Psychological Bulletin
2015, Vol. 141, No. 6, 1205–1227

0033-2909/15/$12.00 http://dx.doi.org/10.1037/a0039650

The Effects of Feedback on Energy Conservation: A Meta-Analysis

Beth Karlin and Joanne F. Zinger
University of California, Irvine

Rebecca Ford
University of Otago

Feedback has been studied as a strategy for promoting energy conservation for more than 30 years, with studies reporting widely varying results. Literature reviews have suggested that the effectiveness of feedback depends on both how and to whom it is provided; yet variations in both the type of feedback provided and the study methodology have made it difficult for conclusions to be drawn. The current article analyzes past theoretical and empirical research on both feedback and proenvironmental behavior to identify unresolved issues, and utilizes a meta-analysis of 42 feedback studies published between 1976 and 2010 to test a set of hypotheses about when and how feedback about energy usage is most effective. Results indicate that feedback is effective overall, $r = .071$, $p < .001$, but with significant variation in effects (r varied from $-.080$ to $.480$). Several treatment variables were found to moderate this relationship, including frequency, medium, comparison message, duration, and combination with other interventions (e.g., goal, incentive). Overall, results provide further evidence of feedback as a promising strategy to promote energy conservation and suggest areas in which future research should focus to explore how and for whom feedback is most effective.

Keywords: feedback, proenvironmental behavior, meta-analysis, energy conservation

National and global focus on energy use is at an all-time high. Although physical scientists are working to develop alternative energy sources, psychologists can also contribute to this issue by developing and testing interventions for demand-side reduction through behavior change. Energy use in identical homes was found to vary by up to 260% (Parker, Mazzara, & Sherwin, 1996), indicating that, in addition to the building itself, the behavior of occupants within the building impacts overall energy use. As such, interventions that target behavior may result in substantial energy savings.

Residential energy use accounts for over 20% of annual emissions (Environmental Protection Agency, 2011), making it a prime target for intervention. Dozens of changes in the use of energy within the home can be made in the immediate term, without economic sacrifice or loss of well-being on the part of consumers (Dietz, Gardner, Gilligan, Stern, & Vandenbergh, 2009; Gardner & Stern, 2008). Household energy conservation was identified as an efficient and effective means of reducing emissions, using currently available technology, with potential cost savings of up to 25% that could yield up to $300 billion in gross energy savings through 2020 (Granade et al., 2009). This savings potential, or "behavioral wedge," provides "both a short-term bridge to gain time for slower-acting climate mitigation measures and an important component of a long-term comprehensive domestic and global climate strategy" (Dietz et al., 2009, p. 18455).

Although a variety of energy conservation actions are technically and economically viable, widespread energy conservation is lagging and policymakers are increasingly looking to psychologists for guidance (Lutzenhiser et al., 2009; C. Wilson & Dowlatabadi, 2007). Thirty years ago, Bittle, Valesano, and Thaler (1979–1980) said that

> the need for conservation of existing resources presents social scientists with an opportunity to develop techniques for guiding human behavior in such a way as to enable us to exist in greater harmony with our environment and its natural limitations. (p. 188)

This is now truer than ever, and the analysis of psychological interventions that promote residential energy conservation is a vital and important topic of study.

One such promising intervention is the provision of feedback to individuals and groups about their energy use. Feedback refers to the process of giving people information about their behavior that can be used to reinforce and/or modify future actions. It is considered an important dimension of behavior change (Bandura, 1969; Skinner, 1938), and has been used to influence behavior in a wide variety of fields, including education (Bridgeman, 1974; Hanna, 1976), public health (Becoña & Vázquez, 2001; Tate, Wing, & Winett, 2001), and organizational behavior (Guzzo, Jette, & Katzell, 1985; Pearce & Porter, 1986).

Feedback in the energy domain has received increasing attention in recent years because of changes in sensing technology and energy infrastructure that allow for energy information to be

This article was published Online First September 21, 2015.

Beth Karlin, School of Social Ecology, University of California, Irvine; Joanne F. Zinger (formerly Joanne Frattaroli), Department of Psychology and Social Behavior, University of California, Irvine; Rebecca Ford, Centre for Sustainability, University of Otago.

We thank the following students for their valuable assistance with data collection and analysis: Christina Baghdikian, Kristen Figueira, Eleonora Goneva, and Anita Au. In addition, we thank Wesley Schultz, Sarah Darby, and Karen Herter for sharing hard-to-find studies, as well as Dan Stokols and Nora Davis for valuable input on previous drafts.

Correspondence concerning this article should be addressed to Beth Karlin, 5548 Social & Behavioral Science Gateway, University of California, Irvine, CA 92697-7075. E-mail: bkarlin@uci.edu

collected, processed, and sent back to consumers quickly, cheaply, and often in real time. Billions of dollars are being spent to upgrade electricity infrastructure across the globe to a smart grid, which includes the replacement of traditional electricity meters with digital meters that allow for wireless communication of information back to the utility and potentially to the consumer as well. The U.S. government is trying to accelerate this transition through programs like the American Reinvestment and Recovery Act of 2009, which allocated $3.4 billion for smart-meter installations, and the White House Green Button Initiative, which encourages utility providers to supply consumers with real-time access to their energy information (Chopra, 2011).

This type of government support, as well as over 200 feedback products and services that have emerged in recent years, demonstrate both political and technical potential for wide-scale provision of improved energy feedback to consumers (Karlin, Ford, & Squiers, 2014). However, many questions remain as to how and for whom energy feedback works. Furthermore, previous research on energy feedback has been criticized for its lack of theoretical rigor, and critics have called for more attention to the conditions under which theories are successful in explaining conservation behavior (Katzev & Johnson, 1987; P. W. Schultz, 2010; Steg & Vlek, 2009). An improved understanding of the variables that moderate and the mechanisms that underlie energy feedback would be of benefit at both a theoretical and a practical level.

The current article presents a meta-analysis of research on the effects of feedback on energy conservation. It aims to integrate analytical and synthetic analyses to determine what is known about energy feedback interventions in a residential setting and to make suggestions for future research. As such, the following sections will (a) analyze past theoretical and empirical research on both feedback and proenvironmental behavior to identify unresolved issues; (b) apply feedback theory to the unique context of residential energy use to derive a set of specific hypotheses; (c) test these hypotheses on the existing body of empirical literature about residential energy feedback, examining the overall effectiveness of feedback on energy conservation and what variables moderate this effect; and (d) integrate findings with the literature to offer concrete suggestions for future research and practice in the use of residential energy feedback.

Literature Review

Changes in electricity infrastructure over the past decade have enabled new methods of both collecting and providing energy information to consumers, leading to increased interest in residential energy feedback. Although there is great technological promise, it is vital that such efforts maximize not only the technological potential of feedback but also its psychological potential. As such, a look back at past theory and research about both feedback and proenvironmental behavior can enable us to apply this information to the specific context of residential energy feedback.

Psychological Theories of Feedback

Before coming to a deeper understanding of how feedback on energy use might affect behavior, it is important to examine feedback more broadly—how does any feedback about performance in any domain affect behavior? Psychological research dating back over a century provides some insights. The earliest psychological research related to feedback focused on knowledge of results (KR) studies (e.g., Jones, 1910; Judd, 1905; Wright, 1906); these studies provided information back to the research subject about the results of the experimental task (e.g., "You answered 80% of the questions correctly") and often found a positive relationship between KR and performance. Early work in behaviorism (e.g., Skinner, 1938; Thorndike, 1927) related KR to feedback through operant conditioning, which introduced the concepts of reinforcement and punishment (such that a desired response to a behavior serves as behavioral reinforcement, and an undesired response serves as punishment). Knowledge of desired results can be seen as a reinforcement of behavior and knowledge of undesired results as a punishment, thus serving to encourage or discourage behavior.

Later work (Bandura, 1969) expanded to include feedback about not only the *results* of a behavior but also the *process* of engaging in that behavior (e.g., "You attended three classes this week"), as well as information relating results to a goal (e.g., "You are on track to earn an A this semester") or peer performance (e.g., "You are in the top 10% of your class"). Bandura (1969), who contributed seminal research on feedback, found that providing a goal and information about progress toward that goal could serve as a form of behavior modification, much like providing a reward or punishment. Similarly, goal-setting theory (Locke & Latham, 1990) views behavioral feedback as a form of self-regulation, asserting that behavior is inherently goal-directed, and feedback about performance is needed to evaluate behavior in relation to these goals. Action-identification theory (Vallacher & Wegner, 1987) asserts that different levels of meaning can be attributed to an action, and the level of meaning can change for a given behavior; as mastery is gained from performing a behavior, its meaning moves up from action-related identity and goals (e.g., run a mile without stopping) to self-related identity and goals (e.g., improve physical fitness).

Kluger and DeNisi (1996) conducted a comprehensive review of psychological theories about feedback and a statistical meta-analysis of feedback studies across multiple behaviors (e.g., test performance, attendance, memory tasks).[1] They found that despite many previous assertions of feedback effectiveness, the empirical evidence was mixed; some studies found strong positive effects for feedback, whereas others found null or negative effects. They introduced feedback intervention theory (FIT) to explain this variation. FIT integrates a series of arguments derived from their analysis of past empirical and theoretical work, focused around three main concepts: standards, goals, and attention.

Standards. The primary argument of FIT is that behavior is regulated by comparisons made between the feedback and the preexisting or intervention-provided *standards*. These standards can be personal goals (Carver & Scheier, 1981; Latham & Locke, 1991) or comparisons with past behavior or others in a social group (Festinger, 1954). When behavior differs from the standard, a *feedback-standard gap* is created, and an individual's desire to decrease this gap mediates the effectiveness of feedback. A standard can be created and provided by the intervention, but it is only effective if the individual accepts and values the standard. Four

[1] Feedback about energy use was not included in their analysis.

options are therefore available to individuals when provided with a feedback-standard gap. They can respond by changing their behavior to match the standard, changing the standard to match their behavior, rejecting the feedback, or leaving the situation altogether. The specific response to feedback can be affected by variables related to the feedback information or by individual-level differences (e.g., level of self-efficacy or anxiety). Both the source and strength of the standard and the size and direction of the feedback-standard gap can impact this choice. For example, research has found that negative feedback is more likely than positive feedback to influence behavior (e.g., Anderson & Rodin, 1989; Campion & Lord, 1982; Podsakoff & Farh, 1989).

Goals. A second key characteristic of feedback is the presence of the feedback provided or an existing goal that the individual accepts and values. According to FIT, feedback goals are organized in a hierarchy. Consistent with action identification theory (Vallacher & Wegner, 1987), the action–goal relationship spans from low-level correspondence, to specific action (e.g., attending lectures), to higher level identities that focus on self-salient outcomes (e.g., becoming a scientist). Goals relating to the focal task (e.g., passing university exams) sit between the self-salient, or metatask, goals and specific action goals.

Attention. FIT proposes that attention is generally directed in the middle of the hierarchy (focal task goals) and that the output of higher level feedback loops may impact lower level goals. Feedback-standard gaps that are salient to the self (e.g., the gap between current perception and desired scientific identity) can be resolved in a number of ways, one of which may be to focus on the focal task (e.g., passing university exams) and the lower level actions (e.g., attending lectures). However, these gaps may also be resolved by other activities (e.g., taking an internship at a scientific institute), which may result in the focal task (passing university exams) receiving less attention or being abandoned altogether. Alternatively, unattained metatask goals may cause people to respond by increasing the standard of focal task goals; if scientific identity standards are not met, one may respond by raising goals related to passing university exams by aiming for an even higher grade. Satisfying these new task goals can further the higher, self-salient goal. This view also provides a supporting explanation for why positive feedback can impact behavior, even though it did not induce a feedback-standard gap: An even higher-level goal can be set that creates a new standard.

Combining these three elements, FIT asserts that feedback is effective insofar as it directs the locus of individuals' attention to a feedback-standard gap that is relevant to a preestablished or feedback-provided goal that is relevant to the individual. Only feedback-standard gaps that receive attention contribute to behavior regulation. The simple presence of feedback is not enough to regulate behavior—the feedback must draw the attention of individuals' to a feedback-standard gap that they have identified as self-relevant. Although attention is generally directed at a level somewhere above physical action (Carver & Scheier, 1981) and below ultimate self-goals (Wicklund, 1975), this can vary as a function of task familiarity and goal attainment (Vallacher & Wegner, 1987). Feedback may direct attention to a specific action or standard and connect that action to self-related goals, serving not only to provide information about the behavior-standard gap but also to draw attention to a behavior in the first place, and place it in context with those goals. As such, the visibility and availability of feedback are also essential and serve as key factors in its effectiveness.

Task Characteristics of Energy Behavior

In developing FIT, Kluger and DeNisi (1996) successfully integrated past research on feedback and provided a coherent set of theoretical assumptions that have implications for interventions across a wide variety of behavioral domains. However, it is important to consider the specific task characteristics of energy use to apply this work successfully. Past research has discussed this need but has done little to address it. Kluger and DeNisi (1996) note that feedback researchers have largely "ignored the theoretical importance of task characteristics." We have identified four key task characteristics of energy use that may affect the applicability of the FIT model to energy feedback, namely, that energy use is abstract, nonsensory, comprised of multiple behaviors, and of low personal relevance to most individuals.

Abstract. Energy use is abstract in nature. People do not consciously use energy with the goal of impacting the environment; they use appliances in the home that require energy. That energy is generated in power plants by burning fossil fuels, which releases greenhouse gases into the environment. Thus, an individual's abstract notions about the concept of environmental impacts are at least one step removed from their concrete (observable) behaviors that use energy.

Although this is a minor distinction from a technical point of view, it can be seen as an important psychological distinction when considering strategies to promote behavior change. Markowitz and Shariff (2012) studied climate change behaviors and found that their abstractness and cognitive complexity make efforts to promote energy-conserving behaviors difficult. Related to this point, they introduced an explanatory construct regarding the *blamelessness* of unintentional action. Most individuals are not trying to emit carbon when watching TV or cooking dinner. Rather, it is seen as a necessary by-product of these actions that is not worthy of blame and does not need to be changed.

Nonsensory. Energy use is nonsensory in nature. Many forms of energy use, such as electricity, are invisible, silent, and untouchable. One cannot see energy directly or touch it. We cannot pick up a kilowatt-hour like an apple. Environmental products like reusable shopping bags and hybrid vehicles can become elements of a lifestyle, as they are visible and easily seen by others. Alternatively, energy use is less visible by peers and even by the user. As such, receiving and paying attention to feedback about one's energy use is optional. That is, people have the option to view feedback or not view it, or even to purchase or not purchase an energy feedback device or product. Kluger and DeNisi (1996) suggest that the issue of locus of attention is "about the what (will receive attention) and not about the if (it will be perceived at all)" (p. 262). However, because energy feedback is optional for people most of the time, "the if" also matters a great deal—user experiences and perceptions are crucial.

