-label (classes A, B, C) than for comparable houses without such a label. In the same vein, Jensen et al. (2016) show that the price mark up for houses with efficient EU-labels (A to C) increased from 2.4% to 10.1% compared to houses with inefficient labels (D to G) after the energy performance rating was made mandatory in real estate advertisements in 2010. Houde (2014) exploits two natural experiments concerning the Energy Star Label. He observes a significantly higher WTP for refrigerators when they are awarded the label. His analysis is based on a comparison of the same models of refrigerators before and after the criteria of the Energy Star Label were tightened. In the second study, the measured positive effect of the Energy Star Label on the WTP for refrigerators is not significant. In a similar vein, Allcott and Taubinsky (2015) do not find a significant effect of a label on the probability to purchase an energy-saving lightbulb in their RCT in a big electrical store in the US. In the study by Kurz et al. (2005), which provided the treatment group with labels on the consumption levels of the household's own appliances, effects range from zero to a reduced energy consumption of 23%.

The choice experiment of Heinzle (2012) shows that the WTP for a more efficient TV significantly increases when consumers are given information on absolute operating costs over the course of ten years. Yet, it significantly decreases when the label provides information on annual operating costs. Newell and Siikamäki (2014) find in a further hypothetical choice experiment that insufficient information can lead to considerable undervaluation of energy efficiency. Simple information on the monetary value of energy savings from a more efficient warm water processor was the most important element guiding a cost-efficient energy efficiency investment, with information on physical energy use and carbon dioxide emissions having additional but lesser importance.

The results of Waechter et al. (2015), moreover, indicate that energy efficiency classes induce people to make decisions based on the

energy efficiency class even when concrete and readily apparent consumption information contradicts it. Indeed, in a recent study based on actual efficiency class ratings, Andor et al. (2017b) find that a majority of participants value efficiency classes beyond the economic value of the underlying differences in energy usage. Furthermore, they show that adding annual operating cost information to the EU energy label promotes the choice of energy-efficient durables. Their results also indicate that displaying operating cost affects choices through two distinct channels: it increases the attention to operating cost and reduces the valuation of efficiency class differences. Ölander and Thøgersen (2014) provide evidence for the positive effect of visualizing information in form of a simple label: changing an energy efficiency scale from a complex "A+++-D"-system to a simpler "A – G" more than doubled the probability that an energy-efficient device is chosen.

In sum, even though up to now there is not a vast amount of research on labeling, we identify a remarkable potential of the intervention. Not only do hypothetical purchase decisions in choice experiments confirm their effectiveness, but also evaluations of labels in the field. Future research should focus on different elements of labels, like the framing of costs or the mode of ranking, and test their separate and combined effects in the field.

5. Discussion and Conclusion

This paper conducted a systematic review of behavioral interventions to induce residential energy conservation. In contrast to the existing literature, this review focused on studies that permit the identification of the causal relationship between the intervention and the outcome of interest. In addition, it is the first review to cover labeling, an intervention that affects millions of people worldwide.

We find that all four reviewed interventions – social comparison, commitment devices, goal setting, and labeling – have the potential to significantly reduce energy consumption of private households. While the vast majority of studies documents a significant reduction effect for the social comparison intervention, results for the other interventions are mixed. Social comparison has been the most researched intervention, in terms of both quality and quantity. In particular, the several "Opower studies" that investigate the causal effect of social comparison based home energy reports (HERs) on energy consumption deliver broad empirical evidence and can be seen as best practice for program evaluations of energy conservation interventions. Yet, even for this intervention, many open questions remain: What is the actual pure effect of social comparison? Under which circumstances do social comparisons cause adverse effects? How much does the effect of social comparison interact with the medium by which it is delivered? First evidence on these questions indicate that adverse effects seem to matter and that it can make a difference if the social comparison is delivered by a letter, electronically online or via an In-Home-Display (IHD).

Commitment devices and goal setting have not yet been researched extensively and existing studies show major methodological shortcomings mainly in terms of underpowered samples. Yet, studies that document significant effects show conservation potential of around 10%. It is thus a promising avenue for further research to evaluate these interventions with a sufficiently large study sample, preferably in the field.

Although energy labels are applied worldwide, labeling is a very recent field of research. First evidence shows that labels can be effective with regard to the perception of and the willingness to pay (WTP) for energy efficiency. Early results by laboratory experiments are confirmed by some first field studies. Because the existing empirical evidence is mostly based on stated preference approaches, though, a promising field of future research is the evaluation of the effects of labels by analyzing revealed preferences.

The review moreover clearly shows that the amount of studies that satisfy the quality criterion of causal inference has increased in the last few years. There are presumably three reasons for this development: the improving quality of empirical research, an increasing interest in behavioral economics, and the need to find effective interventions to trigger energy conservation. However, there is still a lot to be done. Although we focused on studies that are potentially able to identify causal effects, only few studies within the sample contain evidence to back up concrete policy recommendations. Obstacles are, for instance, insufficiently powered studies, poor reporting of statistical tests, and study populations that are different from the target population. Furthermore, the minority of studies discusses the benefits and costs of the intervention, a prerequisite to give policy recommendations.

To sum up, we are surprised how little we know. We therefore call for systematic evaluations of these and similar interventions potentially able to trigger energy conservation before a large-scale rollout. It seems surprising that an intervention such as labeling is applied worldwide but there is little knowledge about the actual impact.

Future studies should focus on at least four crucial points: first, they should assess the causal effect of the intervention with a suitable methodology and a sufficiently powered sample that enables identification of even small effects with statistical precision. Second, researchers should refrain from combining too many different treatments. From a research perspective, we need evidence on the pure effect of a given intervention before we begin to potentially reinforce this effect by additional treatments. Third, long-term effects of the intervention should be analyzed and the intervention costs should be documented in order to enable cost-benefit analyses on the treatment's effectiveness as well as on its relative effectiveness compared to traditional policy instruments, such as taxes or bans (cf. Benartzi et al. 2017).

Fourth, a growing literature discusses that for a final decision whether an intervention is welfare enhancing and thus should be applied at large-scale, a classic cost–benefit analysis is not sufficient but a welfare analysis that takes into account *all* benefits and costs of an intervention is needed (cf. Allcott and Kessler 2015, Andor et al. 2017a, Ito 2015). Behavioral interventions might address externalities as well as internalities but can also cause hidden costs (Damgaard and Gravert, 2018). Allcott and Kessler (2015), for instance, show that while the majority of households has a positive WTP for continuing to receive HERs in the future, a certain percentage of households reveals a negative WTP. This indicates a disutility from receiving the HER. Future systematic evaluations should therefore aim at providing a comprehensive welfare analysis. At best, interventions should be evaluated ex-ante before they are rolled out at large-scale, for instance by randomized field experiments within the target population.

Acknowledgements

This article is a further developed and extended extract of a project report published in German (Andor and Fels 2017). We are grateful for invaluable comments and suggestions by the editor Dale S. Rothman, two anonymous referees, Gunther Bensch, Andreas Gerster, Jörg Langbein, Jörg Peters, Christoph M. Schmidt, Stephan Sommer, and Colin Vance. We also thank Marvin Gleue, Nadine Kneppel, Marc Teipel, and Maximilian Zettler for excellent research assistance. We gratefully acknowledge financial support by the National Academy of Science and Engineering (acatech). Furthermore, this work has been partly supported by the Collaborative Research Center "Statistical Modeling of Nonlinear Dynamic Processes" (SFB 823) of the German Research Foundation (DFG), within Project A3, "Dynamic Technology Modeling". Fels gratefully acknowledges the support of a special grant (Sondertatbestand) from the German Federal Ministry for Economic Affairs and Energy and the Ministry of Innovation, Science, and Research of the State of North Rhine-Westphalia.

Appendix A

The identified articles have been categorized according to the interventions social comparison, commitment/goal setting, labeling, and are ordered based on their publication date within each category. n = sample size; n.a. = not available; CG = control group; ***, **, * significance levels of 1%, 5% and 10%, respectively.

Table A1
"Final results table"

Study	Type of intervention [+ combination with additional interventions] and intervention design	Effect (in %) in comparison to: control group/ baseline	n	Dependent variable	Country	Method of causal analysis	Remarks
Social comparison							
Kantola et al. (1984)	(1) Feedback, Social Comparison [+ Framing, Energy saving tips] Households were informed once by letter that they showed an above-average electricity consumption despite having stated in an earlier survey that they feel an obligation to save electricity (the message pointed to the dissonance between answers in the survey and their actual consumption) + users were provided with pamphlets and cards on energy saving tips	(1) significantly lower electricity consumption/ n.a.	118	Electricity	AU	RCT	Long-run effects: n.a. Costs: n.a.
	(2) Feedback, Social Comparison [+ Energy saving tips] Households were informed once by letter that they showed an above-average electricity consumption + users were provided with pamphlets and cards on energy saving tips	(2) insignif./ n.a.					
	(3) Control Group [+ Energy saving tips] users were provided with pamphlets and cards on energy saving tips	(3) insignif./ n.a.					
	(4) Control Group Participants received a Thank-You letter for taking part in the experiment	(4) CG/n.a.					
Haakana et al. (1997)	(1) Feedback, Social Comparison [+ Energy saving tips] Participants received feedback and other information according to their wishes. Most households opted for comparisons of their own energy consumption with similar households in Finland in addition to personalized tips and an energy conservation video(2) Feedback, Social Comparison [+ Energy saving tips] As in treatment (1) but with tailored information in written form instead of a video	a) (1) n.a./ (−7) (2) n.a./(−5) (3) n.a./(−5) (4) n.a./(+1) b) (1) n.a./ (−4) (2) n.a./0 (3) n.a./(+1) (4) n.a./(+1)	105	a) Electricity b) Water c) Gas	FI	RCT	No information about significance of effects. Long-run effects: n.a. Costs: n.a.

