Big, hot, or bright? Integrating cues to perceive home energy use

Eleanor B. Schille-Hudson[1] (erbrower@iu.edu)**, Tyler Margehtis[1,2]** (tyler.marghetis@gmail.com)**,
Deidra Miniard[2]** (deidraminiard@gmail.com)**, David Landy[1,3]** (dhlandy@gmail.com)**, Shahzeen Z.
Attari[2]** (sattari@indiana.edu)
[1]Department of Psychological and Brain Sciences, Indiana University; [2]O'Neill School of
Public and Environmental Affairs, Indiana University; [3]Netflix

Abstract

Despite constantly using energy and having extensive interactions with household appliances, people consistently mis-estimate the amount of energy that is used by home appliances. This poses major problems for conservation efforts, while also presenting an interesting case study in human perception. Since many forms of energy used are not directly perceptible, and since the *amount* of energy that is being used by an appliance is often difficult to infer from appearances alone, people often rely on cues. Some of these cues are more reliable than others and previous literature has investigated which of these cues people rely on. However, past literature has always studied these proximal cues in isolation—despite the fact that, during real-world perception, people are always integrating a variety of cues. Here, we investigate how people rely on a variety of cues, and how individual differences in the reliance on those cues predicts the ability to estimate home energy use.

Keywords: energy; perception; estimation; home appliances; multi-dimensional scaling

Introduction

Despite its importance in the face of catastrophic climate change, energy and energy use are not well or widely understood by the public. For many home appliances, we have only indirect access to the appliances' energy use and energy units are difficult to understand. However, people frequently make choices as energy consumers: When should I turn off the lights? For how long should I take a hot shower? To what temperature should I set the thermostat? These daily decisions all depend on a perception of energy use. Given people's poor understanding of energy use and the difficulty of perceiving energy use, how do people make these daily decisions about using energy?

In some contexts, people have access to explicit information about appliances' energy use. For instance, some smart meters are digital devices that indicate, in real time, how much energy is being used by an appliance; other appliances may have labels indicating their average energy use (e.g., "Energy Star" labels on efficient appliances). However, explicit information about energy use is the exception, not the rule.

In the absence of direct, explicit information about energy use, people must rely on indirect indices of energy use. (For reference on similar work done in the HCI community see He, Greenberg, and Huang, 2010 and Heller, Konstantinos, Borchers, 2013.)

Vacuums are noisy. Lightbulbs are luminous and sometimes hot. It is these observable features that are typically available to individuals when they are making decisions about their energy use. Some of these cues, however, are more reliable than others. For instance, generating and extracting heat requires a lot of energy; mechanical movement, while perhaps more perceptually salient, can be accomplished with far less energy. Good judgements and decisions about energy use, therefore, requires a good sense of which proximal cues to rely on, and which to ignore. Understanding and improving these judgments can translate to energy conservation, as illustrated by the conservation benefits of in-home smart devices that give real-time feedback on energy-use (Darby, 2006; Delmas, Fischlein, & Asensio, 2013), although these energy technologies may be years away from becoming mainstream.

Past work on situated perception and decision making has advocated for similar approaches to understanding how people make judgments about entities that cannot be perceived directly. For instance, Brunswick (1956) proposed a "lens model" of perception, in which people must integrate across proximal cues in order to decide whether some target entity or property exists in the world; on this account, learning to perceive correctly involves learning how best to weight these different cues, so that more reliable cues (i.e., those that most often co-occur with the target phenomenon) are weighted more. A similar perspective has been advocated by researchers in the Judgement and Decision Making world, who have argued that, for many difficult decisions, people deploy 'replacement heuristics' — relying on some simpler or more easily perceived property or feature to make decisions about some target phenomenon that is more complex or difficult to perceive (Kahneman & Frederick, 2002). On all these approaches, understanding how people make complex perceptual judgments about 'invisible' entities, such as energy use, requires understanding the proximal cues or features they are relying on.