Multiple behaviors. Energy use does not consist of a single target behavior; rather, it consists of a large set of behaviors that can vary from watching TV to turning on the lights. The principle of compatibility (Ajzen & Fishbein, 1977) suggests behaviors and their influences should be measured at the same level of specificity. Although proenvironmental behavior is often addressed holistically with encouragements to "go green" as if it were a single action, there is great diversity in the types of environmental actions that a person can choose. Even within a subset of energy use such as lighting efficiency, we can differentiate between turning off lights, installing energy efficient lighting, or setting light timers in the home. Although the result of all three behaviors is a decrease in home energy, they may be quite different in terms of influencing factors, environmental impact, and psychological consequences. These behaviors vary widely in task characteristics such as cost, effort, and required knowledge; research suggests they are predicted by different motivations as well as different demographic characteristics (Karlin et al., 2014).

Furthermore, the energy savings of individual conservation behaviors are generally not well understood by the public, making it hard for individuals to identify which ones may lead to meaningful energy savings. Research has shown an interest, among members of the U.S. public, in engaging in behaviors that are aimed at reducing their environmental impact. However, the specific behaviors in which Americans overwhelmingly report engaging, such as turning off lights when leaving a room, have a minimal impact on energy savings compared with, for example, installing solar panels (Attari, DeKay, Davidson, & Bruine de Bruin, 2010).

Personal relevance. Finally, FIT asserts that feedback interventions "are unlikely to be ignored because any FI (feedback intervention) has potentially serious implications for the self" (Kluger & DeNisi, 1996, p. 262). This is not necessarily the case with proenvironmental behaviors such as electricity use, as the implications are often minimal to the self, given that home energy is relatively inexpensive and that energy use causes no immediate personal harm. Americans do report concern for environmental issues, but this concern often ranks lower than other concerns that are related to the economy, health care, and terrorism, which have more serious and immediate implications for the self (Leiserowitz, 2008).

Some behaviors and related feedback-standard gaps are more important or motivationally significant to individuals than others. Individuals are less likely to pay attention to, or to try to resolve, feedback-standard gaps associated with activity domains that they consider trivial or insignificant compared with those associated with subjectively important activity domains (Stokols, 1979). An understanding of the unique predictors, including the motivational underpinnings, of environmental behavior is therefore an important topic to consider.

Determinants of Proenvironmental Behavior

Because of these unique task characteristics, a theoretical understanding of energy use and its predictors is important for maximizing the potential utility of a feedback intervention. Proenvironmental behavior refers to individual or collective actions that result in decreased resource use and/or environmental impact, which may be achieved by engaging in behaviors that benefit the environment as well as reducing behaviors that harm the environment. Such behaviors can range from personal consumption of resources to collective action on large-scale political and social issues. A substantial body of research has been conducted on the determinants of proenvironmental behavior (see Bamberg & Moser, 2007, for review). These models often include a combination of attitudinal and contextual variables.

Attitudes. Attitudes play a significant role in determining behaviors related to sustainability. Two psychological approaches, rational (individualistic) theories and moral (altruistic) theories, have been tested for their utility in predicting and explaining such behaviors (Bamberg & Moser, 2007). Rational theories, such as the rational choice model (Simon, 1955) and the theory of planned behavior (Ajzen, 1991), focus on intentional behavior choices resulting from a process of evaluating expected utility. Such theories presume that individuals are primarily driven by self-interest and that favorable beliefs about attitudes, norms, and/or efficacy will translate into favorable behavior intentions and, ultimately, favorable behaviors.

Because environmental issues generally involve the use of natural resources, which are both collective and limited, the optimal choice for an individual is often in direct conflict with the common interest (Hardin, 1968). As such, altruistic or moral motives are also important for understanding proenvironmental behavior. The norm activation model (Schwartz, 1977), for example, stipulates that the activation of a *personal norm* or a sense of moral obligation influences prosocial behavior. Although originally applied to behavior toward other people, later work suggested that this concern for the well-being of others could extend to nonhuman species and nature (Van Liere & Dunlap, 1978). Such norms were found to explain many proenvironmental behaviors, including vegetarianism, recycling, and energy conservation.

Although this contrast between rational and moral approaches to understanding behavior is a recurring theme in psychology, it is important to acknowledge that the two are not mutually exclusive, and their integration can yield greater theoretical and explanatory value than either can alone (Turaga, Howarth, & Borsuk, 2010). Psychological variables that were found to predict proenvironmental behaviors included those representing both a rational and a moral approach, such as concerns about the environment, price sensitivity, both personal and social norms, and efficacy (e.g., Barr, Gilg, & Ford, 2005; Samuelson & Biek, 1991; Verhallen & van Raaij, 1981).

Context. Attitudes, although important for predicting and influencing behavior, may not be sufficient to override individual and structural barriers to behavior. Individual barriers include lack of time, money, or knowledge required to engage in proenvironmental behaviors. Prior research points to home ownership, income, family size, and age as the most significant demographic predictors of energy conservation, such that older, high-income families who own their homes are the most likely to engage in such behaviors (Black, Stern, & Elworth, 1985; Cialdini & Schultz, 2004; Dillman, Rosa, & Dillman, 1983; Karlin et al., 2014; Nair, Gustavsson, & Mahapatra, 2010; Poortinga, Steg, Vlek, & Wiersma, 2003). Physical characteristics of homes, as well as knowledge, may exert influence over proenvironmental behavior (Steg & Vlek, 2009). Understanding these contexts is critical to predicting behavior because variables such as home, location, and size strongly predict a household's carbon footprint, and variables such as income and home ownership are closely related to one's ability to engage in certain environmental behaviors (Stern, 2011). Cultural differences may also influence behavior (Stephenson et al., 2010). Home heating behaviors, for example, may reflect different energy cultures comprised of the material culture (e.g., housing materials), related practices (e.g., time spent in the home), and norms (e.g., thermal comfort levels).

A theory known as the A-B-C (attitude-behavior-context) model (Guagnano, Stern, & Dietz, 1995) posits that environmental behavior is influenced by both attitudes and contextual factors, and that the stronger one set of factors is in predicting behavior, the less force the other exerts. If there are sufficient contextual barriers to engaging in a behavior, then individuals are unlikely to engage in it regardless of attitudes or motivations. For example, Black et al. (1985) found that some behaviors, such as adding home insulation, were not associated with normative beliefs when constrained by contextual factors, such as household infrastructure and homeownership. Likewise, Guagnano et al. (1995) found that the explanatory power of personal norm beliefs (i.e., motivational elements) decreased for recycling behavior when convenient curbside pickup was available.

However, even for highly motivated individuals who are able to perform a target behavior, change may not occur without the provision of a *trigger* to highlight when and where it is needed (Fogg, 2009). According to Fogg (2009), there are three types of forms that a feedback trigger can take. Feedback can serve as a *spark* to help emphasize self-salient goals and increase the importance of proenvironmental behaviors for the self. Feedback can also act as a *facilitator*, helping individuals who are already motivated identify ways to engage in their target behavior, increasing knowledge through the provision of information about proenvironmental behaviors. Finally, feedback may act as a *signal*, serving as a reminder to perform the behavior for individuals who are already motivated and able to do so. This is especially important in the case of environmental behavior, in which impacts are often invisible and untouchable. In this context, eco-feedback, defined as "feedback on individual or group behaviors with a goal of reducing environmental impact" (Froehlich, Findlater, & Landay, 2010, p. 1), can provide the necessary trigger to engage people in proenvironmental behavior change.

Past Research on Eco-Feedback

Eco-feedback, particularly in the form of energy feedback, has been studied extensively over the past 40 years, and several reviews of this literature have appeared in recent years. Four of these reviews (Darby, 2006; Ehrhardt-Martinez, Donnelly, & Laitner, 2010; Electric Power Research Institute, 2009; Fischer, 2008) analyzed past empirical studies on energy feedback through qualitative literature review methods, in which a set of empirical studies on a topic are collected, classified, and synthesized (Manten, 1973). Their overall findings were that feedback is effective, with an average of 10% energy savings; effects were found to range from negative (i.e., an increase in energy consumption) to up to 20% in energy savings. In addition to discussing the general effects of energy feedback, these reviews also suggested that the effectiveness of feedback may vary depending on the moderation of variables related to the feedback intervention.

All four reviews discussed immediacy as a moderator of feedback effectiveness. Darby (2006) distinguished feedback primarily as direct and indirect: Direct feedback is available immediately, whereas indirect feedback is processed in some way before it is provided to the consumer (e.g., utility bill). Darby emphasized that the immediacy of information provision is the key variable moderating the effectiveness of feedback, and suggested that direct and immediate feedback may lead to greater savings (5% to 15% for direct/immediate vs. 0% to 10% for indirect). Much of the literature following Darby (2001, 2006) also discusses direct and indirect feedback, extending beyond immediacy to discuss the frequency with which feedback is provided.

Fischer (2008) argued that frequent feedback was more effective than infrequent feedback, because it helps to improve links between actions and consequences. The Electric Power Research Institute (2009), however, concluded that there was very little difference in energy savings from studies using various levels of feedback frequency, with 9% savings for monthly feedback, 8% savings for daily/weekly feedback, and 7% savings for real-time feedback. Likewise, Ehrhardt-Martinez et al. (2010) suggested that real-time feedback results in *lower* conservation efforts (6.9%) than daily/weekly feedback (10.8%).

Additionally, Darby (2006) argues that indirect feedback effectively conveys the effects of behavior on specific energy use (e.g., heating, appliances), and Fischer (2008) argues that the main way of showing a link between action and results is to provide a breakdown that corresponds to individual appliance end-use. Both Fischer (2008) and Electric Power Research Institute (2009) argue that greater energy reductions in studies that provided individual appliance feedback over whole-home feedback. However, because of the nature of existing studies, it is not possible to fully separate this effect from other possible moderators (Ehrhardt-Martinez et al., 2010).

Feedback duration was also highlighted as an important feature in previous reviews (Darby, 2006; Ehrhardt-Martinez et al., 2010; Fischer, 2008), though none of the authors explain why this is the case and results are inconsistent across reviews. Darby (2006) and Fischer (2008) both argue that feedback is more effective when provided over a long time period of more than three months, whereas Ehrhardt-Martinez et al. (2010) concluded that feedback is more effective for shorter (<6 months, 10.1%) rather than longer (>6 months, 7.5%) studies. As such, the findings regarding feedback duration are still unclear.

The use of comparisons was also identified as a key variable for feedback effectiveness. Darby (2006) concluded that providing feedback with comparisons with past use was more effective than a peer group or target figure. Fischer (2008) suggested that comparisons may work by stimulating specific motives for conserving or by providing context within which to interpret usage. However, none of the studies that she analyzed demonstrated an effect caused by social comparisons, and all of the studies provided a historical comparison, so its effect could not be determined. Similar design issues prevented energy measurement (e.g., kilowatt-hours [kWh], cost) from being evaluated, which Fischer proposed as a possible moderator but was unable to analyze.

The combination of feedback with other interventions, such as goal setting, financial incentives, or conservation information, was also hypothesized to increase effectiveness. Darby (2006) stated that a combination of interventions may be more effective than feedback alone. Fischer's (2008) analysis, however, reveals mixed findings; she suggests that these additional interventions may overload users with too much information, and that their impact will be affected by how the information is presented and how appropriate and relevant it is to the audience. As such, there is no current consensus regarding the impact of combined interventions.

Current Study

The current study seeks to integrate FIT (Kluger & DeNisi, 1996) with past research and theory on proenvironmental behavior, through a statistical meta-analysis of eco-feedback. As the most widely studied form of eco-feedback is residential energy feedback (Froehlich et al., 2010), this is an ideal behavioral domain with which to apply and test FIT for proenvironmental behavior.

Study Justification

Although prior literature reviews suggest a positive effect of feedback on residential energy conservation, there are several reasons why further study in the form of a meta-analysis is still needed. First, although qualitative reviews can list and describe findings, results must be interpreted with caution because effect sizes are not calculated, reported effects are not weighted, and inferential tests are not performed to determine whether observed effects are statistically significant across studies (Rosenthal & DiMatteo, 2001). Differences between studies related to research settings, methodology, and characteristics of the feedback provided (i.e., feedback format, type, frequency) were observed in the literature reviews and, in some cases, descriptive statistics (e.g., averages) were provided. However, effects were not analyzed inferentially to make determinations as to whether they *significantly* moderated the effectiveness of the interventions studied. Because both differences in effects and the number of studies that included each level of a variable may be relatively small, especially when compared with overall effect sizes, the techniques of meta-analysis are useful because they estimate the statistical significance of the differences. These key differences in measuring both main effects and moderator effects lead to more reliable conclusions than eyeballing, self-reported findings, or vote counting (Cooper & Hedges, 1994).

Second, the reviews conducted to date present conflicting findings about several feedback moderators. Meta-analysis allows for statistical analysis of both the overall effect of feedback and differences in findings, caused by various moderating variables related to study setting, methodology, and treatment. This approach offers a more nuanced understanding of the overall effectiveness of feedback across studies, as well as different variables with regard to the provision of feedback that may be more or less effective. At this point, an analysis including studies that date back over 40 years can inform not only *whether* feedback is effective overall but also *how* and *for whom* it is most effective. Such comparative analysis is potentially useful for identifying the most promising areas of future research on this important behavioral intervention.

Finally, previous reviews have not integrated psychological theory into their analyses of energy feedback. They present hypotheses and results, but do not integrate the significant contribution of psychology over the past century into understanding the role of feedback in behavior change. Thus, an approach that integrates psychological theory on feedback in general with an empirical analysis of energy feedback studies is both overdue and needed. This work aims to address the identified limitations by drawing on psychological theory to examine the key moderators identified in previous reviews, propose hypotheses for their effect on energy behavior, and test these hypotheses meta-analytically.

Main Effects of Feedback

Both psychological theory and past empirical research suggest that feedback may serve a key role in engaging individuals in residential energy conservation by highlighting and making consumers aware of the otherwise abstract and invisible energy impacts of household behaviors. As such, we hypothesize a significant and positive main effect of feedback on residential energy conservation behavior. However, past literature reviews have also suggested that the effects of feedback may vary based on both how and to whom it is provided. Therefore, we also hypothesize significant variation in treatment effects of feedback across studies and propose an additional set of moderator variables, to be discussed in the following sections.

Treatment Moderators

Analysis of psychological theories of feedback and past reviews of the energy feedback literature identified the following potential moderator variables: frequency, medium, measurement, combination with other interventions, comparison, granularity, and duration. These treatment moderators along with the hypothesized effect on energy behavior are discussed further.

Frequency. The first variable discussed in previous reviews relates to the immediacy, or *frequency*, with which feedback is provided to users. Fischer (2008) suggests that frequent feedback is more effective than infrequent feedback because it helps link actions with consequences. This may be true for immediate feedback in the instance that users are able to refer to the feedback directly after taking action, particularly when their attention is focused on specific action goals during a learning phase (Kluger & DeNisi, 1996). However, it is unlikely to hold true beyond this, because of the multiplicity of energy behaviors accounting for total demand and the sheer number undertaken during the day.