	(3) Feedback, Social Comparison As in treatment (1), but without any tailored information or energy saving tips (4) Control Group Participants of the Control Group did not know they were part of an experiment	c) (1) n.a./ (−9) (2) n.a./(−7) (3) n.a./(−4) (4) n.a./(−3)					
Staats et al. (2004)	(1) Feedback, Social Comparison [+ Energy saving tips] Groups consisting of six to ten people met as „EcoTeams" once a month to discuss environmentally relevant behavior. During the sessions they received feedback about their energy savings. Topics at the meetings were: garbage, gas, electricity, water, transport and consumer behavior	a) n.a./(−4.6) [insignif.] b) n.a./ (−20.5)** c) n.a./(−2.8) [insignif.]	482	a) Electricity b) Gas c) Water	NL	Diff-in-Diff	Long-run effects: (2 years after stop of the EcoTeam-meetings) (−7.6)** (−16.9)** (−6.7)** Costs: n.a.
Schultz et al. (2007)	(1) Feedback, Social Comparison [+ Comparison to usage history, Energy saving tips] Handwritten messages with information about how much electricity the households consumed in the previous week + descriptive normative information about the energy consumption of the average household in their neighborhood + preprinted suggestions for how to conserve energy; sample restricted to high users	(1) n.a./ (−5.68)*	287	Electricity	US	RCT	Long-run effects: (−4.61) [insignif.] (+9.66)** (−5.97)* (+1.0) [insignif.] Costs: n.a.
	(2) Feedback, Social Comparison [+ Comparison to usage history, Energy saving tips] Same intervention as in (1) but sample restricted to low users	(2) n.a./ (+8.57)*					
	(3) Feedback, Social Comparison [+ Comparison to usage history, Energy saving tips, Graphical enhancement] Same intervention as in (1) + sad face, because the households' consumption was above average	(3) n.a./ (−8.34)**					
	(4) Feedback, Social Comparison [+ Comparison to usage history, Energy saving tips, Graphical Enhancement] Same intervention as in (2) + happy face, because the households consumed less than average	(4) n.a./ (+2.32) [insignif.]					
Peschiera et al. (2010)	(1) Feedback [+ Comparison to usage history, Reminder] Online-feedback about individual past and present electricity consumption + three weekly reminders to check the online-tool	(1) (−16) [insignif.]/ (−30) [insignif.]	83	Electricity	US	RCT	Long-run effects: n.a. Costs: n.a.
	(2) Feedback, Social Comparison [+ Comparison to usage history, Reminder]	(2) (−6) [insignif.]/					

	As in treatment (1) + comparison of the individual electricity consumption with the average consumption of the rest of the building (3) Feedback, Social Comparison [+ Comparison to usage history, Reminder] As in treatment (2) + information about the electricity consumption of "peers" (i.e. occupants of the building who in a pre-survey the treated households classified as "known") (4) Control group Participants knew they were part of an experiment	(−22) [insignif.] (3) (−21)***/ (−34)*** (4) CG/n.a.					
Allcott (2011)	(1) Feedback, Social Comparison [+ Comparison to usage history, Energy saving tips] Monthly/bimonthly/quarterly feedback via mail (Home Energy Report with personal history of consumption, comparison to neighbors and tips to save energy) (2) Control group No information about participation in study	(1) (−2.03)***/ n.a. (2) CG/n.a.	588,446	Electricity	US	RCT	Short-term effect measured one year after intervention. Long-run effects: According to the author there is no evidence of a reduction of the effect with the treatment lasting for two years, but no specific numbers stated (although, see Allcott and Rogers (2014)) Costs: Costs of production and shipping of the reports divided by kWh of saved energy: 3.31 cents/kWh
Ferraro et al. (2011) + Ferraro and Price (2013)	(1) Control group [+ Energy saving tips] Onetime letter with tips to save energy (2) Feedback [+ Energy saving tips, Framing (social norm)] Onetime personalized letter with heavily norm-based language about saving water, feedback from the consumption bill, tips to save energy (3) Feedback, Social Comparison [+ Energy saving tips, Framing (social norm)] Like (2) plus a comparison to the median water consumption of the preceding summer and a classification of the household into one of the consumption groups (4) Control group No information about participation in experiment	(1) (−0.66) [insignif.]/ (−8.41) (2) (−2.7)***/ (−10.08) (3) (−4.8)***/ (−12.01) (4) CG/ (−7.83)	106,669	Water	US	RCT	Long-run effects: (measured two years after intervention) (+0.9) [insignif.] (−0.22) [insignif.] (−1.3)** Costs: $ 0.575 per saved 1000 gal (= 3785.41 L) of water
Peschiera and Taylor (2012)	(1) Feedback, Social Comparison [+ Comparison to usage history] Participants get access to their electricity consumption data through an online-tool + information about consumption of the	(1) 0 [insignif.]/n.a.	88	Electricity	US	RCT	Effects are calculated for above-average users only. If below-average users are included in the sample, the effects turn insignificant.

	last seven days + a comparison with the average residential electricity use + weekly reminders to check the online-tool (2) Feedback, Social Comparison [+ Comparison to usage history] Same intervention as in (1) + data about the electricity use of 'peers' "(i.e. these occupants of the building who the treated households mentioned to "know" in a pre-survey)" (3) Control group Participants knew they were part of an experiment	 (2) (−8.8)**/ n.a. (3) CG/n.a.					Long-run effects: n.a. Costs: n.a.
Ayres et al. (2013)	(1) Feedback, Social Comparison [+ Energy saving tips, comparison to usage history, graphical enhancement] Home Energy Reports (HER) about electricity consumption via mail, monthly for heavy users, quarterly for light users, with tips to save energy, personal history of consumption and comparison to neighbors (additionally laughing or sad smiley, depending on consumption being below- or above-average) (2) Feedback, Social Comparison [+ Energy saving tips, Comparison to usage history, Graphical enhancement] Like (1), the HER additionally contained information about the gas consumption, frequency random (monthly/quarterly) (3) Feedback, Social Comparison [+ Energy saving tips, Comparison to usage history, Graphical enhancement] Like (2), gas consumption measured instead of electricity consumption (4) Control group No information about participation in experiment	(1) (−2.02)***/ n.a. (2) (−1.22)***/ n.a. (3) (−1.20)***/ n.a. (4) CG/n.a.	84,000	(1) + (2): Electricity (3) Gas	US	RCT	Long-run effects: n.a. Costs: 4.94 Cents per kWh saved 1.78 Cents per kWh saved n.a.
Costa and Kahn (2013)	(1) Feedback, Social Comparison [+ Energy saving tips, Comparison to usage history] Home Energy Reports (HER) about electricity consumption via mail, monthly for heavy users, quarterly for light users, with tips to save energy, personal history of consumption and comparison to neighbors (2) Control Group No information about participation in experiment	(1) Liberals: (−2.4)***/n.a. Conservatives: (−1.7)***/n.a. (2) CG/n.a.	81,772	Electricity	US	RCT	The study especially analyzes the intervention's heterogeneous effects ("Liberals" vs. "Conservatives"). The average treatment effect is (−2.1)*** compared to the control group. Long-run effects: n.a. Costs: According to the authors, the intervention might be cost-effective, but no specific statement.
Dolan and Metcalfe (2013)		(1) (−4.4)***/ n.a.	569	Gas	GB	RCT	Experiment 2 of the article fits thematically and is generally suitable