A number of past studies have tried to do exactly that. Previously, in the energy literature, different replacement heuristics have been studied. Past work has suggested that novices base their estimates of home energy use on perceptions of appliances' size (Cowen & Gatersleben, 2017), frequency of use (Schley & DeKay, 2015), effect on temperature (heating or cooling) (Attari, DeKay, Davidson, & de Bruin, 2010), and type of appliance (Lesic, Bruin, Davis, Krishnamurti, & Azevedo, 2018). But these past studies have focused on a single dimension

of experience (e.g., size), in isolation from the many other features which that dimension may be correlated (e.g., frequency of use). As a result, we still do not know how people weight the range of features to which they have access, or whether there is one or a subset of features that are driving most of people's energy estimates.

Moreover, all these approaches share the prediction that *better* judgements will involve better weighting of proximal cues. How do individual differences in weighting these features relate to individual differences in estimation ability?

Here, we attempt to answer these three outstanding questions: Which features are people relying on to make energy estimates? How do individual differences in cue-weighting relate to estimation skill? And how can we capture people's feature representation of appliances in a way that accounts for correlations among features?

In the following studies, we first surveyed participants for the most important or relevant features of energy in home appliances. We then took the most frequently cited features and used them to create feature rating scales for participants to rate multiple home appliances along. A multiple regression was performed on a few theoretically-driven features to determine how they competed with one another. Multi-dimensional scaling (MDS) was performed on all the features to capture the structure in how people perceive appliances and their energy use.

By performing these analyses on multiple features at once, we can establish which features matter *most* in the larger context of available appliance features. We also hope to paint a more clear and nuanced — and thus complete — picture of how these features are combined with one another. MDS affords us a look at categories of appliances that emerge and have implications for why some categories matter. These targeted analyses in concert with the larger picture of appliance feature perception, will hopefully inform future projects on how to help people better understand and use energy (Marghetis, Attari, and Landy, under review).

Methods

Participants
We recruited adults (*N* = 299) from the United States through Amazon Mechanical Turk, an online labor market that has been used previously for online studies. Each subject participated in return for $5. Only the data from those participants who completed the entire study were analyzed (*N = 261*). We also removed participants who repeated the exact same response for their estimates of all appliances (*n* = 1), giving us a final sample of *N* = 260.

Feature Selection
Participants rated features that were selected based on a previous study with different participants (*N = 17*) in which people were asked to list all features that they would use to estimate an appliance's energy use. On the basis of these free response features, we compiled a list of features that were most frequently cited and most widely applicable to our list of home appliances (*N =13*, see Appendix).

Procedure
Participants first completed a feature rating task, in which they were presented with typical home appliances (*N = 36*) and asked to judge each appliance in terms of a set of perceptual or experiential features (e.g., brightness, loudness). They were first instructed "For each question, [to] please imagine a typical version of that appliance *while it is in use* and answer accordingly." The survey was organized by feature. For each feature, e.g. "How loud is each appliance?", participants were given a Likert scale from 1-10 as well as a Not Applicable box for each appliance. Both appliances and features were presented in a random order. Participants supplied ratings for the following features: how frequently the appliance is used, how big the appliance is, how long the appliance is used, how much light the appliance produces, how much the appliance heats itself/its environment, how much sound it makes, how much water it uses, how much it cools itself/its environment, how big its motor is, how much it heats water, how complex its software is, how complex its internal electronic components are, how complex its internal mechanical components are, how much movement it generates in itself/environment. Each participant rated each appliance along each feature dimension, totaling 36 x 13 ratings for each participant.

After the feature rating task, participants were asked to make energy estimates for each appliance. They were given a point of reference: "A 100-watt incandescent light bulb uses 100 units of energy in one hour." Then they were asked to make an estimate for each appliance, "How many units of energy do you think each of the following devices typically uses in one hour?" Appliances were presented in a random order. This task has been used in prior studies to investigate and elicit accuracy in energy perceptions (e.g., Attari et al., 2010).

Analysis
A multiple regressions analysis was run on features that have been identified in past research as important for energy estimation use (Cowen & Gatersleben, 2017; Schley & DeKay, 2015; Marghetis, Attari, and Landy, under review), namely: size, how "electronic" the appliance is, frequency of use, and how much the appliance changes the temperature (i.e., the maximum of the heating and cooling ratings). Feature ratings were z-scored across participants. In a mixed effects model, there were fixed slopes for the interaction of features and feature ratings of every participant, random intercepts on every appliance, and random slopes on feature ratings by participant. The random slopes for every participant's ratings were extracted and used to investigate individual differences in energy estimating accuracy.