Such a proposition also makes the assumption that feedback will receive attention in the first place; because energy and other environmental behaviors largely hold low personal relevance (Leiserowitz, 2008), this assumption may be invalid. Thus, in the specific context of eco-feedback, it becomes increasingly important to consider the ability of feedback to engage users and draw their attention to the feedback information. Feedback that is provided more frequently do provide more opportunities to engage users' attention; thus, our first hypothesis is that more frequent feedback will be more effective, insomuch as it can engage users' attention and direct it to the appropriate feedback-standard gap for behavior regulation.

Medium. A second and often interlinked variable, which may also be responsible for attracting users' attention toward feedback, is the *medium* through which it is provided. These mediums are categorized according to the commonly used categories (Electric Power Research Institute, 2009) of enhanced billing (e.g., feedback provided via an enhanced utility company bill, such that the feedback was part of the utility bill, but the bill contained more detailed information or feedback than the standard utility bill), card (e.g., door hanger or other card/sign provided to the household by the researchers), monitor (e.g., electronic device or product that provides energy information), or computer (e.g., software or web-enabled program on the subjects' personal home computer).

Although energy bills may do a good job at capturing householders' attention, they are typically received monthly or quarterly;

as such, consumers' attention can drop away in the interim (Allcott, 2011). Cards placed outside of doors and in-home monitors that sit on kitchen counters or other visible places in the home may work better and capture attention as householders go about their daily routines. However, in recent years, the quantity of time spent interacting with computers has increased (Pew Research Center, 2014). Feedback that is computerized or interactive may provide greater opportunities to engage users for longer or more frequent periods of time, suggesting that digitized media may augment feedback effectiveness. Therefore, our second hypothesis is that computerized media will be more effective at stimulating behavior change, insomuch as it enhances the opportunities for users to engage with feedback.

Measurement. A further issue raised by previous reviews of feedback is about mechanisms for stimulating individuals' motivation to conserve energy. The abstract nature and cognitive complexity of energy consumption means that it is viewed as a by-product of daily activities and not something to which blame can be attributed or changes made (Markowitz & Shariff, 2012). However, a number of psychological variables including concern about the environment, price sensitivity, personal and social norms, and efficacy (Barr et al., 2005; Samuelson & Biek, 1991; Verhallen & van Raaij, 1981) were found to predict proenvironmental behavior. As such, we expect feedback that stimulates these attitudes to be more effective at overcoming the blamelessness associated with energy consumption and at motivating changes in demand.

The *measurement* by which feedback information is provided might help to stimulate individuals' motivations to engage, by providing energy measurements in terms of carbon emissions (drawing attention to environmental concern) or in terms of financial spending (tapping into price sensitivity). We hypothesize that feedback presented in these terms will be more effective at engaging people and stimulating behavior change than feedback presented in terms of kWh.

Combination with other interventions. Previous reviews have also suggested that the combination of feedback with other interventions, such as financial incentives and goal setting, should increase effectiveness. Both of these methods might provide additional motivation, by either stimulating price sensitivity or providing a personal norm for comparison. In fact, the generation of a standard to which feedback can be compared lies at the heart of Kluger and DeNisi's (1996) FIT; it is an individual's desire to decrease the gap between their desired level of consumption (the goal they have set) and their current consumption that mediates the effectiveness of feedback. Thus, we hypothesize that the combination of feedback with such other interventions will positively impact its effectiveness.

Comparison. The presence of a *comparison* refers back to the concept in FIT of a feedback-standard gap. In this case, the standard refers to a comparison with a goal, past usage, or peer usage. The provision of historical usage data may provide households with a personal norm against which to directly measure and compare current levels of consumption, thereby engaging users in attending to their consumption behaviors. Overlaying this with an explicit comparison of current performance with a predefined goal may serve to provide a larger feedback-standard gap, motivating greater reductions in energy use.

Additionally, research applied from social comparison theory (Festinger, 1954) discusses the role of normative framing and social comparisons in motivating reductions in energy use. This approach involves providing households with consumption data from others—friends, neighbors, or a community—for comparison. Drawing on the focus theory of normative conduct, several recent studies have shown that normative feedback can lead to significant reductions in energy use (Aitken, McMahon, Wearing, & Finlayson, 1994; Nolan, Schultz, Cialdini, Goldstein, & Griskevicius, 2008; P. W. Schultz, Nolan, Cialdini, Goldstein, & Griskevicius, 2007; W. P. Schultz, Khazian, & Zaleski, 2008). The presence of a comparison (historic, goal based, or social) is thus hypothesized to positively moderate the effectiveness of energy feedback.

However, further reference to FIT (Kluger & DeNisi, 1996), particularly the discussion around the hierarchical organization of feedback-standard gaps, suggests that not all comparisons are equal and that the type of comparison used may direct attention to different levels of the hierarchy. Kluger and DeNisi (1996) found that feedback directed toward the metatask level serves to attenuate feedback effects on performance, whereas those directed toward lower levels augment feedback effects on performance. Thinking about the effects of energy feedback on performance, this suggests that feedback at the metatask level (e.g., focusing on identity or values) will be less effective than feedback directed at the focal task level (e.g., reducing energy consumption), or at task learning processes. As such, *comparison type* is an important moderator hypothesized to affect feedback effectiveness. Goal comparisons are considered the most effective, as they focus attention on focal task goals. Feedback that utilizes normative comparisons may drive users' attention toward metatask processes, attenuating the effects of feedback. As such, we hypothesize that goal comparisons will be more effective than social norm comparisons.

Granularity. It is important that feedback information, once provided, can lead to a specific behavioral response. Individuals need to be able to identify at least one action they can engage in that they associate with the feedback—in this case, decreased energy use. *Granularity* defines the level of detail represented in the feedback. Studies were coded as providing either whole-home feedback data or data disaggregated by appliance use. Because energy use can be addressed by a multiplicity of behaviors, and the links between such behaviors and their energy impact is generally not well understood, it is particularly important that the feedback helps to provide links between energy use and specific actions or behaviors that reduce energy use. In this way, people can start to see how both everyday actions lead to consumption and how specific behavior changes may help them to reduce feedback-standard gaps. Disaggregated feedback may help with task learning processes, and is hypothesized to positively moderate the effects of energy feedback. Therefore, feedback that is appliance-specific is hypothesized to be more effective than whole-home feedback because it connects energy usage directly to specific actions.

Duration. The final treatment variable discussed in previous reviews is the *duration* over which feedback is provided. Over time, users' attention may shift as they move from initial task learning to higher level self-salient goal satisfaction; thus, the duration over which feedback is provided may impact how the feedback message is interpreted and where the users' attention is subsequently directed. When users are engaged with task learning, feedback that enables them to identify links between actions and consequences may be more successful at increasing knowledge gains and shifting attention back toward the feedback-standard gap. Once learning has taken place, the type of feedback used for this learning process has served its purpose and users may begin to disengage with it (Karlin, 2011). If users experience success in reducing the feedback-standard gap, or if the feedback directs attention toward higher level motivations and self-related processes, affective processes may be triggered and users may look for opportunities to obtain other personal goals.

Although this reallocation of cognitive resources may result in a short-term performance reduction, long-term performance would be expected to improve as users become more familiar with eco-feedback and behavioral responses become more automatic. Initially, the expectation is that users will engage with the feedback they are receiving because it is novel and interesting, and tasks of short duration will have greater effectiveness than is observed with tasks over the medium term. However, as users' attention moves along different levels of the hierarchy over time, their performance should improve as cognitive demand decreases. As such, feedback is hypothesized to be most effective in the short term, when task learning is most likely to occur, and over the long term as behavioral regulation becomes more automatic and new goals are set as previous ones are met.

Study Quality and Publication Bias

Although our primary goal was to determine variations in energy feedback that moderate its effects on conservation behavior, variables related to study design may also moderate results and therefore are recommended for inclusion in meta-analysis (Stock, 1994). Examining methodological variables can inform us about the extent to which this intervention is robust (Cooper & Hedges, 1994), and can also be informative to future researchers as they make decisions about setting and methodology in their own studies. The inclusion criteria ensured that the studies included in this analysis passed at least a minimum standard of quality, by excluding studies that did not have a control group as well as those with clear confounding variables. Additional study characteristics were identified to test for any bias that could result from threats to validity. Therefore, the following five study characteristics were also coded and analyzed: sampling strategy, response rate, random assignment, baseline data collection, blind control group, and empty control group.

Sampling strategy. *Sampling strategy* refers to the way that subjects were recruited to participate; if samples were recruited by convenience rather than systematically (e.g., whole population or random sample), this could threaten external validity. *Response rate* refers to the percentage of those contacted who elected to participate in the study; a lower response rate could suggest self-selection bias among participants, potentially inflating effects. *Random assignment* refers to whether participants were randomly assigned to treatment conditions; an absence of random assignment could threaten internal validity. *Baseline data* refers to the collection of energy use information before the beginning of treatment in order to establish a baseline to compare treatment energy use; collecting baseline data controls for the threat of history; therefore, a failure to do so could introduce bias.

Control group. The type of *control group* used was also examined, as comparing a blind control with an active treatment group could result in a Type I error caused by the Hawthorne effect, in which awareness of being in the study (rather than the proposed intervention) affects participants' responses. Also, in some feedback studies, the control groups were not completely neutral; some studies also used *information-only* as a control group instead of an empty control group. Seventeen studies included conditions in which information was provided to subjects without feedback, and in seven of those, the information-only group served as the only control group for the study. As such, we included all of these types of control groups in the main effects analysis and tested both blind versus aware, and information versus empty, control groups as study characteristic variables.

Publication bias. Finally, two variables were included to test for publication bias: *publication type* and *sample size*. Publication type was tested because it was found that published studies have larger effect sizes than unpublished studies (Smith, 1980), and that studies with smaller effect sizes tend to take longer to get published (Rosenthal, 1991). Sample size is another variable that can be analyzed to test for publication bias. Studies with fewer participants have a greater likelihood of sampling error (Shadish & Haddock, 1994), but this error should be equally distributed among larger than average and smaller than average effect sizes, especially when an effort is made to include unpublished studies. However, studies with both a small effect size and a small sample size may be less likely to get published and circulated; even though a great effort was made to obtain unpublished studies, one cannot completely avoid the problem of unsuccessful studies being hidden away in file drawers.

Method

The purpose of the present study was twofold: (a) to estimate the overall effect size of energy feedback on energy conservation using all available studies, to evaluate the precision of this effect size estimate by the confidence interval around the estimate, and to subject the obtained effect size to null hypothesis significance testing; and (b) to examine the potential impact of treatment and study variations using moderator analysis of the aforementioned variables. As such, the following methods were employed.

Literature Search

Following the procedures and guidelines suggested by Cooper (2010) and Rothstein and Hopewell (2009), the following six methods were used to locate relevant studies: a keyword search in reference databases, a conference program search, a backward search, a forward search, e-mails to study authors, and personal contacts. This search included articles published between 1976 (the year the first identified study was published) and 2010. The original source and inspiration for this study was the Darby (2006) literature review on feedback and energy conservation. An examination of the reference list of this review identified 28 relevant articles.

Next, keyword searches were conducted in PsycINFO, JSTOR, Web of Science, PubMed, and Google Scholar using the keywords *energy conservation* (or *energy efficiency* or *energy use* or *energy savings* or *energy consumption*) and *feedback* simultaneously. This returned 27 relevant results, including two additional review articles (Abrahamse, Steg, Vlek, & Rothengatter, 2005; Fischer, 2008). Because of the nature of this research, governments, utilities, and private firms have also performed studies, many of which do not appear as academic publications. Therefore, a general Google search was also performed using the same keywords, resulting in five additional studies.

Searches were also conducted of the proceedings for the European Council for an Energy Efficient Economy (ECEEE) conference, the American Council for an Energy Efficient Economy (ACEEE) conference, and the programs of the Behavior Energy and Climate Change (BECC) and Home Energy Display (HED) conferences, which are considered the leading conferences in this field. Eleven new articles were found using this method.

Next, backward searches were performed on all articles that were identified as either an empirical study or a review of energy feedback. In the backward searches, the reference sections of selected articles were reviewed for additional potential studies. Forty-seven articles were identified by this method. In particular, the reference sections of the following three review articles included new and useful references: Abrahamse et al., 2005 (14 articles), Ehrhardt-Martinez et al., 2010 (nine articles), and Fischer, 2008 (six articles).

In addition to the backward searches, forward searches were conducted on the five primary literature reviews of energy feedback conducted to date (Abrahamse et al., 2005; Darby, 2006; Ehrhardt-Martinez et al., 2010; Electric Power Research Institute, 2009; Fischer, 2008). This search method utilized Web of Science and Google Scholar to identify articles that have cited these review articles since their publication. Nine articles were located through this method.

At this point, the preliminary list of 127 potential feedback studies was compiled and sent to the corresponding authors of all of the identified studies and literature reviews for which contact information was available. The e-mail request asked for any additional published and/or unpublished articles, information on relevant studies, or active researchers in the field. Thirty-one articles were identified using this method. Three active researchers in this area (Sarah Darby, Corinna Fischer, and Wesley Schultz) were especially helpful at this stage.

Finally, informal inquiries via e-mail and discussion were conducted with personal contacts including university colleagues, researchers at three energy-related conferences, and the demand–response manager of a local electricity provider. This identified 14 additional articles, bringing the total number of articles initially compiled and reviewed to 172.

Inclusion Criteria

Of the 172 articles originally collected, 69 were identified as review articles or unrelated research articles and set aside for reference. The remaining 103 were identified as empirical studies on energy feedback, and were examined independently by the first and second author for inclusion in the meta-analysis. Discrepancies regarding the inclusion of a particular article were resolved by discussion. To be included in the meta-analysis, a study had to meet the following criteria (the number of studies excluded because of each criterion is in parentheses):

1. The study must have been conducted using an experimental design. Case study, survey data, and purely qualitative studies were excluded (5).

2. The study must have been conducted as a naturalistic field study measuring actual energy use in the home. Studies that were conducted in a lab or office setting were excluded (7).

3. The study must have used the quantity of household energy use (appliance-specific or overall/household energy usage) for its dependent variable. Studies that measured only load-shifting behavior (from peak to nonpeak hours) were excluded (9).

4. The study must have used feedback as the independent variable; it had to include at least one group in which feedback was provided and could be isolated for analysis (9).

5. The study must have included a control group that did not receive feedback. Control participants could be an empty control (e.g., no treatment) or a have received a nonfeedback intervention (e.g., information) (16).

6. The study must have presented statistics or reported a statistically significant (or nonsignificant) effect for feedback. Studies that reported group means but did not conduct inferential tests (or failed to provide the standard deviations or standard errors that would allow us to conduct inferential tests) were excluded. (5).