	(1) Feedback, Social Comparison Biannual feedback and comparison to neighbors' consumption via mail (2) Feedback, Social Comparison [+ Energy saving tips] Like (1) + Tips to save energy (3) Control Group Feedback about personal electricity consumption	(2) (−10.8)***/ n.a. (3) CG/n.a.					to identify causal effects, but the information in the working paper is contradictory and is therefore not listed here. Long-run effects: (18 months after first Intervention) (−7.0)* (−6.0) [insignif.] Costs: According to the authors 333 kWh were saved for every Pound spent
Mizobuchi and Takeuchi (2013)	(1) Feedback [+ Financial Incentive, Comparison to usage history] Monthly feedback via mail, reward of 200 Yen (2$) for 1% reduction of energy consumption (2) Feedback, Social Comparison [+ Financial Incentive, Comparison to usage history] Monthly feedback via mail, comparison to other participating households, reward of 200 Yen (2$) for 1% reduction of energy consumption (3) Control group Information about participation in study	(1) (−4.15)**/ (−5.876)** (2) (−6.48)**/ (−8.196)** (3) CG/ (−1.721)	208	Electricity	JP	RCT	Effects as difference from pre-treatment-consumption. The difference between the effects of (1) and (2) is statistically not significant. Long-run effects: n.a. Costs: n.a.
Tiefenbeck et al. (2013)	(1) Feedback, Social Comparison [+ Energy saving tips, Framing (social norm)] Weekly feedback about the water consumption per capita and tips to save water, partly complemented by a social comparison with the 10% saving the most water or a social appeal to contribute one's share to the mutual goal of energy conservation (2) Control group Information that water consumption will be monitored in the course of a scientific study	(1)a) (−6.0)**/n.a. b) (+5.6)**/ n.a. (2) CG/n.a.	154	a) Water b) Electricity	US	RCT	Study shows adverse effect of (intended) savings in water consumption for consumption of electricity. Since tenants privately pay for gas and electricity, while water is billed collectively, income effects can be excluded. Water consumption was measured daily, electricity consumption weekly. Long-run effects: No more significant effects measured after two weeks: Water: (−5.5) Electricity: (+0.3) Costs: n.a.
Allcott and Rogers (2014)	(1) Feedback, Social Comparison [+ Comparison to usage history, Energy saving tips] Monthly feedback via mail (Home Energy Report with personal history of consumption, comparison to neighbors and tips to save energy)	(1) (−1.7)***/ n.a.	234,000	Electricity	US	RCT	Long-run effects: 2009–2013 (following the measurement of the effect after receiving four reports, the treatment groups were randomly re-allocated): Stop of the intervention after two years: (−2)***

	(2) Feedback, Social Comparison [+ Comparison to usage history, Energy saving tips] Quarterly feedback via mail (Home Energy Report like in (1)) (3) Control Group No information about participation in study	(2) (−1.2)***/ n.a. (3) CG/n.a.					Biannual receipt of Home Energy Reports: (−3.1)*** Receipt of the Home Energy Reports in initial frequency: (−3.3)*** Costs: Costs of production and shipping of the reports divided by kWh of saved energy Assuming that the savings-effect does not last: 3.2–4.44 cents/kWh Assuming that the effects last for the long term: 1.35–1.79 cents/kWh
Delmas and Lessem (2014)	(1) Feedback, Social Comparison [+ Comparison to usage history, Reminder] Real-time feedback and comparison with other users through an online-platform, weekly reminder via e-mail (2) Feedback, Social Comparison [+ Comparison to usage history, Reminder] Additionally to (1) public rankings of which student rooms consume below- or above-average (via posters in the entrance hall and via e-mail) (3) Control group No information about participation in experiment	(1) (−5.68) [insignif.]/n.a. (2) (−19.36)**/ n.a. (3) CG/n.a.	66	Electricity	US	RCT	Long-run effects: 17 weeks after first intervention (−6.5) [insignif.] (−24.76)* Costs: n.a.
Komatsu and Nishio (2015)	(1) Feedback, Social Comparison Onetime feedback via mail and comparison to median consumption (2) Feedback, Social Comparison Like (1) plus comparison to highest saving 25% of participants (3) Feedback, Social Comparison [+ Framing (social norm)] Like (2) plus a message about the social acceptance/disapproval of own consumption (4) Feedback, Social Comparison Like (1) (5) Feedback, Social Comparison Like (2)	(1) No effect sizes given; tendencies for assessment of own consumption: higher** (2) higher *** (3) higher *** (4) Motivation to save energy: insignif. (5) insignif. (6) higher ***	3,033	Electricity (indirect)	JP	RCT	In (1)–(3) the participants' assessment of their own electricity consumption in comparison to the neighbors' consumption was analyzed, in (4)–(6) it was the households´ motivation to save energy. Long-run effects: n.a. Costs: n.a.

	(6) Feedback, Social Comparison [+ Framing (social norm)] Like (3) (7) Control group Simple feedback of electricity consumption	(7) n.a.					
Schultz et al. (2015)	(1) Feedback [+ Energy saving tips] Real-time feedback via IHD and provision of a "climate protection video" (2) Feedback [+ Energy saving tips] Real-time feedback via IHD with the consumption being converted to actual costs, video (3) Feedback, Social Comparison [+ Energy saving tips] Real-time feedback via IHD and comparison to other participants, video (4) Control Group Only video, no feedback	(1) (−3.0) [insignif.]/n.a. (2) insignif./ n.a. (3) (−9.0)**/ n.a. (4) CG/n.a.	431	Electricity	US	RCT	Long-run effects: (three months after Intervention) (−0.81) [insignif.] (+1.13) [insignif.] (−7.02)** Costs: n.a.
Seyranian et al. (2015)	(1) Control Group [+ Energy saving tips] Onetime letter with tips to save energy (2) Feedback, Social Comparison [+ Energy saving tips, Graphical enhancement] Like (1) plus information about own personal consumption in comparison to average consumption of neighborhood, complemented by a laughing or sad smiley (3) Control Group [+ Energy saving tips, Framing (social norm)] Like (1) plus a cover letter with the city logo, emphasizing the household's role as a part of the community and stressing on water conservation as a mutual goal (4) Control Group [+ Energy saving tips, Framing (personal norm)] Like (1) plus a cover letter, setting water conservation as a goal, but accentuating the household as a single player and not as a part of the community	(1) CG/ (+40.8) [insignif.] (2) (−29.74)*/ (−0.1) (3) (−34.1)**/ (+1.4) [insignif.] (4) (−24.6) [insignif.]/ (+3.4) [insignif.]	374	Water	US	RCT	Consumption was measured one week after the intervention Long-run effects: (four weeks after intervention) CG/(+16.2) [insignif.] (−11.5)**/(+2.6) [insignif.] (−12.2)**/(+2.6) [insignif.] (−12.1)***/(+3.5) [insignif.] Costs: n.a.
Alberts et al. (2016)	(1) Feedback, Social Comparison [+ Energy saving tips, Graphical enhancement] Weekly emailed energy report consisting of energy conservation tips and information about one's own consumption in comparison to	(1) (−20.9)**/ n.a.	89	Electricity and gas	UK	RCT	Long-run effects: (two weeks after start of intervention) (−30.8)**/n.a. (−9.9) [insignif.]/n.a. Costs: n.a.

	average consumption of all neighbors and to most efficient 20%; with rating as “great”, “good” or “more than average” accompanied by a smiley face (2) Feedback, Social Comparison [+ Energy saving tips, Graphical enhancement, Competition incentive] Like (2) plus information that there is a competition to find the resident with the lowest energy consumption over the course of the project, although no specific price was communicated (3) Control Group [+ Energy saving tips] One-shot e-mail at the start of the experiment containing a series of simple energy saving tips	(2) (−21.2)*/ n.a. (3) CG/n.a.					
Anderson et al. (2017)	Study 1 (graduate students):(1) Control Group [+ Feedback, Energy saving tips] Weekly messages emails with common energy use feedback information along with a few energy conservation tips (2) Feedback, Social Comparison [+ Energy saving tips, graphical enhancement] Like (1) plus information about average consumption of other similar residents and of the most efficient 10%, and a supporting message plus stars depending on the relative performance.	(1) CG/n.a. (2) per week of messaging (−1.2)**/n.a.	146	Electricity	KR	RCT	Long-run effects: (eleven weeks after start of intervention) CG insignif./n.a. CG insignif./n.a. Costs: n.a.
	Study 2 (undergraduate students): (3) Control Group [+ Feedback, Energy saving tips] Like (1) (4) Feedback, Social Comparison [+ Energy saving tips, graphical enhancement] Like (2)	(3) CG/n.a. (4) insignif./ n.a.	118	Electricity	KR	RCT	
Tiefenbeck et al. (2016)	(1) Feedback [+ Graphical enhancement] Smart shower meter that displays water use in tenth of liters, water temperature in degree Celsius, energy consumption in kWh, energy efficiency rating (from A to G), and a polar bear animation (ice floe that progressively shrinks as the amount of energy used increases) during shower (2) Feedback, Social Comparison [+ Graphical enhancement] Like (1) plus total amount of water used in the	(1) (−22.03)***/ n.a. (2) (−22.52)***/ n.a.	636	Electricity	CH	RCT + Diff-in-Diff	Long-run effects: n.a. Costs: n.a.

	previous shower (soc. comp. if more than one person lives in the household) (3) Control Group Smart shower meter displays water temperature only	CG/n.a.					
Hahn et al. (2016)	(1) Control Group [+"gain"-Framing] Letter containing $100 rebate coupon for purchasing drought-resistant plants phrased "You can have a free $100 landscape coupon" (2) Social Comparison [+"gain"-Framing] Like (1) plus a social comparison phrased "We have selected you because you used XX,XXX gallons more water in 2013 than the average home…" (3) Control Group [+"loss"-Framing] Like (1) but the rebate coupons were phrased "Do not lose your chance to have a free $100 landscape coupon" (4) Social Comparison [+"loss"-Framing] Like (2) plus a social comparison phrased 'We have selected you because you used XX,XXX gallons more water in 2013 than the average home…'	CG/n.a. (2) (−1.4)**/ n.a. (3) CG/n.a. (4) (−1.1) [insignif.]/n.a.	23,282	Water	US	RCT	The study looks at three outcomes: number of rebates requested, number of rebates redeemed, and changes in monthly water use. This table only lists the effects on water use. Long-run effects: n.a. Costs: n.a. for direct water use, $10.64 per thousand gallons of estimated water savings via rebate coupons
Shen et al. (2015)	(1) Feedback, Social Comparison [+ Energy saving tips, Graphical enhancement] Mailed home energy reports with energy consumption compared to average consumption of the *entire neighborhood* and of most efficient 20% with rating as 'great', 'good' or 'more than average' accompanied by smiley face and energy conservation tips. (2) Feedback, Social Comparison [+ Energy saving tips, Graphical enhancement] Like (1) but comparison group were street neighbors (3) Feedback, Social Comparison [+ Energy saving tips, Graphical enhancement] Like (1) but comparison group were next-door neighbors (4) Control group No information about participation in an experiment	(1) (−2.8) [insignif.]/n.a. (2) (−5.4)***/ n.a. (3) (−2.1) [insignif.]/n.a. CG/n.a.	475	Electricity	US	RCT	Long-run effects: n.a. Costs: n.a.