Results

What proximal cues do people use to estimate appliances' energy use?

We first zoomed in on those features that have been identified, in past literature, as playing a role in novice's judgements of home energy use. These included how frequently the appliance is used, how "electronic" the appliance is, how much the appliance changes the temperature (the max of the 'heat' and 'cool' ratings), and how large the appliance is. Using a linear mixed effects model, we predicted participants' energy estimates (log transformed) using these four features, with random intercepts and slopes for participants, and random intercepts for appliances. Feature ratings were z-scored within each participant. See Figure 1 for coefficient estimates of reliance on these proximal cues.

Participants' estimates of appliances' energy use were driven almost entirely by how large they judged the appliance to be (b = 0.10 ± 0.01 SEM, p<.001). Most variance in estimates is accounted for by differences in *size*. By contrast, people's judgments of how much the appliance changed the temperature and of how "electronic" an appliance was also had much smaller relations to their energy estimates (b = 0.04 ± 0.01 SEM, p<.001, b = 0.05 ± 0.01 SEM, p<.001). Critically, we found no relation between judgments of how often an appliance is used and estimates of how much energy it uses — despite past work that has argued that frequency-of-use is used as a 'replacement heuristic' for energy estimation (Schley & DeKay, 2015). Note that people's estimates of energy use were explained primarily by judgments of the appliance's size rather than by how much the appliance changed the temperature, even though heat is a more reliable cue to energy use, because heating and cooling use a lot of energy.
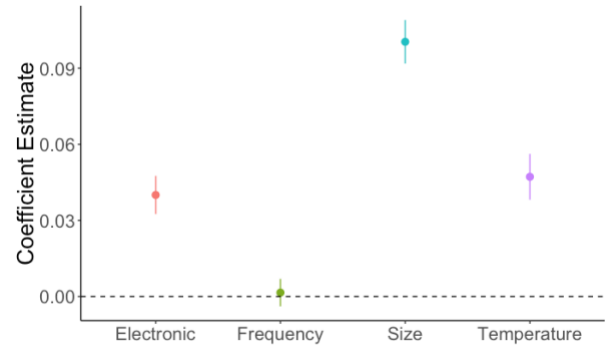
Individual differences in the use of proximal cues to estimate home energy use

We next investigated individual differences in the features that were associated with energy estimates — that is, we asked whether some people relied more on some proximal cues (e.g., size) than on others (e.g., temperature change).

	Size	Electronic	Frequency of use	Temperature Change
Size	1.00	0.183	0.066	0.215
Elect.		1.00	0.103	-0.131
Freq.			1.00	0.023
Temp.				1.00

Table 1: Correlation matrix of key features

Figure 1: Reliance on proximal cues to estimate energy use. Points indicate coefficient estimates from a mixed-effects model of energy estimates. Error lines indicate standard errors.
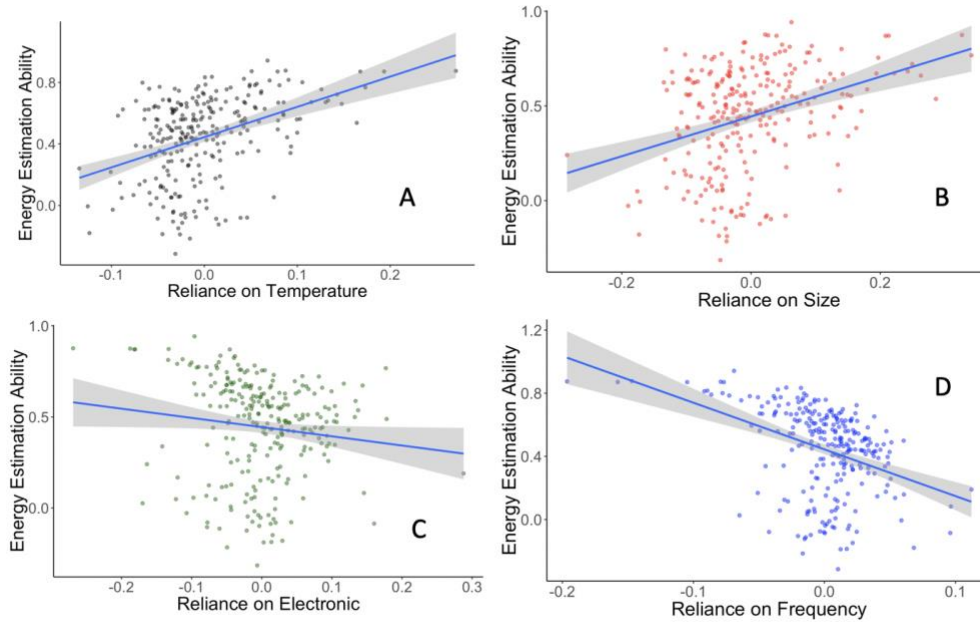