Altogether, 51 articles were excluded according to these criteria, with the remaining 52 included for analysis. Of these, 13 articles were recognized as reports of overlapping data (e.g., two or more articles reporting on the same data set); these articles were grouped together and given the same study code. Conversely, multiple studies from the same article were coded and analyzed separately if different samples were used, as was the case in four of the articles reviewed (three that included two studies, and one that included three studies). A total of 42 independent studies from 52 research articles and reports were included and coded in this meta-analysis.

Coding Procedure

A detailed coding sheet was developed based on the established guidelines for meta-analysis (D. B. Wilson, 2009); each study was coded according to the same criteria. For each study, the following information was extracted and coded:

1. Report identification: publication year, author(s), publication type, funding.

2. Study setting: study year, location, population, home type, sample size.

3. Study participants: demographics, housing characteristics.

4. Methodology: recruitment, assignment, data collection.

5. Treatment: frequency, medium, measurement, combination with other intervention, comparison, granularity, and duration.

6. Statistics: cell means and standard deviations, inferential statistics.

In some cases, the information for a particular study either was not obtainable from the study report (e.g., total number of subjects contacted) or was somewhat ambiguous (e.g., random assignment); therefore, not every variable could be coded for all studies. When information was missing in a study and there was no clue available to support a reasonable estimate, the information was coded as missing data. All study variables were coded by the first author and, because the coding process involved some degree of subjectivity, an additional author coded 12 randomly selected studies (28%) to establish reliability. Interrater reliability was acceptably high ($\kappa > .700$) for all variables.

Calculating Effect Sizes

Because the included studies measured and analyzed variables in several different ways that do not permit direct comparison, all study results were converted into correlation coefficients (r). Because effect size represents the degree to which the tested intervention (e.g., feedback) resulted in a reduction in energy use, a *positive* effect size indicates that feedback resulted in *decreased* energy use (compared with the control), and a *negative* effect size indicates that feedback resulted in *increased* energy use (compared with the control); an effect size of zero indicates that the feedback had no statistically significant effect on energy use. Although the specific feedback intervention in each study was slightly different, and the measurement of the dependent variable varied by frequency (daily, weekly, monthly) and style (meter read, self-report), a correlation coefficient (r) was calculated for each study, and these methodological differences were later analyzed as moderators.

Conversions to correlation coefficients were calculated according to established guidelines and procedures of meta-analysis (Rosenthal, 1991; Rosenthal & Rubin, 2003). In some cases, the study report indicated that a focused test had been conducted (e.g., t test, F test with 1 degree of freedom in the numerator); however, rather than reporting statistical data, it stated only whether the test was significant or nonsignificant. In these cases, if the result was reported as significant, the p value was assumed to be one decimal place smaller than the alpha value (e.g., assumed to be .049 if the test was significant at the .05 level), and the correlation coefficient was calculated according to the procedures described by Rosenthal and Rubin (2003). For two results that were reported only as nonsignificant, the effect size was assumed to be zero. Although this procedure has been widely recommended (e.g., Rosenthal, 1991), modern treatments (e.g., Pigott, 2009) are more cautious. We therefore ran the overall meta-analysis both with and without these two imputed studies, and demonstrated that the results were consistent.

Because it was predicted that feedback would have a positive effect (e.g., feedback groups would decrease energy use more than control), all p values calculated were one-tailed (unless otherwise noted). The first and second author independently calculated effect sizes for all included studies, and discrepancies were resolved through discussion.

Significance Testing

Once the effect size estimates were calculated for the individual studies, both a random effects and fixed effects approach to significance testing of the effect sizes were conducted. Fixed effects analyses treat the participants in each study as the unit of analysis and are typically used when a relatively small number of studies are available (Borenstein, Hedges, Higgins, & Rothstein, 2009). Fixed effects analyses assume that each study has the same underlying effect size, and that observed differences in effect sizes are caused by sampling biases. Fixed effects meta-analysis uses all of the studies together to estimate the effect size. However, the assumption that all studies have the same effect size may be unrealistic, which limits the ability to generalize the results of a fixed effect meta-analysis to additional or future studies. Random effects analysis assumes that the effect sizes of each study come from a distribution of possible effect sizes. Random effects analysis usually results in decreased statistical power, but allows broader generalizability to studies that are not included in the analysis (Field, 2001; Hedges & Vevea, 1998; Hunter & Schmidt, 2000).

The present study includes both analyses to determine whether the effects are robust under a wide range of methodological assumptions; the fixed effects approach was computed to consider a scenario in which the effect sizes of each study are similar, and the random effects approach was computed to increase the generalizability of the findings. Both the fixed and random effects analyses were implemented using the metaphor R software package (Version 1.9-4, R Core Team, 2012). The fixed effect model used standard inverse variance weighting and the random effects model utilized restricted maximum likelihood (REML) to estimate the between study heterogeneity. Effects were considered significant when the p value was less than .05.

Moderator Analysis

In addition to analyzing the overall effect of feedback on energy conservation, moderator analyses were conducted to examine which variables may moderate the effects of feedback on energy conservation. Values for each variable, extracted from each study report in the coding sheet (e.g., feedback duration, energy granularity), were obtained, and moderator analyses were conducted using a mixed effects model in which the moderators were included in the random effects meta-analysis statistical model. The model was fit using REML.

Results

Main Effects of Feedback

A main effect size for feedback on energy conservation was calculated for each of the 42 studies. Main effects were calculated by comparing all of the feedback conditions with all of the control conditions in each study. The total number of participants across the 42 studies was 256,536, with a median of 119 participants per study. A list of all studies included in the meta-analysis is provided in Table 1, along with each study's sample size, reported percent savings, *r* and associated statistical significance, and values for each treatment moderator. A forest plot of the 42 main effect sizes and confidence intervals is shown in Figure 1.

Effect sizes for the main effect of feedback ranged from −.0803 to .4803, and 21 (50%) of the effect sizes were statistically significant at the $p < .05$ level. When taken together, the 42 studies had an unweighted mean correlation coefficient of .1174. However, this estimate does not account for the variability of the sizes of the studies, nor does it take into account the possibility of between study effect size variance. Therefore, we conducted both a fixed effect and random effect analysis.

The fixed effects analysis obtained a mean effect size of .0397 (95% confidence interval [CI] [.0359, .0436]; $z = 20.163$, $p < .0001$), and the random effects analysis obtained a mean effect size of .0712 (95% CI [.0454, .0969]; $z = 5.4148$, $p < .0001$). Both of these analyses, although somewhat smaller than the unweighted mean correlation coefficient, still suggest that feedback interventions in general do significantly decrease residential energy use. We set the effect size of two studies that reported no effect to zero. In order to verify that this did not introduce a bias in the random effects analysis, we also ran the analysis without these two studies and obtained a mean effect size of 0.0754 (95% CI [0.0484, 0.1023]; $z = 5.4839$, $p < .0001$), which is consistent with our previous analyses and suggests that setting these effect sizes to zero did not bias the analysis.

Additionally, a high level of variability was found within the individual effect sizes; five studies (12%) had a negative effect size, two studies (5%) were estimated to have an effect size of zero, and 35 studies (83%) had a positive effect size. A statistical test of the heterogeneity, using the Cochran's Q test, among the effects was significant ($Q = 470.960$, $p < .001$), corresponding to an $I^2 = 91.40\%$, which signifies substantial heterogeneity. As such, the findings suggest that the effect of feedback on energy conservation may vary based on variables related to the study setting, quality, methodology, and treatment. Thus, the findings justify additional analyses to identify which specific variables may moderate this effect.

Moderator Analysis

Although analyses revealed a significant positive effect for feedback, the 42 studies analyzed contain a great deal of variation in terms of the study setting, methodological approach, and treatment provided to participants. Therefore, a series of moderator analyses were performed to better understand when, how, and to whom feedback is most effective. All of the proposed variables introduced were examined as potential moderators of the overall effect of feedback on energy conservation. Descriptions of analyses for each moderator variable are provided in the sections below; Table 2 presents the results of all moderator analyses.

Study quality variables and publication bias. First, we tested for any biases in the study results that were caused by study characteristics or publication bias. The following study characteristics were analyzed: sampling strategy, response rate, random assignment, baseline data, and two variables for control group—whether they were blind or aware of the study (testing for Hawthorne effects) and whether they received no contact from the researcher ("empty control") or served as an "information-only" control. We examined each of the six study characteristic variables in relation to feedback effect size; there were no significant relationships between any of the study characteristics and feedback effectiveness. This provides confidence that these qualities, which are theoretically extraneous to the effect of feedback on energy conservation, are not directly influencing the overall average effect size.

Additionally, two variables were tested for publication bias: publication type and sample size. Differences for publication type were not significant, suggesting no bias according to publication type ($p = .477$). Sixteen studies contained more than 300 individuals, and we performed moderator analysis comparing these studies with the remaining studies with less than 300 individuals, and study size revealed a significant negative relationship ($p < .007$); studies with larger samples had smaller effect sizes than those with smaller samples.

The finding that larger studies had significantly smaller effect sizes than smaller studies could suggest a biased sample—one that is missing studies that had both a small effect size and a small sample size. Therefore, a second analysis was conducted to assess whether this effect represents a *file drawer* bias, which asks the question, If it was possible to access all of the unsuccessful studies hiding away in file drawers, would the effect for feedback no longer be significant? To help answer this question, we performed a trim and fill analysis to estimate the effect of the studies that are missing from the literature because of publication bias. In the analysis, we estimated that there are 10 missing studies; when incorporating these studies, as expected, our random effects model effect size decreased from .071 to .053 (95% CI [.026, .080]). However, the value is still significant ($p < .0001$), suggesting that the reported mean effect is not simply an artifact of publication bias.

Treatment variables. Finally, the hypothesized treatment variables were tested for moderation of feedback effectiveness. The following variables that described differences in the way that feedback was provided were tested: (a) frequency, (b) medium, (c) measurement, (d) combination with other intervention, (e) comparison, (f) granularity, and (g) duration.

Feedback frequency was categorized as monthly or less (eight studies), 1 to 4 times per week (eight studies), daily (four studies), or continuous (seventeen studies); five studies could not be categorized because frequency was mixed. No significant differences in effect size based on feedback frequency was observed ($p < .624$). Although this result was initially surprising, we observe that feedback frequency is confounded by several of the other moderators, such as feedback medium, which may explain the results.

Feedback medium was categorized as enhanced billing (bill or home energy report sent by utility company with more detailed information than the standard utility bill; five studies); card (door hanger, sign, or other card provided by the researchers; 15 studies); monitor (electronic device or product that provides energy information; 16 studies), or computer (software or web-enabled program on the subjects' personal home computer; two studies). Four studies could not be categorized because the medium was mixed. Studies with enhanced billing feedback had the lowest effect size, followed by monitor, card, and, finally, computer, which had the highest effect size. An overall comparison of these groups was not significant ($p = .316$), but feedback provided via computer was

Table 1
Main Effects of Feedback and Treatment Moderators

Author & Year of publication	*n*	Reported % saving	*r*	*p*	Frequency	Medium	Measurement	Comparison	Combination	Granularity	Duration
Alcott (2010)	78,492	2.4%	.0096	.0036	Monthly or less	Bill	kWh & Cost	Historical	None	Whole-home	6–12 months
Allen & Janda (2006)	60	—	.0007	.4980	Continuous	Monitor	kWh & Cost	None	None	Whole-home	1–3 months
Arvola (1993, 1996a, 1996b); Arvola et al. (1994)	696	2.9%	.1018	.0036	Monthly or less	Card	Mixed	Mixed	None	Whole-home	>12 months
Ayres et al. (2013)	84,000	1.2%	.0091	.0045	Monthly or less	Bill	kWh & Cost	Mixed	None	Whole-home	6–12 months
Battalio et al. (1979); Winett et al. (1978)	70	.9%	.0303	.4017	1–4 times/week	Card	kWh	Historical	None	Whole-home	<1 month
Becker (1978); Seligman et al. (1978) Study 2	80	13.0%	.3094	.0022	1–4 times/week	Card	Goal only	Goal	Goal	Whole-home	<1 month
Becker & Seligman (1978); Seligman et al. (1978) Study 3	20	15.7%	.1899	.2113	1–4 times/week	Card	kWh	Mixed	None	Whole-home	<1 month
Bittle et al. (1979–1980)	353	—	−.0164	.3794	Daily	Card	kWh & Cost	None	None	Whole-home	<1 month
Bittle et al. (1979)	30	—	.0366	.4212	Daily	Card	kWh & Cost	None	None	Whole-home	1–3 months
Brandon & Lewis (1999)	120	—	.1602	.0403	Mixed	Mixed	Mixed	Mixed	None	Whole-home	6–12 months
Dobson & Griffin (1992)	100	12.9%	.1968	.0243	Continuous	Computer	kWh & Cost	None	None	Appliance	1–3 months
Haakana et al. (1997)	755	19.0%	.0715	.0245	Monthly or less	Card	kWh & Cost	Mixed	None	Appliance	>12 months
Harrigan (1992); Harrigan & Gregory (1994)	71	.0%	.0000	.5000	Continuous	Monitor	kWh	None	None	Whole-home	6–12 months
Hayes & Cone (1981)	40	7.0%	.0427	.3968	Monthly or less	Card	kWh & Cost	Historical	None	Whole-home	3–6 months
Hutton et al. (1986) Study 1	371	4.1%	.1369	.0042	Continuous	Monitor	Cost	None	None	Whole-home	3–6 months
Hutton et al. (1986) Study 2	377	5.0%	.1387	.0035	Continuous	Monitor	Cost	None	None	Whole-home	3–6 months
Hutton et al. (1986) Study 3	336	6.8%	.0235	.3340	Continuous	Monitor	Cost	None	None	Whole-home	3–6 months
Kasulis et al. (1981)	390	—	.0461	.1822	Monthly or less	Bill	kWh & Cost	None	Pricing	Whole-home	3–6 months
Katzev et al. (1980–1981)	22	15.0%	.1508	.2525	Mixed	Card	kWh & Cost	Mixed	Other	Whole-home	<1 month
Kurz et al. (2005)	423	.0%	.0000	.5000	1–4 times/week	Card	kWh	Social	None	Whole-home	3–6 months
Mansouri & Newborough (1999); Wood & Newborough (2003)	31	20.0%	.2567	.0817	Continuous	Monitor	kWh	Historical	None	Appliance	1–3 months
Matsukawa (2004)	319	1.8%	−.0266	.3180	Continuous	Monitor	kWh	Historical	None	Whole-home	3–6 months
McClelland & Cook (1979–1980)	101	12.0%	.1535	.0637	Continuous	Monitor	Cost	None	None	Whole-home	6–12 months
Midden et al. (1983)	95	13.2%	.2148	.0173	1–4 times/week	Card	kWh & Cost	Social	Incentive	Whole-home	1–3 months
Mountain (2007) Study 1	118	18.1%	.1816	.0245	Continuous	Monitor	kWh & Cost	None	None	Whole-home	>12 months
Mountain (2007) Study 2	110	2.7%	.1882	.0245	Continuous	Monitor	kWh & Cost	None	None	Whole-home	>12 months
Mountain Economic Consulting and Associates (2006)	552	6.5%	.0838	.0245	Continuous	Monitor	kWh & Cost	None	None	Whole-home	>12 months
Nexus Energy Software (2006)	249	19.0%	.1420	.0125	Mixed	Mixed	kWh & Cost	Mixed	Goal	Whole-home	3–6 months
Pallak & Cummings (1976); Pallak et al. (1980)	109	16.0%	.2538	.0039	1–4 times/week	Monitor	kWh	None	Commitment	Whole-home	1–3 months
Parker et al. (2008)	17	7.0%	.4803	.0219	Continuous	Monitor	kWh & Cost	Historical	None	Whole-home	>12 months
Robinson (2007)	141	—	−.0830	.1640	1–4 times/week	Mixed	Mixed	Mixed	None	Whole-home	3–6 months
Seaver & Patterson (1976)	75	—	.0617	.2971	Mixed	Card	kWh & Cost	Historical	None	Whole-home	3–6 months