Commitment devices and goal setting

Becker (1978)	(1) Goal Setting [+ Energy saving tips] Participants were asked to reduce their energy consumption by 20% compared to their baseline usage + they received energy conservation tips specific to their own household appliances	(1) (−1.3) [insignif.]/n.a.	100	Electricity	US	RCT	Long-run effects: n.a. Costs: n.a.
	(2) Feedback, Goal Setting [+ Energy saving tips] Same intervention as in (1) + they received feedback about their energy consumption three times a week	(2) (−13) [insignif.]/n.a.					
	(3) Goal Setting [+ Energy saving tips] Participants were asked to reduce their energy consumption by 2% compared to their baseline usage + they received energy conservation tips specific to their own household appliances	(3) (+1.2) [insignif.]/n.a.					
	(4) Feedback, Goal Setting [+ Energy saving tips] Same intervention as in (3) + they received feedback about their energy consumption three times a week	(4) (−4.6) [insignif.]/n.a.					
	(5) Control group Participants of the control group knew they were part of an experiment	(5) CG/n.a.					
Winett et al. (1979)	(1) Feedback, Goal Setting [+ Comparison to usage history, Graphical enhancement, Energy saving tips] Daily feedback sheet with information about electricity usage + comparison with own usage on the preceeding day + a happy or frowning smiley for a decrease or increase in consumption + usage change compared to baseline + feedback whether they achieved a goal, which they had set themselves in a meeting before the start of the experiment + energy conservation information	(1) (−13)***/ n.a.	71	Electricity	US	RCT	Long-run effects: (1) (−11) [insignif.] (2) (−7) [insignif.] Effects are calculated relative to the mean consumption of the two control groups. Costs: Feedback condition: – total expenditure per household = $26 – Savings per household = $44 from expected expenditures based on the comparison group's use during this same period and the marginal cost per kWh. Self-monitoring: – total expenditure per household = $22 – Savings per household = $26
	(2) Feedback [+ Energy saving tips] Households had to read and report their daily electricity consumption + received energy saving information	(2) (−7)*/n.a.					
	(3) Control Group Households agreed to participate in the study	(3) CG/n.a.					
	(4) Control Group Households denied to participate in the study	(4) CG/n.a.					
Winett et al. (1982)			83	Electricity	US	RCT	

	Study 1 (during winter): (1) Feedback, Goal Setting [+ Energy saving tips] Participants took part in a meeting and received information about energy conservation + they had to sign a commitment to reduce their energy consumption by 15% within the next 35 days + they received written feedback about their energy use + they were provided with a tape which presented information on energy conservation in form of a discussion	(1) n.a./ (−13)***					Long-run effects: n.a. Costs: n.a.
	(2) Goal Setting [+Energy saving tips] Participants took part in a meeting and received information about energy conservation + they had to sign a commitment to reduce their energy consumption by 15% within the next 35 days + they received a video which presented model homes similar to the participants´ homes and explained possible ways to reduce energy consumption.	(2) n.a./ (−11)***					
	(3) Feedback, Goal Setting [+ Energy saving tips] As in treatment (2) + feedback	(3) n.a./ (−14)***					
	(4) Goal-Setting [+ Energy saving tips] As in treatment (1) but without feedback	(4) n.a./(−1) [insignif.]					
	(5) Control Group Participants of the control group knew they were part of an experiment	(5) CG/(+4)					
	Study 2 (during summer): (6) Feedback, Goal Setting As in treatment (3) but without information on energy conservation	(6) n.a./ (−19)***					
	(7) Feedback, Goal Setting [+ Energy saving tips] As in treatment (1) but without discussion tape	(7) n.a./ (−22)***					
	(8) Goal Setting [+ Energy saving tips] As in treatment (2)	(8) n.a./(−12) [insignif.]					
	(9) Control Group Participants of the control group knew they were part of an experiment	(9) CG/(−2)					
Katzev and Johnson (1983)	(1) Goal Setting „Foot-in-the-door-treatment": Households were asked to fill in an energy conservation questionnaire. Afterwards, they were told to reduce their electricity consumption by 10% within the next two weeks	(1) n.a./(+11) [insignif.]	66	Electricity	US	RCT	The subjects were recruited through a door-to-door solicitation procedure in which they were asked for permission to read their electricity meters as part of a study on residential energy consumption.
	(2) Goal Setting Households were asked to reduce their	(2) n.a./(+12) [insignif.]					Long-run effects:

	electricity consumption by 10% within the next two weeks (3) Control Group (4) Control Group Households agreed to have their electricity meters read by the experimenter	 (3) n.a./(+7) [insignif.] (4) CG/0					(+1)** (−2)*** (−5)** (+5) (all effects compared to baseline period; significance compared to control group) Costs: n.a.
Katzev and Johnson (1984)	(1) Goal Setting [+ Energy saving tips] Households were asked to reduce their energy consumption by 15% within the next two weeks + received a utility guide for household energy conservation (2) Goal Setting Households were asked to reduce their energy consumption by 15% within the next two weeks + had to fill in an energy conservation survey (3) Control Group [+ Financial incentive, Energy saving tips] Households received a financial reward depending on the amount of electricity saved + a utility guide for household energy conservation (4) Goal Setting [+ Financial incentive, Energy saving tips] Households received a combination of the interventions in (2) and (3) (5) Control Group Households were asked to fill in a short energy conservation survey (6) Control Group Households agreed to have their electricity meters read by the experimenters	(1) (−6.1) [insignif.]/ (−13.4) [insignif.] (2) (+1.4) [insignif.]/ (−5.9) [insignif.] (3) (+2.0) [insignif.]/ (−5.3) [insignif.] (4) (−3.7) [insignif.]/ (−11.0) [insignif.] (5) (+7.5) [insignif.]/ (+0.2) [insignif.] (6) CG/(−7.3) [insignif.]	90	Electricity	US	RCT	Long-run effects: (1) (−2.8)/(−14.3) (2) (−1.3)/(−12.8) (3) (+0.2)/(−11.3) (4) (−10.7)/(−22.2) (5) (−1.5)/(−13.0) (6) CG/(−11.5) All effects are insignif. Costs: n.a.
van Houwelingen and van Raaij (1989)	(1) Feedback, Goal Setting [+ Energy saving tips] SmartMeter and IHDs were installed and households were asked to monthly jot down their preferred and actual consumption of gas. Externally imposed goal: 10% consumption reduction in comparison to preceding year. Additionally tips to save energy. (2) Feedback [+ Energy saving tips] Monthly feedback (not stated through which canal) and tips to save energy	(1) n.a./ (−12.3)*** (2) n.a./ (−7.7)***	325	Gas	NL	RCT	Long-run effects: (one year after end of the intervention) (1) insignif./(−2.1)* (2) insignif./(−3.2)** (3) n.a./(−1.5) (insignif.) (4) n.a./(+1.4) (insignif.) (5) n.a./(−2.2)* (6) n.a./(−2.9)* Costs: n.a.