To capture these individual differences, we used the random by-participant slopes from our mixed effects model of energy estimates; for each participant, therefore, we had four random slopes, one for each feature (size, frequency-of-use, temperature change, and electronic-ness). Positive values of these random slopes indicate that a participant relies on that feature more than the group average; negative values indicate that they rely on that feature less than average.

In general, there was considerable variability in how strongly these features were associated with individuals' energy estimates (Fig. 2, panels A, B, C, and D). Some individuals' energy estimates were explained primarily by their judgments of the frequency of an appliance's use, despite the fact that frequency of use is a poor proxy for energy use. Others, however, appeared to ignore frequency and instead relied on temperature change, a reliable cue to energy use. Indeed, participants who relied more on temperature change tended to rely less on frequency of use (*R* = -0.60). Size and temperature change, both fairly good proxies for energy use, were highly correlated (*R* = 0.95), suggesting that people who use one feature to evaluate appliances' energy use are also likely to use the other.

All this together suggests that individual difference in the reliance on proximal cues might be associated with variability in how *good* people were at estimate home energy use. To quantify individual differences in estimation ability, we calculated, for each individual, the correlation between their estimates and the true energy used by each appliance. As predicted, participants who relied more on how much an appliance changed the temperature were also, overall, significantly better at estimating home energy use (b = 1.97 ± 0.27 SEM, p<.001); the same held for participants who relied more on the appliance's size, though to a lesser degree. Indeed, past work has found that lay people reliably underestimate the energy used by large appliances that heat or cool (Attari et al, 2010); here, our results suggest that there may be important variability in people's sensitivity to appliances' size and temperature change (Fig. 2A, 2B). By contrast, participants

Figure 2: Energy estimation ability as predicted by reliance on select features

who relied more on electronic-ness and frequency-of-use were overall worse at estimating home energy use (b = -0.51 ± 0.24 SEM, p<.05, b = -2.96 ± 0.43 SEM, p<.001). We also ran a correlation on the participants' reliance on each of these four features (Table 1). We found reliance on frequency of use and electronic-ness to be positively correlated, while frequency of use and temperature change were negatively correlated.

Characterizing the complex structure of the full appliance space

Finally, we combined ratings of all thirteen features (e.g., size, brightness, movement, etc.) to characterize lay perception of home appliances. To do so, we used multi-dimensional scaling (MDS). This technique takes the similarity between paired appliances and uses that to generate a reduced dimensional representation that captures how similar or different appliances are to each other. This approach gets at the rich structure that exists in how people perceive appliances as varying along multiple dimensions, many of which covary with each other. This approach is also necessary, because when dimensions are treated as independent, classic approaches like multiple regression do not account for collinearity of dimensions.