Table 1 (*continued*)

Author & Year of publication	*n*	Reported % saving	*r*	*p*	Frequency	Medium	Measurement	Comparison	Combination	Granularity	Duration
Seligman & Darley (1977); Seligman et al. (1978) Study 1	29	10.5%	.4317	.0199	Daily	Card	Goal only	Goal	None	Whole-home	<1 month
Sexton et al. (1987); Sexton et al. (1989); Sexton & Sexton (1987)	269	—	−.0803	.0946	Continuous	Monitor	Cost	None	Pricing	Whole-home	6–12 months
Sipe & Castor (2009) Study 1	305	—	.0702	.1108	Continuous	Monitor	kWh & Cost	None	None	Whole-home	6–12 months
Sipe & Castor (2009) Study 2	588	—	−.0156	.3529	Continuous	Monitor	kWh & Cost	None	None	Whole-home	6–12 months
Summit Blue Consulting (2009)	85,000	—	.0941	.0001	Monthly or less	Bill	kWh & Cost	Mixed	None	Whole-home	6–12 months
Ueno et al. (2005); Ueno et al. (2006)	19	12.0%	.4099	.0407	Continuous	Computer	kWh & Cost	Mixed	None	Appliance	6–12 months
van Houwelingen & Van Raaij (1989)	235	6.2%	.1206	.0325	Mixed	Mixed	Mixed	Mixed	Goal	Whole-home	6–12 months
Wilhite & Ling (1995)	1,284	10.0%	.0549	.0245	Monthly or less	Bill	Mixed	Mixed	None	Whole-home	>12 months
Winett et al. (1982) Study 1	49	—	.1598	.1364	Daily	Card	kWh & Cost	Goal	None	Whole-home	1–3 months
Winett et al. (1982) Study 2	35	—	.0202	.4541	1–4 times/week	Card	kWh & Cost	Goal	None	Whole-home	<1 month
Unweighted average *r*			.1174								
Weighted average *r*			.0396								
Fixed effects *p* value				<.001							
Random effects *p* value				<.001							
Reported % savings		9.0%									

Note. A dash (–) indicates "not reported"; kWh = kilowatt-hours.

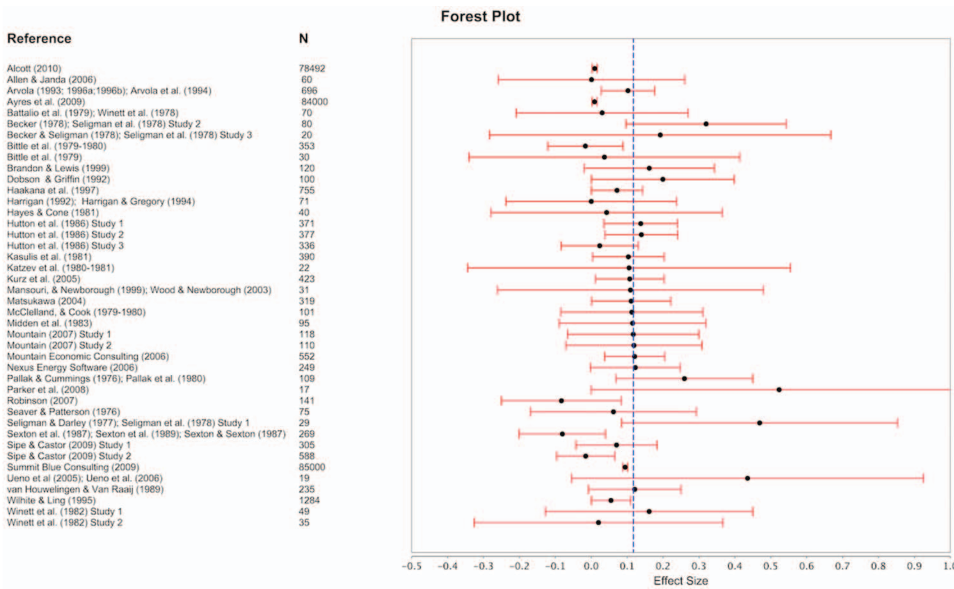

Figure 1. A forest plot of study effect sizes. See the online article for the color version of this figure.

marginally more effective than feedback provided via any of the other media combined ($p = .083$).

Energy measurement was coded as cost in five studies, as kWh in seven studies, and as kWh and cost combined in 23 studies. Five studies could not be categorized because their energy measurement was mixed, and an additional two did not provide an energy measurement (goal only). Analyses indicated no significant differences among these three energy measurement groups ($p = .899$). In addition, three studies that combined environmental information with cost or energy measurements were compared with the 32 studies that did not, but no significant difference was found ($p = .150$).

The next treatment variable tested was combination with other intervention strategies. Three studies were identified in which feedback was combined with a goal intervention, such that this Feedback + Goal Combo intervention was compared with a control group. Two studies were identified in which feedback was combined with an incentive intervention, such that this Feedback + Incentive Combo intervention was compared with a control group. Effect sizes for these *combined* interventions were compared with the remaining 37 feedback-only effect sizes. The effect of combining interventions was significant ($p = .037$); the Feedback + Goal Combo interventions and Feedback + Incentive Combo interventions both had higher effect sizes than studies that used feedback alone.

Comparison (e.g., historical, social, goal) was analyzed in two ways. First, the 19 studies whose feedback had a comparison was tested against the 17 studies that did not have a comparison (six studies could not be categorized because the comparison message was mixed); the overall effect of having a comparison was not significant ($p = .907$). Six studies could not be categorized because the comparison message was mixed. Second, different comparison types were examined among the studies that did have comparisons; the four studies with goal comparisons had the highest average effect size, followed by the two studies with social comparison, further followed by seven studies with historical comparison. Twelve studies could not be coded because comparison type was mixed. Although the overall effect of comparison type was not significant ($p = .199$), the effect of goal only comparisons was significant ($p = .016$).

Energy granularity was coded as whole-home (38 studies) or disaggregated by appliance or use (four studies). Energy granularity was not found to be a significant moderator of feedback effectiveness ($p = .255$). This is not unexpected because of the very small number of studies in our analysis, even though studies that provided disaggregated feedback had a higher effect size than the ones that provided whole-home feedback.

Finally, feedback duration was categorized as less than a month (seven studies), 1 to 3 months (seven studies), 3 to 6 months (10 studies), 6 to 12 months (11 studies), or more than 1 year (seven studies); the effects of duration on feedback were marginally significant ($p < .756$). We noticed an apparent increase in effectiveness in studies that are longer than 3 months compared with studies that are up to 1 year, and a decrease in effectiveness of studies that are longer than 1 year, as seen in Figure 2. We applied moderator analysis to evaluate each of these observations and found that the difference in effect of studies of less than 3 months compared with up to 1 year was significant ($p < .041$), and the difference in effect size between studies of less than 1 year and longer than 1 year was also marginally significant ($p < .094$).

Discussion

As feedback technologies become increasingly ubiquitous in our society and the capacity to leverage personalized energy information grows, there is an urgency to ensure that these technologies are utilized to their full potential. Although there is much research addressing *whether* feedback works, there has been little research into the more nuanced questions of *how* it works best. The current study integrated 42 empirical analyses in an attempt to assess the

Table 2
Moderator Analyses

Variable	Grouping: Group	k	Mean r	Moderator analysis: Q	p	Lower limit	Upper limit
Treatment variables							
Feedback frequency	Monthly or less	8	.052	2.610	.624	−.048	.152
	1–4 times/week	8	.090			−.029	.209
	Daily	4	.064			−.084	.211
	Continuous	17	.075			−.026	.175
	Unknown	5	.129			.038	.219
Feedback medium	Bill	5	.042	4.730	.316	−.058	.142
	Card	15	.084			−.019	.186
	Computer	2	.238			.024	.452
	Monitor	2	.074			−.024	.172
Energy measurement	Cost only	5	.071	.212	.899	.001	.140
	kWh only	7	.049			−.057	.154
	kWh and cost	23	.067			−.010	.145
Environmental info	Environmental info	3	−.018	2.070	.150	−.133	.097
	No environmental info	32	.054			−.044	.151
Comparison message	No comparison	17	.072	.014	.907	.026	.116
	Comparison message	19	.076			.014	.137
Comparison	Goal comparison	4	.244	7.310	.199	.168	.320
	Social comparison	2	.060			−.037	.157
	Historical comparison	7	.030			−.054	.114
Comparison goal	Goal comparison	4	.195	5.770	.016	.085	.306
	No goal comparison	32	.060			.032	.089
Combined intervention	Feedback only	34	.062	6.013	.049	−.037	.159
	Feedback + Goal	3	.163			.068	.258
	Feedback + Incentive	2	.203			−.012	.417
Energy granularity	Appliance-specific	4	.129	2.732	.256	−.013	.267
	Whole home	38	.072			−.032	.176
Feedback duration	<1 months	7	.095	8.475	.076	.001	.188
	1–3 months	7	.178			.044	.313
	3–6 months	10	.054			−.051	.159
	6–12 months	11	.042			−.059	.143
	>12 months	7	.096			−.011	.203
Study quality							
Sampling strategy	Convenience	3	.033	.070	.790	−.037	.210
	Systematic	35	.069			−.057	.196
Response rate	Below 50%	12	.073	.180	.660	.035	.112
	50% or higher	16	.059			−.004	.123
Random assignment	Not random	7	.076	−.686	.246	.012	.141
	Random	35	.071			.001	.142
Baseline	No baseline	7	.079	.528	.767	.012	.145
	Beginning of study	13	.052			−.039	.143
	Previous year	22	.077			.003	.152
Control group – aware	Blind	11	.085	1.177	.277	.046	.124
	Aware	20	.054			−.001	.109
Control group – info	Information only	6	.056	.179	.672	.031	.089
	Empty	25	.078			−.005	.160
Publication bias							
Publication type	Journal	24	.077	2.49	.477	.008	.150
	Conference	8	.077			.016	.139
	Report	9	.072			−.006	.150
	Thesis	1	−.083			−.289	.123
Sample size	<300	26	.123	7.403	.007	.069	.176
	>300	16	.048			.020	.076

Note. kWh = kilowatt-hours; info = information.

overall effectiveness of feedback, as well as specific variables that moderate this effect. Research suggestions are also provided for the most promising directions of future work.

Review of Findings

Analyses were conducted to test the main effect of feedback on energy conservation, and to assess the moderating role of several internal and external variables that have been hypothesized to interact with the effectiveness of feedback.

Main effects of feedback. As suggested by previous qualitative literature reviews on eco-feedback, the main effect of feedback on energy conservation across all 42 studies was significant. Although the mean effect size was .115, which supports previous literature reviews that reported average savings of 8% to 12%, the

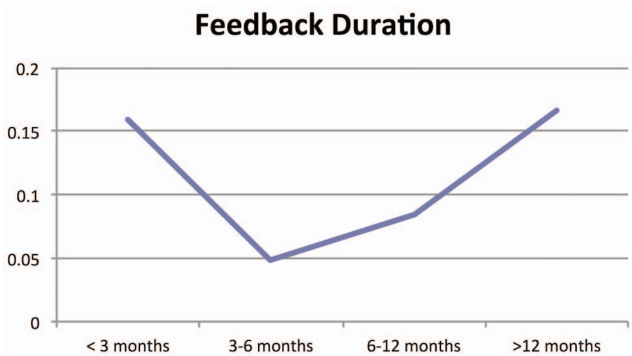

Figure 2. Feedback duration (*x*-axis) as moderator of feedback effectiveness (*y*-axis). See the online article for the color version of this figure.

simple mean effect size does not take into account study size or other causes of variance in the effect size estimates. When considering a random effects model, we found a mean effect size of .071, which is a more conservative estimate.

Although feedback was found to be effective, the significant heterogeneity in effects among studies justified further analysis into moderating variables related to treatment, study setting, methodology, and publication. These findings provide empirical support for the role of feedback in energy conservation, serve to clarify the direction and magnitude of the moderating variables discussed in previous literature reviews, and suggest directions and opportunities for future programs.

Treatment moderators. A number of moderators of feedback were identified in this analysis. It is important to note the limitations of moderator analyses of this type; because we did not randomly assign studies to different conditions or different levels of each moderator, we do not have the ability to infer cause. Although questions of directionality are not an issue (it is clear that—with the exception of publication type—the moderator variable came before the dependent variable), effects caused by untested variables cannot be ruled out. Moderator findings in the current study, therefore, should be viewed as a starting point for future testing rather than a known determinant of the effect.

Because energy is nonsensory in nature, invisible when consumed, and lacks high personal relevance to users, *frequency* was proposed to be more effective if it could draw users' attention to feedback-standard gaps more often, and thus encourage greater engagement and greater savings. Although the results of this study negate this hypothesis and also go against earlier suggestions by Darby (2001, 2006) and Fischer (2008), uncertainties remain because of the discrepancies around the definition of frequency; feedback studies define frequency according to how often energy information is updated; however, from a theoretical standing, frequency refers to how frequently users receive the feedback. These may not be the same thing, which is partly because of the confounding effect of medium; although weekly and less frequent feedback tends to be provided by bills or cards and *pushed* out to users by some means in a way that makes it hard to ignore, the more frequent continuous feedback tends to be provided electronically and requires users to *pull* information from the system. Consequently, more frequent feedback—feedback that is categorized as continuous—may only be accessed occasionally, despite the higher frequency with which it is delivered. This suggests that it may be worth exploring the impact of how frequently users engage with feedback in future studies. It also implies that there may be an upper limit to the amount of time in a week that people spend evaluating and responding to energy feedback for the purpose of reducing overall energy consumption.