	(3) Feedback [+ Energy saving tips] Households were asked to document their monthly gas consumption on observation sheets, additionally tips to save energy (4) Control Group [+ Energy saving tips] Provision of tips to save energy (5) Control Group No information about participation in experiment (6) Control Group Contacted households did not want to participate in the experiment	(3) n.a./ (−5.1)*** (4) n.a./ (−4.3)** (5) n.a./(−0.3) [insignif.] (6) CG/(−0.2) [insignif.]					
McCalley and Midden (2002)	Laboratory study, in which the participants had to set up washing machines on a computer: the first six washing cycles for an evaluation of the „base consumption“, which the participants received feedback about. The following interventions were randomly assigned, afterwards another 20 washing cycles. (1) Control Group No Feedback, no goal (2) Feedback Feedback, no goal (3) Feedback, Goal Setting, Commitment Feedback, self-imposed goal (Options: 0%, 5%, 10%, 15%,20%) (4) Feedback, Goal Setting Feedback, externally imposed goal (20%)	(1) CG/no difference to (2) (2) n.a./(−9.6) [insignif.] (3) (−12.3***)/ (−21.9)*** (4) (−9.9)**/ (−19.5)**	100	Electricity	NL	Laboratory Experiment	Long-run effects: n.a. Costs: n.a.
Abrahamse et al. (2007)	(1) Feedback, Goal Setting Externally set goal of energy conservation (5%), online feedback about personal consumption after two and five months (2) Feedback, Goal Setting [+ Framing (social norm)] Like (1) plus a mutual goal (all together saving 5%) and online feedback about the group's consumption after two and five months (3) Control Group Twice filling out a questionnaire, no other intervention	(1) + (2) significant difference to control group***/ (−8.3) (3) CG/(+0.4) [insignif.]	189	Electricity, Gas (and gasoline)	NL	RCT	Effects refer to direct energy consumption, not to the as well measured indirect consumption, because the indirect consumption strongly varied and was not seen as robust by the authors. For the analysis of the direct energy consumption both treatment groups were taken together, as there were no significant differences between them. Long-run effects: n.a. Costs: n.a.
Harding and Hsiaw (2014)	Offer for customers to participate in program to save electricity. Range of goals to set for oneself between 0%, 0–15%, 15–50%, over 50% (1) Feedback, Goal Setting, Commitment [+ Comparison to usage history]	(1) (−1.5) [insignif.]/n.a. (2) (−11.0)**/ n.a. (3) (−1.0)	12,451	Electricity	US	Matching	Households with optimistic goals (15–50%) save quite a lot shortly after start of the program. This effect wanes after two to three months, presumably because the consumers realized that they would not be able to reach their

	Self-set goal: 0%, daily access to website with monthly bills of consumption (2) Feedback, Goal Setting, Commitment [+ Comparison to usage history] Self-set goal: 0–15%, daily access to website with monthly bills of consumption (3) Feedback, Goal Setting, Commitment [+ Comparison to usage history] Self-set goal: 15–50%, daily access to website with monthly bills of consumption (4) Feedback, Goal Setting, Commitment [+ Comparison to usage history] Self-set goal: > 50%, daily access to website with monthly bills of consumption (5) Control Group No information about participation in experiment	[insignif.]/n.a. (4) 0/n.a. (5) CG/n.a.					very optimistic goals. Long-run effects (after 18 months) for (2): Significant effects with significance level of 95%; no specific statements about size Costs: n.a.
Jaeger and Schultz (2017)	(1) Social Comparison [+ Energy saving tips] During a drought, residents received doorhangers with the social norm: “Over 80% of households in your community use efficient landscape irrigation techniques” plus water conservation information (2) Social Comparison, Commitment [+ Energy saving tips] Like (1) plus commitment “I, resident of ___ commit to efficiently watering my landscape. __ (sign.)” (3) Framing (warning) [+ Energy saving tips] Doorhanger including the warning “During the drought, please keep in mind that using excessive and unnecessary amounts of water can result in penalties, including fines of up to $500, water service interruption, and/or prosecution.”, additionally: water conservation information. (4) Commitment [+ Energy saving tips, Framing (warning)] Like (3) plus the commitment ‘I, resident of ___ commit to efficiently watering my landscape. __ (sign.)’ (5) Control Group [+ Energy saving tips] Doorhanger with information and some tips for water conservation behavior (6) Control Group No information control.	(1) n.a. (2) for compliers (−3.53)***/ n.a. (3) n.a. (4) for compliers: (−5.56)***/ n.a. (5) CG/n.a. (6) CG/n.a.	5,363	Water	US	RCT	Long-run effects: (four months post intervention) (1) n.a. (2) (−8.00)*** (3) n.a. (4) (−3.2) [insignif.] (5) CG (6) CG Costs: n.a.

Labeling							
Verplanken and Weenig (1993)	Laboratory experiment: Choice between four refrigerators (1) Control Group Information on consumption in kWh, no time pressure (2) Labeling [+ graphical enhancement] Annual energy consumption as graphical label (measured in local currency), no time pressure (3) Control group Information on consumption in kWh, time pressure (5 min for decision) (4) Labeling [+ graphical enhancement] Annual energy consumption as graphical label, time pressure	(1) 27 (2) 50* (3) 33 (4) 23	120	Electricity (indirect)	NL	Choice-experiment with random assignment to experimental groups	Effect = share of group that prefers an efficient refrigerator to an inefficient one in % Long-run effects: n.a. Costs: n.a.
Kurz et al. (2005)	(1) Labeling A series of labels was placed on different appliances in the household: refrigerators, air conditioners, showers, washing machines, clothes dryers, dishwashers, toilets and outdoor taps. The labels provided information about the water and energy-consumption levels of the labeled appliances (2) Control Group Households were provided with the same information as in treatment (1) but in the form of information leaflets instead of labels (3) Feedback, Social Comparison [+ Graphical enhancement] Households received e-mails with graphical feedback on their levels of water and energy consumption and a comparison to other households of similar size who participated in the study	(1) n.a./from insignif. to (−23)** (2) n.a./ insignif. (3) n.a./ insignif.	166	Electricity and Water	AU	RCT	The study reports the effects for seven weeks separately, without documenting an average treatment effect. Therefore, we report the interval of the documented effects. Long-run effects: n.a. Costs: n.a.
Brounen and Kok (2011)	Labeling Analysis of the effect of the EU label for energy efficiency on the sales price of houses with a „green “label (A, B or C)	(+3.7)***	31,993	Electricity Water Gas (indirect)	NL	Observation study, controlled for self-selection with two-step Heckman-procedure	Effect = average mark-up in % on houses with "green" label relative to comparable houses Long-run effects: n.a. Costs: n.a.
Heinzle (2012)	<u>Experiment 1:</u> Participants are asked to evaluate the energy saving potential of different TVs. The provided information varies:	(1) 18.6/66.3/ 15.1	(1)–(3) 252	Electricity (indirect)	DE	Choice-Experiments with random assignment to experimental	Effects (1)–(3) = Proportion of participants who correctly estimate the

	(1) Labeling Consumption information in Watt and current price for electricity (2) Labeling Consumption information in Watt (3) Labeling Actual annual operating costs of the devices <u>Experiment 2:</u> Choice between two TVs with differing information: (4) Control Group Consumption information in Watt (60 W vs. 225 W) (5) Labeling Absolute operating costs for the course of ten years (180€ vs. 660 €) (6) Labeling Annual operating costs (18€/year vs. 66€/year)	(2) 19.8/64.0/16.3 (3) 92.5/3.8/3.8 *** (significant difference to the other treatment groups) (4) 481.22 (5) 641.96** (6) 353.97**	(4)–(6) 208			groups within a survey	energy saving potential in %/who overestimate/who underestimate Effects (4)–(6) = Median willingness to pay for a more efficient TV in Euro (actual saving potential: 480 Euro). Long-run effects: n.a. Costs: n.a.
Houde (2014)	<u>Natural Experiment 1:</u> (1) Labeling The tightening of the criteria for the receipt of the Energy Star Label in 2008 in the US allowed to observe the willingness to pay for the same models of refrigerators in the same store with and without the label (before and after the change of regulation) <u>Natural Experiment 2:</u> (2) Labeling In January 2010 it was published that some refrigerators received the Energy Star Label, although they did not actually fulfill the criteria (incorrect appliance of the test procedure). Because the following withdrawal of the label could not be anticipated by either the producers or the customers, the label's value can be determined by the willingness to pay for these models of refrigerators before and after January 2010.	(1) (+19)**/1.5 (2) (+89.6) [insignif.]/7.07	(2) 184,645	Electricity (indirekt)	US	Two natural experiments	Effects (1) + (2) = Average willingness to pay for the Energy Star Label in $/share of total price for refrigerator in % Long-run effects: n.a. Costs: The opportunity costs of imperfect information amount to about $15 per refrigerator sold. Extrapolated on the whole market for refrigerators that makes a sum of $135 mil/year. According to the author that is twice as much as the annual costs for the Energy Star Label program.
Newell and Siikamäki (2014)	Survey (online) about six different hypothetical purchase decisions for warm water processors. There are three boilers to choose from. The boilers differ in price and each boiler is labeled with a different version of the Energy Star		1,184	Electricity and Gas (indirect)	US	Choice-experiment with random assignment to experimental groups	Effects = relative willingness to pay for energy-efficient devices: a WTP of 1 stands for cost minimizing behavior, where a change in sales price and a change in the operating costs are equally weighted. A value lower than 1