The two-dimensional MDS solution is illustrated in Figure 3. Note the rich structure that emerges bottom-up from this approach, with some appliances clumping together into meaningful groups, with related appliances clustering together into meaningful categories. We used k-means clustering (k=8) to capture these categories (Fig. 2). For example, all the light-bulb appliances (i.e.
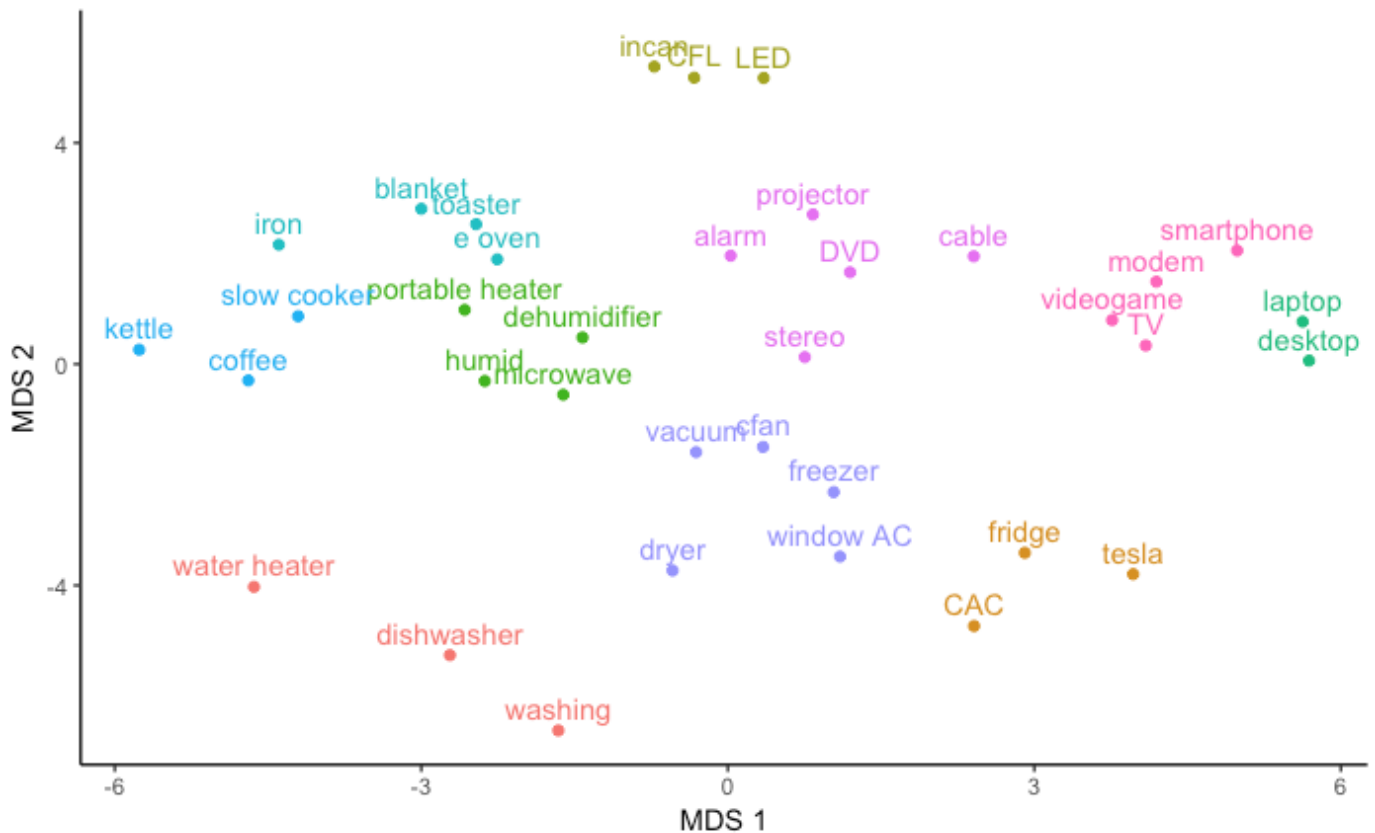
incandescent lightbulbs, Compact Fluorescent Light bulb, and LED bulb) group together because people rated those appliances very similarly.