Medium was hypothesized to increase the effectiveness of feedback insomuch as it enhanced the opportunities for users to engage. Results showed that studies with feedback that used the least engaging medium (a utility bill) reported the lowest average effect size, whereas studies with feedback that used the most engaging and interactive medium (computer) had the highest effect size. Feedback via a dedicated monitor was less effective than feedback provided via a card, but the difference was not significant, so it is not clear if this was a stable effect or not. Although cards push information to users, for a substantial part of most peoples' everyday lives, it is easy to ignore feedback monitors, especially if they are placed out of sight in an attempt to declutter the home (e.g., as found by Hargreaves, Nye, & Burgess, 2010). Further research around the type of technology used, its visibility, and the frequency with which users engage with the feedback information is needed to disentangle these findings.

The third variable tested was energy *measurement*. Fischer (2008) suggested that energy measurement might moderate the effectiveness of feedback, and hypothesized that feedback presented in terms of cost or carbon would be more effective by stimulating price or environmental concerns, both identified as determinants of proenvironmental behavior. However, the present study found no moderating effect for measurement, suggesting that either the unit of measurement did not stimulate such motivating concerns, or that these concerns did not actually motivate those participating in the study. Additionally, past research has shown that although people may cite concerns about price and the environment as key motivators, they are less effective at stimulating action than a peer-comparison messages (Nolan et al., 2008). Further issues may arise because of the magnitude of differences between the different units; 10 kWh may correspond to $1.00 or 0.007 metric tons of carbon dioxide (which seems like a very small amount), so translating energy use into alternative units may also have the effect of reducing—instead of increasing—the apparent magnitude of feedback-standard gaps. Additional research is required to disentangle this issue.

Although price and environmental messaging did not lead to savings, the *combination* of feedback with external incentives and goal setting did increase the effectiveness of feedback. However, only a small number of studies included such combinations, and none of them compared multiple treatment conditions in a fully crossed factorial design. Thus, further research is needed to explore the full interaction effect between feedback and other behavior-based energy interventions. This is a highly promising finding that should be explored further through additional study.

Providing a feedback-standard referent through the use of *comparison* messages was also hypothesized to moderate feedback effectiveness by providing an explicit standard against which users could reference their current consumption. The overall effect of having a comparison message (vs. no comparison message) was not significant, but differences were found between comparison types. Social and historical comparisons were nonsignificant; however, studies using goal-based comparisons did have a significant (and positive) effect size, suggesting that receiving feedback in the form of goal-based comparisons does reduce energy use. This finding was in line with our hypothesis that goal-based comparisons, which are more likely to keep attention focused at the task-motivation level, would be more effective than social comparisons (as these may drive attention to the metatask level, attenuating the effect of feedback). Our finding that social comparisons were not effective also supports the findings of Fischer (2008), who found no effect related to social comparison in her literature review.

This is a highly relevant finding—although our results suggest that goal comparisons (which tend to be overlaid on one's own historical performance) are the most effective and social comparisons are the least effective, the latter are the most commonly used comparison type of industry feedback (Allcott & Mullainathan, 2010). In fact, the California Public Utilities Commission (2010) restricted *behavior-based energy efficiency programs* to reports that included social comparisons. One reason for this is that energy reports with social comparisons have been validated in large random control trials, providing assurance to utilities and regulatory bodies. However, these trial results only validate that such reports work better than a control group that receives no report at all. The current meta-analysis findings suggest that additional research into different types of comparison messages and other elements within energy reports may lead to higher savings, with no relative increase in cost to providers. Because there is a current opportunity for utilities to provide comparisons to users in regular billing and/or online at a near-real-time frequency (caused by smart meters), this is an important topic that requires further investigation in randomized experiments.

Our finding that historical comparison did not impact feedback effectiveness contrasts with Darby (2006), who concluded that historic feedback is more effective than goal-based comparisons. However, there may be a missing third variable that correlates with type of comparison, such as the size of the feedback-standard gap or who selected or created the standard. Further research is needed to separate the impact of comparison type, relevance, and feedback-standard gap size.

The multiplicity of energy behaviors that contribute to total demand and the ambiguity between action and consequence, lead to our penultimate moderator—*granularity* (i.e., whether feedback information was provided about whole-home usage or specific appliances/devices). Our hypothesis suggested that more granular feedback positively impacts effectiveness by enabling users to identify specific behavior changes to make. One reason for the lack of significance may be that appliance specific information is not useful to users all of the time, and may only be necessary at particular points in time, such as when users are going through a learning process. Additionally, even if users could see where energy is being consumed, this does not necessarily provide them with knowledge around how to reduce consumption; although the feedback may help identify target behaviors, it may not provide users with either general knowledge or knowledge of specific ways to decrease use, both of which are influential on behavior (Hutton, 1982).

The impact of the *duration* over which feedback studies are conducted is the final moderator assessed. Our hypothesis suggested that users' attention to feedback may vary in both quantity and direction (i.e., toward different levels of the hierarchy) over time as they move from initial task learning to higher level self-salient goal satisfaction. The findings support this hypothesis and suggest that users engage with feedback initially, while they go through a process by which they learn how to reduce their consumption. The findings also suggest that after an initial learning phase, users stop engaging with the feedback and energy conservation decreases. Again, this is supported by studies that report how users spend less and less time interacting with feedback as studies progress (Ueno, Inada, Saeki, & Tsuji, 2006). However, feedback provided for longer time periods may allow habits to be created and maintained, which may lead to a rebound in effect size. Further research examining the amount and quality of users' interaction with feedback technologies over time would be useful to explore this finding in more depth.

Limitations and Suggestions for Future Research

As with all meta-analyses, issues related to missing data, limited studies for each moderator, correlations among moderator variables, and differing procedures between studies all decrease the ability to make definitive declarative statements. However, the results presented clearly meet the requirements of the Promising Practices Network (2012), in that (a) they represent an associated change in the dependent variable of more than 1%, (b) changes are significant at the $p < .10$ level, and (c) the samples exceed 10 people in both the treatment and control groups. Additionally, they provide benefits over other forms of literature synthesis because of the use of normalized effect sizes and statistical analysis.

As meta-analysis is used to aggregate findings among results of multiple studies that use different procedures to test a common hypothesis, results are often referred to as *synthesis-generated evidence* as opposed to the *study-generated evidence*, which comes from the analysis of individual studies (Cooper, 2010). Only study-generated evidence is able to make causal attributions, as the variation between study procedures presents potential third variables and confounding results. However, synthesis-generated evidence is extremely useful for exploring associations that were not tested in individual studies, thus providing nuanced and guided suggestions for future empirical research. As explored in the following paragraphs, the current meta-analysis identified five such primary suggestions: (a) the use of factorial designs isolating treatment variation between conditions, (b) greater attention to the physical design and presentation of feedback displays, (c) collection of multiple dependent variables to allow moderation testing, (d) repeated and persistent data collection to assess long-term impacts, and (e) comprehensive presentation of methodology and results to enable greater replication and interpretation of findings.

Factorial designs. One major limitation of the current literature is the failure to test hypotheses through isolation of variables within treatment conditions. Moderator analysis in meta-analysis is essentially correlational; given that studies were not randomly assigned to different levels of the moderators, causation cannot be inferred. However, the treatment variables that seem the most promising can and should be directly tested in the future by incorporating them into the research designs of primary studies. Among the included studies, the treatment conditions were often confounded (e.g., goal-setting and incentives), preventing study authors from determining which strategy was responsible for treatment effects. Of the 22 studies that had more than a single treatment group, 17 featured designs in which treatment groups received different conditions (e.g., control, feedback, feedback plus rebate) but did not fully cross conditions in order to isolate the treatment effect of each variable. An additional nine studies were excluded from analysis because feedback was tested in a between-subjects design, but it could not be isolated for analysis because of confounding variables.

To correct for this, factorial designs are recommended to test research hypotheses and to isolate treatment conditions. To fully understand the interaction between feedback and incentives, for example, it is important to not only have a group that receives feedback and incentives but also a group that receives only incentives and one that receives only feedback. Completely balanced designs allow for the variables themselves, as well as the interactions between variables, to be better understood. Only five studies utilized a complete multifactor ANOVA analysis or multivariate regression model to isolate and analyze the relationship between conditions. Four of them (Becker & Seligman, 1978; Kurz, Donaghue, & Walker, 2005; Mansouri & Newborough, 1999; Winett et al., 1982) tested a factorial design with feedback and another intervention strategy, and one (Robinson, 2007) included a factorial design of Comparison Message (historic vs. social) × Medium (e-mail vs. mail); more studies like this are suggested.

Design and presentation. As suggested by our findings, the way in which feedback information is presented to users can have an impact on the way in which it is perceived and interpreted, and a subsequent impact on motivation and action. However, there has been limited work investigating responses to different types of feedback displays, beyond energy measurement and comparison messages. Froehlich et al. (2010) found that research in "environmental psychology has largely focused on the effect of the feedback intervention itself" and not on "the production of the eco-feedback artifact" (p. 5). Specifically, they found that only half of the environmental psychology articles included a graphic or a description of the feedback interface itself. Of those that did describe the interface, the most common designs were bar or line graphs with usage breakdowns and simple LCD displays that lacked the interactivity and complexity present in both of the new types of feedback in the marketplace, as well as in articles coming out of the human-computer-interaction field.

The few studies that have investigated displays did find differences in the effects of feedback based on the type of graph used (Egan, 1998; Ford & Karlin, 2013) and by comparing ambient (e.g., light changing color) with factual (numbers indicating kWh consumption) feedback (Ham & Midden, 2010). As indicated by these studies, the successful design of energy feedback technologies can greatly benefit from psychological testing of the designs that are used most often in practice, so that feedback design can take into account principles drawn from cognitive and social psychology. As such, it is suggested that psychologists work more closely with designers and design researchers to test theoretically derived design parameters in experimental settings.

Moderator analysis. Savings between homes varied widely, both between and within studies, suggesting differences in context or response that lead to a more effective intervention for some households than others. Although all of the studies included in this analysis collected data on energy use (i.e., kWh savings) for measuring the effectiveness of feedback, few included additional measure of potential individual-level moderators of feedback effectiveness. As discussed in the literature review, several individual-level differences have been found to predict proenvironmental behavior, including both demographic and psychographic variables.

In particular, contextual barriers may impede people from engaging in a behavior regardless of attitudes or motivation (Guagnano et al., 1995). Hypotheses relating to the contextual constraints that impede action could be tested via measures such as home and appliance ownership, financial resources, home location, house size, and relevant cultural constraints. Additionally, consumers may have other personal considerations with regard to energy use that are more self-relevant than conservation (e.g., comfort, security). Understanding these motivations and constraints is vital to the successful use of feedback for energy conservation.

Repeated and persistent data collection. Most studies measured behavior during or immediately after the intervention took place; only five of 42 studies tested for persistence of effects after the intervention had ceased (Hayes & Cone, 1981; Katzev, Cooper, & Fisher, 1980−1981; Kurz et al., 2005; Winett et al., 1982). For those studies, the effect size was higher during the feedback intervention (r = .0790) than during the follow-up period (r = −.0121); however, this difference was not significant (p = .1850). It is unclear whether feedback across other studies would remain effective over the lifetime of a consumer or household. We suggest that future research collect data more often and for a longer period of time, to examine the long-term effects of feedback as an intervention strategy, both during and after the provision of feedback.

Such studies may also help to identify the psychological determinants of behaviors. FIT suggests that over time users may respond to feedback in different ways, shifting their attention between different motivational and learning processes. This hypothesis is supported by survey research of naturalistic feedback users, which revealed a distinction between the use of feedback for tracking (e.g., monitoring ongoing energy use) and learning (e.g., gaining specific information about energy use; Karlin, 2011). If feedback serves as a learning tool that provides knowledge about specific behaviors, then one may expect feedback interventions to provide diminishing returns, such that the effects begin to fade after the subjects have learned everything they can from the information. However, if the role of feedback is to provide ongoing motivation for continued behavior change, then one would expect energy savings to correlate directly with the provision of feedback, remaining stable as long as it is provided and tapering off when the intervention is removed. Repeated and persistent data collection, along with additional self-report data collection about motivation and user experience, could help to provide clarity around the various mechanisms by which feedback interventions operate over time.

Improved reporting. The final suggestion is a more comprehensive presentation of the methodology and results to enable greater replication and interpretation of findings. Many studies failed to present a clear and comprehensive report of the methodologies employed in recruiting and assigning subjects to conditions, as well as the specific details of the intervention strategies tested. As indicated in the Results section, several studies could not be coded on key variables because of missing data (e.g., 33% did not report response rate). Such omissions prevent thorough and comprehensive analysis and replication. It is imperative that authors are clear about the target populations, recruitment and assignment strategies, response rates of participants, and specific details of both the independent (treatment) and dependent (outcome) variables in the study.

Additionally, the presentation of statistical data was inconsistent. Only a few studies reported means and standard deviations for the treatment groups, which should be standard practice in the presentation of experimental research, and seven studies were excluded for not providing sufficient data to calculate an effect size. The presentation of methodology and results of any statistical analyses (or qualitative analyses, for that matter) should be clear and comprehensive, in order to allow transparency in assessing and analyzing study findings. Simply saying that an intervention was "effective" is not nearly as precise as providing the means and standard deviations for the treatment and control conditions or telling the reader which inferential tests were used (e.g., *t* test, ANOVA), along with provision of the test statistics and associated *p* values. More than a suggestion, this is a strong request of all future researchers in this area.

Conclusion

The current paper served to estimate the overall effect size of energy feedback on energy conservation and to examine the potential impact of treatment and study variation. Overall, we found significant evidence that feedback is an effective strategy for promoting energy conservation behavior, with a mean effect size of .071 across 42 studies. We also found that feedback is most effective when it is combined with goal-setting or external incentive interventions, when it provides goal-based comparisons, when gives feedback via a computer, and when the feedback intervention is somewhat brief (e.g., less than 3 months) or quite long (e.g., longer than 1 year).

Because of the nature of synthetic analysis and the limited number of studies that tested each condition, any and all of the variables mentioned herein may be good candidates for future research. Further investigation into how and for whom feedback works best—and the ways in which to administer it most efficiently (considering the cost to savings ratio)—will most certainly increase our knowledge about how to best deploy this strategy in a wider population.

References

References marked with an asterisk indicate studies included in the meta-analysis.

Abrahamse, W., Steg, L., Vlek, C., & Rothengatter, T. (2005). A review of intervention studies aimed at household energy conservation. *Journal of Environmental Psychology, 25,* 273–291. http://dx.doi.org/10.1016/j.jenvp.2005.08.002

Aitken, C. K., McMahon, T. A., Wearing, A. J., & Finlayson, B. L. (1994). Residential water use: Predicting and reducing consumption1. *Journal of Applied Social Psychology, 24,* 136–158. http://dx.doi.org/10.1111/j.1559-1816.1994.tb00562.x

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50,* 179–211. http://dx.doi.org/10.1016/0749-5978(91)90020-T

Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin, 84,* 888–918. http://dx.doi.org/10.1037/0033-2909.84.5.888

*Allcott, H. (2010). *Social norms and energy conservation*. Working paper, Massachusetts Institute of Technology. http://dx.doi.org/10.1016/j.jpubeco.2011.03.003

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics, 95,* 1082–1095. http://dx.doi.org/10.1016/j.jpubeco.2011.03.003

Allcott, H., & Mullainathan, S. (2010). Behavior and energy policy. *Science, 327,* 1204–1205. http://dx.doi.org/10.1126/science.1180775

*Allen, D., & Janda, K. (2006). The effects of household characteristics and energy use consciousness on the effectiveness of real-time energy use feedback: A pilot study. *Proceedings from ACEEE '92: American Council for an Energy Efficient Economy Summer Study on Energy Efficiency in Buildings*. Pacific Grove, CA: American Council for an Energy-Efficient Economy.