	Label, which differ in the kind of information that is provided: (1) Labeling Only the simple consumption information is provided (2) Labeling Like (1) + relative consumption costs in comparison to different devices (3) Labeling Like (2) + Information about the CO2 emissions (4) Labeling Like (2) + specific information on energy consumption (5) Labeling [+ graphical enhancement] Wie (4) + awarded with an "Energy-Star-Label “ (6) Labeling [+ graphical enhancement] Like (1) + Information about energy consumption in relation to other devices in overseeable nuances (A-G ranking: „EU-Style“)	(1) 0.7** (2) 0.68** (3) 0.93 [insignif.] (4) 1.02 [insignif.] (5) 1.36** (6) 1.34**					shows an underestimation of savings through energy efficiency, while a value > 1 shows an overestimation. The significance is measured as the difference to 1. Long-run effects: n.a. Costs: n.a.
Ölander and Thøgersen (2014)	Choice between four TV devices with different labels; question which device the customer would buy (1) Labeling Labels with energy efficiency grades from A - G (2) Labeling Labels with energy efficiency grades from A+ + + - D	(1) + (2) The scale from A – G more than doubled the probability to choose an energy-efficient device in comparison to the scale from A+ + + − D	151	Electricity (indirect)	DK	Choice-experiment with random assignment to experimental groups	Long-run effects: n.a. Costs: n.a.
Allcott and Taubinsky (2015)	<u>Experiment 1:</u> Choice between an energy-saving bulb and three conventional bulbs with comparable power. (1) Labeling Information about consumption, cost saving with energy-saving bulbs, downsides of energy-saving bulbs (+ Information like in (2)) (2) Control Group Information about amount of energy-saving bulbs sold and sales development in the past.	(1) (+2.30)***/ n.a. (2) CG	(1) + (2) 1,533	Electricity (indirect)	US	<u>Experiment 1:</u> „Artefactual field experiment“: Choice - Experiment with random assignment to experimental groups. One of 30 choices (randomly selected) leads to	For experiment 1 several treatment groups were summarized for the analysis. Effects (1) = Willingness to pay in Dollar for an energy-saving bulb in comparison to a conventional bulb and in comparison to the control group (2)

	Experiment 2: Customers in a big electrical store were asked about their consumption behavior and with their individual information a comparison of the energy consumption of energy-saving bulbs and conventional bulbs was provided via iPad					actual buying decision.	
	(3) Labeling [+ Financial incentive] Annual and total energy costs + discount coupon (10% on all bulbs)	(3) (−2.2) [insignif.]	(3)–(6) 1,087			Experiment 2: RCT	Effects (3)–(6) = Probability of the purchase of an energy-saving bulb in comparison to conventional bulbs after the information treatment and in comparison to the control group in percentage points.
	(4) Labeling [+ Financial incentive] Annual and total energy costs + discount coupon (10% on all bulbs, 30% on energy-saving bulbs)	(4) (+11.0) [insignif.]					Long-run effects: n.a. Costs: n.a.
	(5) Control Group [+ Financial incentive] No interview/information, discount coupon (10% on all bulbs, 30% on energy-saving bulbs)	(5) + (7.8)*					
	(6) Control Group [+ Financial incentive] No interview/information, discount coupon (10% on all bulbs)	(6) CG; Probability for purchase of an energy-saving bulb: 38%					
Waechter et al. (2015)	Experiment 1: On the basis of a scale from 0 (not efficient) to 100 (very efficient) the participants had to estimate the energy consumption of a TV. Four different versions of a label (randomly assigned) provided information on the grade of energy efficiency (A – high, B – low) and the electricity consumption (high, low).	Experiment 1:	Experiment 1: 166	Electricity (indirect)	Online-experiment, which was conducted in Switzerland	Experiment 1: Choice-Experiment with random assignment to experimental groups	Dependent variables: Experiment 1: Estimation of the energy efficiency of the TV devices
	(1) Labeling Information about the grade of energy efficiency of the TV devices	(1) The higher the grade of energy efficiency, the lower the estimates for the electricity consumption (significant***)					
	(2) Control Group Information about the electricity consumption of the TV devices	(2) No effect on estimates of electricity consumption					
	Experiment 2: On the basis of two devices with different labels, participants had to decide which device they would recommend to an energy	Experiment 2:	Experiment 2: 305			Experiment 2: no groups	Experiment 2: Share of participants who wrongly recommended the device with the higher actual energy consumption in %

	conserving person. The device with the higher grade of energy efficiency thereby featured the higher actual energy consumption. (3) Labeling Device = TV (4) Labeling Device = Freezer	(1) 44.6 (2) 72.8				
	Experiment 3: On the basis of two different labels participants had to evaluate the energy consumption of a freezer in comparison to a reference refrigerator, whereas the labels differed only in the grade of energy efficiency, but not in actual energy consumption. (5) Labeling Label: "Grade of energy efficiency: A+++" (6) Control Group "Grade of energy efficiency: A+"	Experiment 3: (1) 67.72** (2) 77.07 (CG)	Experiment 3: 166		Experiment 3: Choice-Experiment with random assignment to experimental groups	Experiment 3: Estimation of the energy consumption of the freezer compared to the reference refrigerator on a scale from 1 to 100 Long-run effects: n.a. Costs: n.a.
Jensen et al. (2016)	Natural experiment that examines the effect of the EU-requirement to include an energy performance rating from A to G in the real estate advertisement, which went into effect in 2010. The documented effect sizes refer to the change in property sales price before and after 2010 when compared to a certain reference group. Model 1: categorizes good and bad ratings in two exclusive groups (1) Labeling good energy performance rating of buildings (A to C) (2) Control Group energy performance rating of buildings D to G Model 2: compares each energy rating to the reference rating G (3) Labeling Energy performance rating A or B (4) Labeling Energy performance rating C (5) Control Energy performance rating D (6) Labeling Energy performance rating F (7) Labeling Energy performance rating G	(1) from (+2.4)*** to (+10.1)***/ n.a. (2) CG (1) from (+6.6)*** to (+6.2)***/n.a. (2) from (+0.2) to (+5.1)***/n.a. (3) CG (4) from (−1.5)*** to (−5.4)***/n.a. (5) from (−9.3)*** to (−24.3)***	117,483	DK	FE	Long-run effects: n.a. Costs: n.a.

Appendix B

The keyword search in EconLit comprised the categories "Title", "Abstract", and "Subjects (SU)" for both journal articles and working papers. In ScienceDirect, we searched with keywords in the category "Abstract, Title, Keywords" and included articles from the disciplines 'Economics, Econometrics, Finance', 'Psychology', 'Social Sciences', 'Environment', and "Energy".

Table B1
"List of keywords".

Intervention	Keywords
Social comparison	"Social norms" OR "social learning" OR "social modeling" OR "social influence" OR "peer comparison" OR "peer information" OR "comparative energy information" OR "feedback" *AND* "energy conservation" OR "energy consumption" OR "energy use" OR "energy usage" OR "energy demand*" OR "energy saving" OR "electricity conservation" OR "electricity consumption" OR "electricity use" OR "electricity usage" OR "electricity demand*" OR "electricity saving" OR "gas conservation" OR "gas consumption" OR "gas use" OR "gas usage" OR "gas demand*" OR "gas saving" OR "water conservation" OR "water consumption" OR "water use" OR "water usage" OR "water demand*" OR "water saving" OR "conservation behavior"
Commitment	"Pre-commitment" OR "precommitment" OR "pledge" OR "behavioral contract" OR "commitment contract" OR "commitment devices" OR "commitment approach*" OR "personal commitment" OR "public commitment" OR "self-control" OR "self-regulation" *AND* 'energy conservation' OR 'energy consumption' OR 'energy use' OR 'energy usage' OR 'energy demand*' OR 'energy saving' OR 'electricity conservation' OR 'electricity consumption' OR 'electricity use' OR 'electricity usage' OR 'electricity demand*' OR 'electricity saving' OR 'gas conservation' OR 'gas consumption' OR 'gas use' OR 'gas usage' OR 'gas demand*' OR 'gas saving' OR 'water conservation' OR 'water consumption' OR 'water use' OR 'water usage' OR 'water demand*' OR 'water saving' OR 'conservation behavior'
Goal-setting	'Goal setting' *AND* 'energy conservation' OR 'energy consumption' OR 'energy use' OR 'energy usage' OR 'energy demand*' OR 'energy saving' OR 'electricity conservation' OR 'electricity consumption' OR 'electricity use' OR 'electricity usage' OR 'electricity demand*' OR 'electricity saving' OR 'gas conservation' OR 'gas consumption' OR 'gas use' OR 'gas usage' OR 'gas demand*' OR 'gas saving' OR 'water conservation' OR 'water consumption' OR 'water use' OR 'water usage' OR 'water demand*' OR 'water saving' OR 'conservation behavior'
Labeling	"Energy labeling" OR "energy labelling" OR "information label*" OR "energy information" OR "energy label" OR "information acquisition" OR "information disclosure" OR "environmental certification" *AND* 'Energy conservation' OR 'energy consumption' OR 'energy use' OR 'energy usage' OR 'energy demand*' OR 'energy saving' OR 'electricity conservation' OR 'electricity consumption' OR 'electricity use' OR 'electricity usage' OR 'electricity demand*' OR 'electricity saving' OR 'gas conservation' OR 'gas consumption' OR 'gas use' OR 'gas usage' OR 'gas demand*' OR 'gas saving' OR 'water conservation' OR 'water consumption' OR 'water use' OR 'water usage' OR 'water demand*' OR 'water saving' OR 'conservation behavior'

Analysis refers to this standardized inclusion decision form in Citavi. If one of the criteria was coded with “no”, the study was excluded from the systematic review.

Table B2
“Inclusion decision form”.

Author (year)	Text
Title	Text
Name of coder	Text
Study includes applied research (not just theoretical models)	Options: yes/no/discuss
Study includes at least one of the chosen interventions	Options: Feedback/Social Comparison/Commitment/Goal-Setting/Labeling/discuss
Study refers to at least one of the outcome variables	Options: gas/water/electricity/discuss
Study targets private households or individuals in private households	Options: yes/no/discuss
Study was carried out in a developed country	Options: yes/no/discuss
Inclusion decision	Options: include/exclude/relevant meta-study/discuss

Table B3
“Coding sheet“.