American American Recovery and Reinvestment Act of 2009 (ARRA), Pub.L. 111–5, 123 Stat. 115, codified as amended at 26 U.S.C. § section 1.

Anderson, S., & Rodin, J. (1989). Is bad news always bad?: Cue and feedback effects on intrinsic motivation. *Journal of Applied Social Psychology, 19,* 449–467. http://dx.doi.org/10.1111/j.1559-1816.1989.tb00067.x

*Arvola, A. (1993). Billing feedback as a means to encourage household electricity conservation: A field experiment in Helsinki. In R. Ling and H. Wilhite (Eds.), *The energy efficiency challenge for Europe*. Retrieved from http://www.eceee.org/conference_proceedings/eceee/1993/Panel_3/p3_2

*Arvola, A. (1996a). The effect of billing feedback on consumption results of an experiment in Helsinki. In *Research program on consumer habits and energy conservation* (pp. 61–77). Helsinki, Finland: LINKKI.

*Arvola, A. (1996b, October). Results from a feedback experiment in Helsinki. In *Energiräkningen som Informationsbärare*. Symposium conducted at the University of Helsinki, Finland.

Arvola, A., Uutela, A., & Anttila, U. (1994). Billing feedback as a means of encouraging conservation of electricity in households: A field experiment in Helsinki. Proceedings from European Council for an Energy Efficient Economy '94: *European Council for an Energy Efficient Economy Summer Study on Energy Efficiency in Buildings*. Toulon/Hyères, France: European Council for an Energy Efficient Economy.

Attari, S. Z., DeKay, M. L., Davidson, C. I., & Bruine de Bruin, W. (2010). Public perceptions of energy consumption and savings. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 107,* 16054–16059. http://dx.doi.org/10.1073/pnas.1001509107

*Ayres, I., Raseman, S., & Shih, A. (2013). Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage. *Journal of Law, Economics, and Organization, 29,* 992–1022.

Bamberg, S., & Moser, G. (2007). Twenty years after Hines, Hungerford, and Tomera: A new meta- analysis of psycho-social determinants of pro-environmental behavior. *Journal of Environmental Psychology, 27,* 14–25. http://dx.doi.org/10.1016/j.jenvp.2006.12.002

Bandura, A. (1969). *Principles of behavior modification*. New York, NY: Holt, Rinehart & Winston.

Barr, S., Gilg, A. W., & Ford, N. (2005). The household energy gap: Examining the divide between habitual-and purchase-related conservation behaviors. *Energy Policy, 33,* 1425–1444. http://dx.doi.org/10.1016/j.enpol.2003.12.016

*Battalio, R. C., Kagel, J. H., Winkler, R. C., & Winett, R. A. (1979). Residential electricity demand: An experimental study. *The Review of Economics and Statistics, 61,* 180–189. http://dx.doi.org/10.2307/1924585

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

*Becker, L. J., & Seligman, C. (1978). Reducing air conditioning waste by signaling it is cool outside. *Personality and Social Psychology Bulletin, 4,* 412–415. http://dx.doi.org/10.1177/014616727800400310

Becoña, E., & Vázquez, F. L. (2001). Effectiveness of personalized written feedback through a mail intervention for smoking cessation: A randomized-controlled trial in Spanish smokers. *Journal of Consulting and Clinical Psychology, 69,* 33–40. http://dx.doi.org/10.1037/0022-006X.69.1.33

*Bittle, R. G., Valesano, R., & Thaler, G. (1979). The effects of daily feedback on residential electricity usage as a function of usage level and type of feedback information. *Journal of Environmental Systems, 9,* 275–287. http://dx.doi.org/10.2190/91AA-P97G-JF92-T7EJ

*Bittle, R. G., Valesano, R., & Thaler, G. (1979–1980). The effects of daily cost feedback on residential electricity consumption. *Behavior Modification, 3,* 187–202. http://dx.doi.org/10.1177/014544557932004

Black, J. S., Stern, P. C., & Elworth, J. T. (1985). Personal and contextual influences on household energy adaptations. *Journal of Applied Psychology, 70,* 3–21. http://dx.doi.org/10.1037/0021-9010.70.1.3

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: Wiley. http://dx.doi.org/10.1002/9780470743386

*Brandon, G., & Lewis, A. (1999). Reducing household energy consumption: A qualitative and quantitative field study. *Journal of Environmental Psychology, 19,* 75–85. http://dx.doi.org/10.1006/jevp.1998.0105

Bridgeman, B. (1974). Effects of test score feedback on immediately subsequent test performance. *Journal of Educational Psychology, 66,* 62–66. http://dx.doi.org/10.1037/h0035803

California Public Utilities Commission. (2010). Decision determining evaluation, measurement and verification processes for 2010 through 2012 Energy Efficiency Portfolios. Report D. 10–14-029. Sacramento, CA: Author.

Campion, M. A., & Lord, R. G. (1982). A control systems conceptualization of the goal-setting and changing process. *Organizational Behavior & Human Performance, 30,* 265–287. http://dx.doi.org/10.1016/0030-5073(82)90221-5

Carver, C. S., & Scheier, M. F. (1981). *Attention and self-regulation: A control-theory approach to human behavior*. New York, NY: Springer-Verlag. http://dx.doi.org/10.1007/978-1-4612-5887-2

Chopra, A. (2011). Modeling a green energy challenge after a blue button. *White House Office of Science and Technology Policy*. Retrieved from http://www.whitehouse.gov/blog/2011/09/15/modeling-green-energy-challenge-after-blue-button

Cialdini, R. B., & Schultz, W. (2004). Understanding and motivating energy conservation via social norms (Research Report prepared for the William and Flora Hewlett Foundation). Retrieved from OPower website: http://opower.com/uploads/library/file/2/understanding_and_motivating_energy_conservation_via_social_norms.pdf

Cooper, H. (2010). *Research synthesis and meta-analysis*. Thousand Oaks, CA: Sage.

Cooper, H., & Hedges, L. V. (1994). *The handbook of research synthesis*. New York, NY: Russell Sage Foundation.

Darby, S. (2001). Making it obvious: Designing feedback into energy consumption. In P. Bertoldi, A. Ricci, & A. de Almeida (Eds.), *Energy Efficiency in Household Appliances and Lighting* (685–696). Berlin: Springer.

Darby, S. (2006). *The effectiveness of feedback on energy consumption: A Review for DEFRA of the literature on metering, billing and direct displays*. Oxford, UK: Environmental Change Institute. Retrieved from http://www.eci.ox.ac.uk/research/energy/downloads/smart-metering-report.pdf

Dietz, T., Gardner, G. T., Gilligan, J., Stern, P. C., & Vandenbergh, M. P. (2009). Household actions can provide a behavioral wedge to rapidly reduce US carbon emissions. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 106,* 18452–18456. http://dx.doi.org/10.1073/pnas.0908738106

Dillman, D., Rosa, E., & Dillman, J. (1983). Lifestyle and home energy conservation in the United States: The poor accept lifestyle cutbacks while the wealthy invest in conservation. *Journal of Economic Psychology, 3,* 299–315. http://dx.doi.org/10.1016/0167-4870(83)90008-9

*Dobson, J. K., & Griffin, J. D. A. (1992). Conservation effect of immediate electricity cost feedback on residential consumption behavior. *Proceedings from ACEEE '92: American Council for an Energy Efficient Economy Summer Study on Energy Efficiency in Buildings*. Pacific Grove, CA: American Council for an Energy-Efficient Economy.

Egan, C. (1998). Graphical displays and comparative energy information: What do people understand and prefer? *Proceedings from ACEEE '98: American Council for an Energy Efficient Economy Summer Study on Energy Efficiency in Buildings*. Pacific Grove, CA: American Council for an Energy-Efficient Economy.

Ehrhardt-Martinez, K., Donnelly, K., & Laitner, J. A. (2010). *Advanced metering initiatives and residential feedback programs: A meta-review for household electricity-saving opportunities*. Washington DC: American Council for an Energy Efficient Economy (ACEEE). Retrieved from http://aceee.org/research-report/e105

Environmental Protection Agency. (2011). Statement on recent findings in West Coast states and Hawaii.

Electric Power Research Institute. (2009). *Residential electricity use feedback: A research synthesis and economic framework*. Palo Alto, CA: Author.

Festinger, L. A. (1954). A theory of social comparison processes. *Human Relations, 7,* 117–140. http://dx.doi.org/10.1177/001872675400700202

Field, A. P. (2001). Meta-analysis of correlation coefficients: A Monte Carlo comparison of fixed- and random-effects methods. *Psychological Methods, 6,* 161–180. http://dx.doi.org/10.1037/1082-989X.6.2.161

Fischer, C. (2008). Feedback on household electricity consumption: A tool for saving energy? *Energy Efficiency, 1,* 79–104. http://dx.doi.org/10.1007/s12053-008-9009-7

Fogg, B. J. (2009). A behavior model for persuasive design. In Proceedings of the 4th International Conference on Persuasive Technology (p. 40). New York, NY: ACM.

Ford, R., & Karlin, B. (2013). Graphical displays in eco-feedback: A cognitive approach. In A. Marcus (Ed.), *Design, User Experience, and Usability. Web, Mobile, and Product Design* (pp. 486–495). Berlin: Springer.

Froehlich, J., Findlater, L., & Landay, J. (2010, April). The design of eco-feedback technology. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1999–2008). ACM.

Gardner, G. T., & Stern, P. C. (2008). The short list: The most effective actions US households can take to curb climate change. *Environment: Science and Policy for Sustainable Development, 50,* 12–25.

Granade, H. C., Creyts, J., Derkach, A., Farese, P., Nyquist, S., & Ostrowski, K. (2009). *Unlocking energy efficiency in the U.S. economy*. Washington, DC: McKinsey.

Guagnano, G., Stern, P., & Dietz, T. (1995). Influences on attitude-behavior relationships: A natural experiment with curbside recycling. *Environment and Behavior, 27,* 699–718. http://dx.doi.org/10.1177/0013916595275005

Guzzo, R. A., Jette, R. D., & Katzell, R. A. (1985). The effects of psychologically based intervention programs on worker productivity: A meta-analysis. *Personnel Psychology, 38,* 275–291. http://dx.doi.org/10.1111/j.1744-6570.1985.tb00547.x

*Haakana, M., Sillanpaeae, L., & Talsi, M. (1997). The effect of feedback and focused advice on household energy consumption. *Proceedings from ECEEE '97: European Council for an Energy Efficient Economy Summer Study on Energy Efficiency in Buildings*. Toulon/Hyères, France: European Council for an Energy Efficient Economy.

Ham, J., & Midden, C. (2010). Ambient persuasive technology needs little cognitive effort: The differential effects of cognitive load on lighting feedback versus factual feedback. In *Persuasive 2010: Proceedings of 5th International Conference on Persuasive Technology,* pp. 132–142. Springer Berlin Heidelberg.

Hanna, G. S. (1976). Effects of total and partial feedback in multiple-choice testing upon learning. *The Journal of Educational Research, 69,* 202–205. http://dx.doi.org/10.1080/00220671.1976.10884873

Hardin, G. (1968). The tragedy of the commons: The population problem has no technical solution; it requires a fundamental extension in morality. *Science, 162,* 1243–1248.

Hargreaves, T., Nye, M., & Burgess, J. (2010). Making energy visible: A qualitative field study of how householders interact with feedback from smart energy monitors. *Energy Policy, 38,* 6111–6119. http://dx.doi.org/10.1016/j.enpol.2010.05.068

Harrigan, M. (1992). *Evaluating consumer energy education: A field test of three education packages in support of a residential conservation and load-shifting program*. Washington, DC: Alliance to Save Energy.

Harrigan, M. S., & Gregory, J. M. (1994). *Do savings from energy education persist?* Washington, DC: Alliance to Save Energy.

*Hayes, S. C., & Cone, J. D. (1981). Reduction of residential consumption of electricity through simple monthly feedback. *Journal of Applied Behavior Analysis, 14,* 81–88. http://dx.doi.org/10.1901/jaba.1981.14-81

Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods, 3,* 486–504. http://dx.doi.org/10.1037/1082-989X.3.4.486

Hunter, J. E., & Schmidt, F. L. (2000). Fixed effects vs. random effects in meta-analysis models: Implications for cumulative research knowledge. *International Journal of Selection and Assessment, 8,* 275–292. http://dx.doi.org/10.1111/1468-2389.00156

Hutton, R. B. (1982). Advertising and the Department of Energy's campaign for energy conservation. *Journal of Advertising, 11,* 27–39. http://dx.doi.org/10.1080/00913367.1982.10672801

*Hutton, R. B., Mauser, G. A., Filiatrault, P., & Ahtola, O. T. (1986). Effects of cost-related feedback on consumer knowledge and consumption behavior: A field experimental approach. *Journal of Consumer Research, 13,* 327–336. http://dx.doi.org/10.1086/209072

Jones, G. M. (1910). Experiments on the reproduction of distance as influenced by suggestions of ability and inability. *Psychological Review, 17,* 269–278. http://dx.doi.org/10.1037/h0074504

Judd, C. H. (1905). Practice without knowledge of results. *The Psychological Review: Monograph Supplements, 2,* 185–198.

Karlin, B. (2011, June). Tracking and learning: Exploring dual functions of residential energy feedback. In Proceedings of the 6th International Conference on Persuasive Technology: Persuasive Technology and Design: Enhancing Sustainability and Health. New York, NY: ACM.