Study	Authors (year)	Text
Type of intervention	e.g. *1* *1 + 2* *1 + 2 + 4* *0*	Open code 0 = Control group 1 = Feedback 2 = Social comparison 3 = Commitment 4 = Goal-setting 5 = Labeling 1 + 2 = Feedb. + Social Comparison 3 + 4 = Commitment + Goal-Setting
Intervention design	e.g. *Onetime letter, external goal (5%)* *Onetime letter, self-set goal* *Onetime letter, external goal, comparison with another treatment group*	Text
Combination with additional interventions	e.g. *None* *Energy saving tips* *None*	Text

Effect (in %) in comparison to: control group/baseline	e.g.	Number
Significance level:	*(−2.2)**/(−5.4)***	
***: p < .01	*(−2.7)*/(−10.0)**	
**: p < .05	*(−0.5) [insignif.]/(−1.9)**	
*: p < .10	*KG/(−1.1) [insignif.]*	
(insignif.): not significant		
N	e.g.	Number
	34,000	
Dependent variable	e.g.	Open Code
	2	1 = Electricity
		2 = Water
		3 = Gas
		For Labeling:
		1A = Electricity (indirectly)
		2A = Water (indirectly)
		3A = Gas (indirectly)
Country	e.g.	Country code
	US	
Method of causal analysis	e.g.	Open code
	1	1 = RCT
		2 = Matched Comparison
		3 = RDD
		4 = Diff-in-Diff
		5 = FE
		6 = Other
Remarks	e.g.	Text
	(1) Limitations	
	(2) Cost–benefit analysis	
	(3) Long-run effect + follow-up period after intervention	

References

Abrahamse, W., Steg, L., Vlek, C., Rothengatter, T., 2005. A review of intervention studies aimed at household energy conservation. J. Environ. Psychol. 25 (3), 273–291.

Abrahamse, W., Steg, L., Vlek, C., Rothengatter, T., 2007. The effect of tailored information, goal setting, and tailored feedback on household energy use, energy-related behaviors, and behavioral antecedents. J. Environ. Psychol. 27 (4), 265–276. http://dx.doi.org/10.1016/j.jenvp.2007.08.002.

Alberts, G., Gurguc, Z., Koutroumpis, P., Martin, R., Muuls, M., Napp, T., 2016. Competition and norms: a self-defeating combination? Energ Policy 96, 504–523.

Allcott, H., 2011. Social norms and energy conservation. J. Public Econ. 95 (9–10), 1082–1095. http://dx.doi.org/10.1016/j.jpubeco.2011.03.003.

Allcott, H., 2015. Site selection bias in program evaluation. Q. J. Econ. 130 (3), 1117–1165.

Allcott, H., 2016. Paternalism and energy efficiency: an overview. Annu. Rev. Econ. 8, 145–176.

Allcott, H., Kessler, J.B., 2015. The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisons (NBER Working Paper No. 21671). National Bureau of Economic Research.

Allcott, H., Mullainathan, S., 2010. Behavior and energy policy. Proc. Am. Assoc. Adv. Sci. 327 (5970), 1204–1205. http://dx.doi.org/10.1126/science.1180775.

Allcott, H., Rogers, T., 2014. The short-run and long-run effects of behavioral interventions: experimental evidence from energy conservation. Am. Econ. Rev. 104 (10), 3003–3037. http://dx.doi.org/10.1257/aer.104.10.3003.

Allcott, H., Taubinsky, D., 2015. Evaluating behaviorally motivated policy: experimental evidence from the lightbulb market. Am. Econ. Rev. 105 (8), 2501–2538. http://dx.doi.org/10.1257/aer.20131564.

Anderson, K., Song, K., Lee, S.H., Krupka, E., Lee, H., Park, M., 2017. Longitudinal analysis of normative energy use feedback on dormitory occupants. Appl. Energy 189, 623–639. http://dx.doi.org/10.1016/j.apenergy.2016.12.086.

Andor, M., Fels, K., 2017. *Energiesparen durch verhaltensökonomisch motivierte Maßnahmen?*. acatech Schriftenreihe Energiesysteme der Zukunft.

Andor, M., Gerster, A., Peters, J., Schmidt, C.M., 2017a. Social norms and energy conservation beyond the US, (Ruhr Economic Papers #714).

Andor, M., Gerster, A., Sommer, S., 2017b. Consumer Inattention, Heuristic Thinking and the Role of Energy Labels, (Ruhr Economic Papers #671).

Angrist, J.D., Pischke, J.-S., 2009. Mostly Harmless Econometrics: An empiricists' Companion. Princeton University Press.

Ayres, I., Raseman, S., Shih, A., 2013. Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage. J. Law Econ. Org. 29 (5), 992–1022. http://dx.doi.org/10.1093/jleo/ews020.

Bandura, A., 1986. Social Foundations of Thought and Action: A Social Cognitive Theory. Prentice Hall.

Becker, L.J., 1978. Joint effect of feedback and goal setting on performance: a field study of residential energy conservation. J. Appl. Psychol. 63 (4), 428–433. http://dx.doi.org/10.1037/0021-9010.63.4.428.

Benartzi, S., Beshears, J., Milkman, K.L., Sunstein, C.R., Thaler, R.H., Shankar, M., Tucker-Ray, W., Congdon, W.J., Galing, S., 2017. Should governments invest more in nudging? Psychol. Sci. 28 (8), 1041–1055. http://dx.doi.org/10.1177/0956797617702501.

Brandon, A., Ferraro, P.J., List, J.A., Metcalfe, R.D., Price, M.K., Rundhammer, F., 2017. Do The Effects of Social Nudges Persist? Theory and Evidence from 38 Natural Field Experiments. (NBER Working Paper No. 23277). National Bureau of Economic Research.

Brounen, D., Kok, N., 2011. On the economics of energy labels in the housing market. J. Environ. Econ. Manag. 62 (2), 166–179. http://dx.doi.org/10.1016/j.jeem.2010.11.006.

Campbell Collaboration, 2014. Campbell Systematic Reviews: Policies and Guidelines, Campbell Systematic Reviews 2014: Supplement 1.

Chetty, R., 2015. Behavioral economics and public policy: a pragmatic perspective. Am. Econ. Rev. 105 (5), 1–33.

Costa, D.L., Kahn, M.E., 2013. Energy conservation 'nudges' and environmentalist ideology: evidence from a randomized residential electricity field experiment. J. Eur. Econ. Assoc. 11 (3), 680–702. http://dx.doi.org/10.1111/jeea.12011.

Damgaard, M.T., Gravert, C., 2018. The hidden costs of nudging: Experimental evidence from reminders in fundraising. J. Public Econ. 157, 15–26.

Delmas, M.A., Lessem, N., 2014. Saving power to conserve your reputation? The effectiveness of private versus public information. J. Environ. Econ. Manag. 67 (3), 353–370. http://dx.doi.org/10.1016/j.jeem.2013.12.009.

Delmas, M.A., Fischlein, M., Asensio, O.I., 2013. Information strategies and energy conservation behavior: a meta-analysis of experimental studies from 1975 to 2012. Energ Policy 61, 729–739. http://dx.doi.org/10.1016/j.enpol.2013.05.109.

Dolan, P., Metcalfe, R., 2013. Neighbors, Knowledge, and Nuggets: Two Natural Field Experiments on the Role of Incentives on Energy Conservation (Economics Working Paper No. 2589269). Institute for Research, Becker Friedman. http://dx.doi.org/10.2139/ssrn.2589269.

Ferraro, P.J., Price, M.K., 2013. Using nonpecuniary strategies to influence behavior: evidence from a large-scale field experiment. Rev. Econ. Stat. 95 (1), 64–73. http://dx.doi.org/10.1162/REST_a_00344.

Ferraro, P.J., Miranda, J.J., Price, M.K., 2011. The persistence of treatment effects with norm-based policy instruments: evidence from a randomized environmental policy experiment. Am. Econ. Rev. 101 (3), 318–322. http://dx.doi.org/10.1257/aer.101.3.318.

Frondel, M., Vance, C., 2013. Energy efficiency: don't belittle the rebound effect. Nature 494(430-430) http://dx.doi.org/10.1038/494430c.

Haakana, M., Sillanpaeae, 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.

Hahn, R., Metcalfe, R.D., Novgorodsky, D., Price, M.K., 2016. The Behavioralist as Policy Designer: The Need to Test Multiple Treatments to Meet Multiple Targets (Experimental Economics Center Working Paper Series No. 2016-05.). Experimental Economics Center, Andrew Young School of Policy Studies, Georgia State University.

Harding, M., Hsiaw, A., 2014. Goal setting and energy conservation. J. Econ. Behav. Organ. 107, 209–227. http://dx.doi.org/10.1016/j.jebo.2014.04.012.

Heinzle, S.L., 2012. Disclosure of energy operating cost information: a silver bullet for overcoming the energy-efficiency gap? J. Consum. Policy 35 (1), 43–64. http://dx.doi.org/10.1007/s10603-012-9189-6.