Karlin, B., Davis, N., Sanguinetti, A., Gamble, K., Kirkby, D., & Stokols, D. (2014). Dimensions of conservation: Exploring differences among energy behaviors. *Environment and Behavior, 46,* 423–452. http://dx.doi.org/10.1177/0013916512467532

Karlin, B., Ford, R., & Squiers, C. (2014). Energy feedback technology: A review and taxonomy of products and platforms. *Energy Efficiency, 7,* 377–399. http://dx.doi.org/10.1007/s12053-013-9227-5

*Kasulis, J. J., Huettner, D. A., & Dikeman, N. J. (1981). The feasibility of changing electricity consumption patterns. *Journal of Consumer Research, 8,* 279–290. http://dx.doi.org/10.1086/208866

*Katzev, R., Cooper, L., & Fisher, P. (1980−1981). The effect of feedback and social reinforcement on residential electricity consumption. *Journal of Environmental Systems, 10,* 215–227. http://dx.doi.org/10.2190/M8W4-JUYV-NB03-CUFL

Katzev, R., & Johnson, T. (1987). *Promoting energy conservation: An analysis of behavioral research*. Boulder, CO: Westview Press.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119,* 254–284. http://dx.doi.org/10.1037/0033-2909.119.2.254

*Kurz, T., Donaghue, N., & Walker, I. (2005). Utilizing a social-ecological framework to promote water and energy conservation: A field experiment. *Journal of Applied Social Psychology, 35,* 1281–1300. http://dx.doi.org/10.1111/j.1559-1816.2005.tb02171.x

Latham, G. P., & Locke, E. A. (1991). Self-regulation through goal setting. *Organizational Behavior and Human Decision Processes, 50,* 212–247. http://dx.doi.org/10.1016/0749-5978(91)90021-K

Leiserowitz, A. (2008). International public opinion, perception, and understanding of global climate change. *Report for Yale project on climate change*. New Haven, CT: Yale University Press.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice Hall.

Lutzenhiser, L., Cesafsky, L., Chappells, H., Gossard, M., Moezzi, M., Moran, D., & Wilhite, H. (2009). *Behavioral assumptions underlying California residential sector energy efficiency programs*. Oakland, CA: CIEE. Retrieved from http://uc-ciee.org/downloads/ba_ee_res_wp.pdf

*Mansouri, I., & Newborough, M. (1999). Dynamics of energy use in UK households: End-use monitoring of electric cookers. *Proceedings from ECEEE '99: European Council for an Energy Efficient Economy Summer Study on Energy Efficiency in Buildings*. Toulon/Hyères, France: European Council for an Energy Efficient Economy.

Manten, A. A. (1973). Scientific literature review. *Scholarly Publishing, 5,* 75–89.

Markowitz, E. M., & Shariff, A. F. (2012). Climate change and moral judgement. *Nature Climate Change, 2,* 243–247. http://dx.doi.org/10.1038/nclimate1378

*Matsukawa, I. (2004). The effects of information on residential demand for electricity. *The Energy Journal (Cambridge, Mass.), 25,* 1–17. http://dx.doi.org/10.5547/ISSN0195-6574-EJ-Vol25-No1-1

*McClelland, L., & Cook, S. W. (1979–1980). Energy conservation effects of continuous in-home feedback in all-electric homes. *Journal of Environmental Systems, 9,* 169–173. http://dx.doi.org/10.2190/L8BU-ECLK-PEC5-KKTW

*Midden, C. J. H., Meter, J. F., Weenig, M. H., & Zieverink, H. J. A. (1983). Using feedback, reinforcement and information to reduce energy consumption in households: A field-experiment. *Journal of Economic Psychology, 3,* 65–86. http://dx.doi.org/10.1016/0167-4870(83)90058-2

*Mountain, D. C. (2007). Real-time feedback and residential electricity consumption: British Columbia and Newfoundland and Labrador Pilots. Ontatio, Canada: Mountain Economic Consulting.

Mountain Economic Consulting and Associates. (2006). *The impact of real-time feedback on residential electricity consumption: The hydro one pilot*. Ontario, Canada: Author.

Nair, G., Gustavsson, L., & Mahapatra, K. (2010). Factors influencing energy efficiency investments in existing Swedish residential buildings. *Energy Policy, 38,* 2956–2963. http://dx.doi.org/10.1016/j.enpol.2010.01.033

*Nexus Energy Software. (2006). *California bill analysis pilot final report*. San Francisco, CA: Calmac. Retrieved from http://www.calmac.org/publications/ca_bill_analysis_pilot_final_04-06es.pdf

Nolan, J. M., Schultz, P. W., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2008). Normative social influence is underdetected. *Personality and Social Psychology Bulletin, 34,* 913–923. http://dx.doi.org/10.1177/0146167208316691

*Pallak, M., Cook, D., & Sullivan, J. (1980). Commitment and energy conservation. In L. Bickman (Ed.), *Applied Social Psychology Annual, 1,* 235–253.

*Pallak, M. S., & Cummings, N. (1976). Commitment and voluntary energy conservation. *Personality and Social Psychology Bulletin, 2,* 27–30. http://dx.doi.org/10.1177/014616727600200105

Parker, D. S., Hoak, D., & Cummings, J. (2008). *Pilot evaluation of energy savings from residential energy demand feedback devices*. Cocoa, FL: Florida Solar Energy Center. Retrieved from http://www.fsec.ucf.edu/en/publications/pdf/FSEC-CR-1742-08.pdf

Parker, D. S., Mazzara, M., & Sherwin, J. (1996). Monitored energy use patterns in low-income housing in a hot and humid climate. Proceedings from IBSHHC '96: *Tenth Symposium on Improving Building Systems in Hot Humid Climates*. Fort Worth, TX: TAMU.

Pearce, J. L., & Porter, L. W. (1986). Employee responses to formal performance appraisal feedback. *Journal of Applied Psychology, 71,* 211–218. http://dx.doi.org/10.1037/0021-9010.71.2.211

Pew Research Center. (2014). *Internet use over time*. Washington DC: Author. Retrieved from http://www.pewinternet.org/data-trend/internet-use/internet-use-over-time/

Pigott, T. D. (2009). Handling missing data. In H. M. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 399–416). New York, NY: Russell Sage Foundation.

Podsakoff, P. M., & Farh, J. H. (1989). Effects of feedback sign and credibility on goal setting and task performance. *Organizational Behavior and Human Decision Processes, 44,* 45–67. http://dx.doi.org/10.1016/0749-5978(89)90034-4

Poortinga, W., Steg, L., Vlek, C., & Wiersma, G. (2003). Household preferences for energy-saving measures: A conjoint analysis. *Journal of Economic Psychology, 24,* 49–64. http://dx.doi.org/10.1016/S0167-4870(02)00154-X

Promising Practices Network. (2012). *How programs are considered*. Retrieved from http://www.promisingpractices.net/criteria.asp

R Core Team. (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/

*Robinson, J. (2007). *The effect of electricity-use feedback on residential consumption: A case study of customers with smart meters in Milton, Ontario* (Master's thesis). University of Waterloo, Waterloo, Canada.

Rosenthal, R. (1991). *Meta-analytic procedures for social research* (Rev. ed.). Newbury Park, CA: Sage. http://dx.doi.org/10.4135/9781412984997

Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology, 52,* 59–82. http://dx.doi.org/10.1146/annurev.psych.52.1.59

Rosenthal, R., & Rubin, D. B. (2003). r equivalent: A simple effect size indicator. *Psychological Methods, 8,* 492–496.

Rothstein, H. R., & Hopewell, S. (2009). Grey literature. In C. Harris, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (vol 2, pp. 103–125). New York, NY: Russell Sage Foundation.

Samuelson, C. D., & Biek, M. (1991). Attitudes toward energy conservation: A confirmatory factor analysis. *Journal of Applied Social Psychology, 21,* 549–568. http://dx.doi.org/10.1111/j.1559-1816.1991.tb00536.x

Schultz, P. W. (2010). Making energy conservation the norm. In K. Ehrardt-Martinez & J. A. Laitner (Eds.), *People-centered initiatives for increasing energy savings* (pp. 251–262). Washington, DC: American Council for an Energy-Efficient Economy.

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science, 18,* 429–434. http://dx.doi.org/10.1111/j.1467-9280.2007.01917.x

Schultz, W. P., Khazian, A. M., & Zaleski, A. C. (2008). Using normative social influence to promote conservation among hotel guests. *Social Influence, 3,* 4–23. http://dx.doi.org/10.1080/15534510701755614

Schwartz, S. H. (1977). Normative influences on altruism. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 221–279). New York, NY: Academic Press.

*Seaver, W. B., & Patterson, A. H. (1976). Decreasing fuel-oil consumption through feedback and social commendation. *Journal of Applied Behavior Analysis, 9,* 147–152. http://dx.doi.org/10.1901/jaba.1976.9-147

*Seligman, C., & Darley, J. M. (1977). Feedback as a means of decreasing residential energy consumption. *Journal of Applied Psychology, 62,* 363–368. http://dx.doi.org/10.1037/0021-9010.62.4.363

*Seligman, C., Darley, J. M., & Becker, L. J. (1978). Behavioral approaches to residential energy conservation. *Energy and Building, 1,* 325–337. http://dx.doi.org/10.1016/0378-7788(78)90012-9

*Sexton, R. J., Johnson, N. B., & Konakayama, A. (1987). Consumer response to continuous-display electricity-use monitors in a time-of-use pricing experiment. *Journal of Consumer Research, 14,* 55–62. http://dx.doi.org/10.1086/209092

*Sexton, R. J., & Sexton, T. A. (1987). Theoretical and methodological perspectives on consumer response to electricity information. *The Journal of Consumer Affairs, 21,* 238–257. http://dx.doi.org/10.1111/j.1745-6606.1987.tb00201.x

*Sexton, R. J., Sexton, T. A., Wann, J. J., & Kling, C. L. (1989). The conservation and welfare effects of information in a time-of-day pricing experiment. *Land Economics, 65,* 272–279. http://dx.doi.org/10.2307/3146671

Shadish, W. R., & Haddock, C. K. (1994). Combining estimates of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 261–282). New York, NY: Russell Sage Foundation.

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics, 69,* 99–118. http://dx.doi.org/10.2307/1884852

*Sipe, B., & Castor, S. (2009). The net impact of home energy feedback devices. Proceedings from IEPEC '09: *International Energy Program Evaluation Conference,* Portland, OR: IEPEC.

Sipe, B., & Castor, S. (2009). The Net Impact of Home Energy Feedback Devices. In *Energy Program Evaluation Conference* (pp. 341–351). Portland, OR: IEPEC.

Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York, NY: Appleton-Century.

Smith, M. L. (1980). Publication bias and meta-analysis. *Evaluation in Education, 4,* 22–24. http://dx.doi.org/10.1016/0191-765X(80)90004-X

Steg, L., & Vlek, C. (2009). Encouraging pro-environmental behavior: An integrative review and research agenda. *Journal of Environmental Psychology, 29,* 309–317. http://dx.doi.org/10.1016/j.jenvp.2008.10.004

Stephenson, J., Barton, B., Carrington, G., Gnoth, D., Lawson, R., & Thorsnes, P. (2010). Energy cultures: A framework for understanding energy behaviours. *Energy Policy, 38,* 6120–6129. http://dx.doi.org/10.1016/j.enpol.2010.05.069

Stern, P. C. (2011). Contributions of psychology to limiting climate change. *American Psychologist, 66,* 303–314. http://dx.doi.org/10.1037/a0023235

Stock, W. A. (1994). Systematic coding for research synthesis. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 125–138). New York, NY: Russell Sage Foundation.

Stokols, D. (1979). A congruence analysis of human stress. In I. G. Sarason & C. D. Spielberger (Eds.), *Stress and anxiety* (Vol. 6, pp. 27–53). New York, NY: Wiley.

Summit Blue Consulting. (2009). *Impact evaluation of OPOWER SMUD pilot study*. Boulder, CO: Author.

Tate, D. F., Wing, R. R., & Winett, R. A. (2001). Using Internet technology to deliver a behavioral weight loss program. *JAMA: Journal of the American Medical Association, 285,* 1172–1177. http://dx.doi.org/10.1001/jama.285.9.1172

Thorndike, E. L. (1927). The law of effect. *The American Journal of Psychology, 39,* 212–222. http://dx.doi.org/10.2307/1415413

Turaga, R. M. R., Howarth, R. B., & Borsuk, M. E. (2010). Pro-environmental behavior: Rational choice meets moral motivation. *Annals of the New York Academy of Sciences, 1185,* 211–224. http://dx.doi.org/10.1111/j.1749-6632.2009.05163.x

*Ueno, T., Inada, R., Saeki, O., & Tsuji, K. (2005). Effectiveness of displaying energy consumption data in residential houses. Analysis on how the residents respond. *Proceedings from ECEEE '05: European Council for an Energy Efficient Economy Summer Study on Energy Efficiency in Buildings* (pp. 1289–1299). Toulon/Hyères, France: European Council for an Energy Efficient Economy.

*Ueno, T., Inada, R., Saeki, O., & Tsuji, K. (2006). Effectiveness of an energy-consumption information system for residential buildings. *Applied Energy, 83,* 868–883. http://dx.doi.org/10.1016/j.apenergy.2005.09.004

Vallacher, R. R., & Wegner, D. M. (1987). What do people think they're doing? Action identification and human behavior. *Psychological Review, 94,* 3–15. http://dx.doi.org/10.1037/0033-295X.94.1.3

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

Van Liere, K. D., & Dunlap, R. E. (1978). Moral norms and environmental behavior: An application of Schwartz's norm-activation model to yard burning. *Journal of Applied Social Psychology, 8,* 174–188. http://dx.doi.org/10.1111/j.1559-1816.1978.tb00775.x

Verhallen, T. M. M., & van Raaij, W. F. (1981). Household behavior and the use of natural gas for home heating. *Journal of Consumer Research, 8,* 253–257. http://dx.doi.org/10.1086/208862

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metaphor package. *Journal of Statistical Software, 36,* 1–48.

Wicklund, R. A. (1975). Objective self-awareness. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 8[3], pp. 233–237). New York, NY: Academic Press. http://dx.doi.org/10.1016/S0065-2601(08)60252-X

*Wilhite, H., & Ling, R. (1995). Measured energy savings from a more informative energy bill. *Energy and Building, 22,* 145–155. http://dx.doi.org/10.1016/0378-7788(94)00912-4

Wilson, C., & Dowlatabadi, H. (2007). Models of decision-making and residential energy use. *Annual Review of Environment and Resources, 32,* 169–203. http://dx.doi.org/10.1146/annurev.energy.32.053006.141137

Wilson, D. B. (2009). Systematic coding. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 159–176). New York, NY: Russell Sage Foundation.

*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. *Journal of Applied Behavior Analysis, 15,* 381–402. http://dx.doi.org/10.1901/jaba.1982.15-381

*Winett, R. A., Kagel, J. H., Battalio, R. C., & Winkler, R. C. (1978). Effects of monetary rebates, feedback, and information on residential electricity conservation. *Journal of Applied Psychology, 63,* 73–80. http://dx.doi.org/10.1037/0021-9010.63.1.73

*Wood, G., & Newborough, M. (2003). Dynamic energy-consumption indicators for domestic appliances: Environment, behavior and design. *Energy and Building, 35,* 821–841. http://dx.doi.org/10.1016/S0378-7788(02)00241-4

Wright, W. R. (1906). Some effects of incentive on work and fatigue. *Psychological Review, 13,* 23–34. http://dx.doi.org/10.1037/h0076041

Received March 23, 2013
Revision received April 28, 2015
Accepted June 22, 2015 ■