Houde, S., 2014. How Consumers Respond to Environmental Certification and the Value of Energy Information (NBER Working Paper No. 20019). National Bureau of Economic Researchhttp://dx.doi.org/10.3386/w20019.

van Houwelingen, J.H., van Raaij, W.F., 1989. The effect of goal-setting and daily electronic feedback on in-home energy use. J. Consum. Res. 16 (1), 98–105. http://dx.doi.org/10.1086/209197.

Imbens, G.W., Wooldridge, J.M., 2009. Recent developments in the econometrics of program evaluation. J. Econ. Lit. 47 (1), 5–86. http://dx.doi.org/10.1257/jel.47.1.5.

Ito, K., 2015. Asymmetric incentives in subsidies: evidence from a large-scale electricity rebate program. Am. Econ. J. Econ. Pol. 7 (3), 209–237.

Jaeger, C.M., Schultz, P.W., 2017. Coupling social norms and commitments: testing the underdetected nature of social influence. J. Environ. Psychol. 51, 199–208. http://dx.doi.org/10.1016/j.jenvp.2017.03.015.

Jensen, O.M., Hansen, A.R., Kragh, J., 2016. Market response to the public display of energy performance rating at property sales. Energ Policy 93, 229–235.

Kahneman, D., 2003. Maps of bounded rationality: psychology for behavioural economics. Am. Econ. Rev. 93 (5), 1449–1475.

Kahneman, D., Tversky, A., 2013. Prospect theory: an analysis of decision under risk. In: Handbook of the Fundamentals of Financial Decision Making: Part I, pp. 99–127.

Kantola, S.J., Syme, G.J., Campbell, N.A., 1984. Cognitive dissonance and energy conservation. J. Appl. Psychol. 69 (3), 416–421. http://dx.doi.org/10.1037/0021-9010.69.3.416.

Karlin, B., Ford, R., Zinger, J., 2015. The effects of feedback on energy conservation: a meta-analysis. Psychol. Bull. 141, 1205–1247.

Katzev, R.D., Johnson, T.R., 1983. A social-psychological analysis of residential electricity consumption: the impact of minimal justification techniques. J. Econ. Psychol. 3 (3–4), 268–284. http://dx.doi.org/10.1016/0167-4870(83)90006-5.

Katzev, R.D., Johnson, T.R., 1984. Comparing the effects of monetary incentives and foot-in-the-door strategies in promoting residential electricity conservation. J. Appl. Soc. Psychol. 14 (1), 12–27. http://dx.doi.org/10.1111/j.1559-1816.1984.tb02217.x.

Komatsu, H., Nishio, K., 2015. An experimental study on motivational change for electricity conservation by normative messages. Appl. Energy 158, 35–43. http://dx.doi.org/10.1016/j.apenergy.2015.08.029.

Kőszegi, B., Szeidl, A., 2012. A model of focusing in economic choice. Q. J. Econ. 128 (1), 53–104.

Kurz, T., Donaghue, N., Walker, I., 2005. Utilizing a social-ecological framework to promote water and energy conservation: a field experiment. J. Appl. Soc. Psychol. 35 (6), 1281–1300. http://dx.doi.org/10.1111/j.1559-1816.2005.tb02171.x.

Laibson, D., 1997. Golden eggs and hyperbolic discounting. Q. J. Econ. 112 (2), 443–478.

Lokhorst, A.M., Werner, C., Staats, H., van Dijk, E., Gale, J.E., 2013. Commitment and behavior change: a meta-analysis and critical review of commitment-making strategies in environmental research. Environ. Behav. 45 (1), 3–34.

McCalley, L.T., Midden, C.J.H., 2002. Energy conservation through product-integrated feedback: the roles of goal-setting and social orientation. J. Econ. Psychol. 23 (5), 589–603. http://dx.doi.org/10.1016/S0167-4870(02)00119-8.

Mizobuchi, K., Takeuchi, K., 2013. The influences of financial and non-financial factors on energy-saving behaviour: a field experiment in Japan. Energ Policy 63, 775–787. http://dx.doi.org/10.1016/j.enpol.2013.08.064.

Newell, R.G., Siikamäki, J.V., 2014. Nudging energy efficiency behavior: the role of information labels. J. Assoc. Environ. Resour. Econ. 1 (4), 555–598. http://dx.doi.org/10.3386/w19224.

Nisan, M., Horenczyk, G., 1990. Moral balance: the effect of prior behaviour on decision in moral conflict. Br. J. Soc. Psychol. 29 (1), 29–42.

Nobelprize.org (2017). Richard H. Thaler - facts. Nobelprize.org. Nobel Media AB 2014. Web. 14 Nov 2017. Link: http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2017/thaler-facts.html.

O'Donoghue, T., Rabin, M., 1999. Doing it now or later. Am. Econ. Rev. 89 (1), 103–124.

O'Donoghue, T., Rabin, M., 2008. Procrastination on long-term projects. J. Econ. Behav. Organ. 66 (2), 161–175.

Ölander, F., Thogerson, J., 2014. Informing versus nudging in environmental policy. J. Consum. Policy 37, 341–356. http://dx.doi.org/10.1007/s10603-014-9256-2.

Peschiera, G., Taylor, J.E., 2012. The impact of peer network position on electricity consumption in building occupant networks utilizing energy feedback systems. Energ. Buildings 42 (8), 1329–1336. http://dx.doi.org/10.1016/j.enbuild.2012.03.011.

Peschiera, G., Taylor, J.E., Siegel, J.A., 2010. Response–relapse patterns of building occupant electricity consumption following exposure to personal, contextualized and occupant peer network utilization data. Energ. Buildings 42 (8), 1329–1336. http://dx.doi.org/10.1016/j.enbuild.2010.03.001.

Reisch, L., Zhao, M., 2017. Behavioural economics, consumer behaviour and consumer policy: state of the art. Behavioural Public Policy 1 (2), 190–206. http://dx.doi.org/10.1017/bpp.2017.1.

Schubert, R., Stadelmann, M., 2015. Energy-using durables – why consumers refrain from economically optimal choices. Front. Energy Res. 3 (7), 1–13. http://dx.doi.org/10.

3389/fenrg.2015.00007.
Schultz, P.W., Nolan, J.M., Cialdini, R.B., Goldstein, N.J., Griskevicius, V., 2007. The constructive, destructive, and reconstructive power of social norms. Psychol. Sci. 18 (5), 429–434. http://dx.doi.org/10.1111/j.1467-9280.2007.01917.x.
Schultz, P.W., Estrada, M., Schmitt, J., Sokoloski, R., Silva-Send, N., 2015. Using in-home displays to provide smart meter feedback about household electricity consumption: a randomized control trial comparing kilowatts, cost, and social norms. Energy 90 (1), 351–358. http://dx.doi.org/10.1016/j.energy.2015.06.130.
Seyranian, V., Sinatra, G.M., Polikoff, M.S., 2015. Comparing communication strategies for reducing residential water consumption. J. Environ. Psychol. 41, 81–90. http://dx.doi.org/10.1016/j.jenvp.2014.11.009.
Shen, M., Cui, Q., Fu, L., 2015. Personality Traits and Energy Conservation. Energy Policy 85, 322–334. http://dx.doi.org/10.1016/j.enpol.2015.05.025.
Staats, H., Harland, P., Wilke, H.A.M., 2004. Effecting durable change: a team approach to improve environmental behavior in the household. Environ. Behav. 36, 341–367. http://dx.doi.org/10.1177/001391650326016.
Thaler, R.H., Sunstein, C.R., 2008. Nudge: Improving Decisions about Health, Wealth, and Happiness. Yale University Press.
Tiefenbeck, V., Staake, T., Roth, K., Sachs, O., 2013. For better or for worse? Empirical evidence of moral licensing in a behavioral energy conservation campaign. Energ Policy 57, 160–171. http://dx.doi.org/10.1016/j.enpol.2013.01.021.
Tiefenbeck, V., Goette, L., Degen, K., Tasic, V., Fleisch, E., Lalive, R., Staake, T., 2016. Overcoming salience bias: how real-time feedback fosters resource conservation. Manag. Sci. https://pubsonline.informs.org/doi/pdf/10.1287/mnsc.2016.2646.
Verplanken, B., Weenig, M.W.H., 1993. Graphical energy labels and consumers' decisions about home appliances: a process tracing approach. J. Econ. Psychol. 14 (4), 739–752. http://dx.doi.org/10.1016/0167-4870(93)90019-H.
Waechter, S., Sütterlin, B., Siegrist, M., 2015. The misleading effect of energy efficiency information on perceived energy friendliness of electric goods. J. Clean. Prod. 93, 193–202. http://dx.doi.org/10.1016/j.jclepro.2015.01.011.
Winett, R.A., Neale, R.A., Grier, H.C., 1979. Effects of self-monitoring and feedback on residential electricity consumption. J. Appl. Behav. Anal. 12 (2), 173–184. http://dx.doi.org/10.1901/jaba.1979.12-173.
Winett, R.A., Hatcher, J.W., Fort, T.R., Leckliter, I.N., Love, S.Q., Riley, A.W., Fishback, J.F., 1982. The effects of videotape modeling and daily feedback on residential electricity conservation home temperature and humidity, perceived comfort and clothing worn: winter and summer. J. Appl. Behav. Anal. 15 (3), 381–402. http://dx.doi.org/10.1901/jaba.1982.15-381.