While this MDS solution can characterize people's mental representations of appliances, it is blind to people's estimates of the appliances' energy use. However, when we regressed the MDS dimensions onto estimation ability, we found both MDS axes were related significantly to energy estimates increase (dimension 1: b = 146.88 ± 67.5 SEM, p<.05; dimension 2: b = -254.68 ± 104.0 SEM, p<.05). This was true despite the fact that these MDS dimensions combine multiple experiential features in complex, non-linear ways. Thus, lay people have structured perceptions of appliances, and these perceptions seem to relate systematically to their perceptions — and misperceptions — of their energy use. Future work should try to leverage this to improve energy decisions and behaviors.

Discussion

We began by asking how it is that people are able to estimate the energy used by appliances, when that energy use is often hidden. We found that estimates of appliances' size accounted for most of the variance in people's energy estimates. People relied, to a lesser extent on temperature change and how "electronic" an appliance, but they did not rely on frequency of use as a cue. Previous literature has claimed that all these features matter. Our results put those findings in a new light because we found that size is the primary driver of energy estimates. Since these replacement cues correlate, previous findings such as 'people use frequency of use as a replacement

Figure 3: Two-dimensional MDS solution for home appliance space

heuristic' might indicate that people tend to use bigger appliances more often. Interestingly, people relied more on size than heat, despite heat being a better indication of energy use. Heating (and cooling) both take a lot of energy but are perhaps not as obvious to people because the energy used to heat (and cool) are often used to achieve homeostasis. Your heating bill is high in the winter because so much energy has to be exerted to maintain your home at a constant temperature.

When we examined individual differences in the reliance on these cues, we found that the degree to which people relied on certain features predicted how good their energy estimates were. People who relied more on temperature change had better energy estimates than people who relied more on size, or any of the other theory-driven features used in our model. The more participants relied on how "electronic" an appliance was or on frequency of use, the worse their energy estimation ability was. When we ran a correlation on individual differences of reliance, we found that reliance on frequency is negatively correlated with reliance on temperature change. We also found that reliance on frequency is positively correlated with reliance on electronic-ness. This suggests that teaching people to use these more reliable cues may have benefits for energy judgments and decisions (Marghetis et al., under review).

Using multi-dimensional scaling, we also sought to characterize the public's mental representation of home appliances. This bottom-up approach found significant structure in people's perceptions of appliances; moreover, this two-dimensional representation was related systematically to people's energy estimates. In Fig. 3, the upper-left quadrant of the graph seems to include all the appliances that heat water, while the lower-left quadrant includes the appliances that heat without water. This suggests that this two-dimensional MDS solution has picked out *heat* as a notable component of one of its major axes. The appliances near the top of Fig. 3 are quite small and increase in size as you go down the MDS 2 axis, suggesting that this MDS solution has picked out size as a major component of its other axis. It is quite notable that even just a two-dimensional solution has, in a bottom-up way, picked out the two most useful and frequently used replacement heuristics. The clustering as shown in Fig. 3, also created through the bottom-up k-means algorithm, is quite remarkable as well. Kitchen appliances that heat water have clustered together on the left (blue); devices that are electronic or involved in entertainment have clustered together on the right (pink and green); appliances that heat or cool and move air around have also clustered together in the middle of the figure (purple). These clusters suggest that this MDS solution is a fruitful way to access the internal structure of people's complex perceptions.

Conclusion

We set out to answer three main questions. The first was 'Which features are people relying on to make energy estimates?' The answer to this is not simple. Our MDS solution shows that people rely on a complex and correlated set of proximal features. However, when comparing a smaller set of theoretically important features, size far outstrips any of them. Among the features we examined, people seem to rely most on size, even though it is not the best indicator of energy use. The best indicator of energy use was heat or temperature change.

We also set out to answer how individual differences in cue-weighting relate to estimation skill. Fig. 2A shows that as people rely on heat as a cue, their estimation skill improves. This is true to a lesser extent of size as well (Fig. 2B). As people rely on how electronic an appliance is, or how frequently it is used, their estimation skill decreases (Figs. 2C, 2D).

Finally, we set out to capture people's feature representation of appliances in a way that accounts for correlations among features. With an MDS solution, we found that meaningful clusters of appliances emerge, even from bottom-up clustering methods, and that the dimensions of this representation were related systematically to estimates of energy use.

This study speaks to previous energy literature that has attempted to identify the most predictive cue of people's energy estimates. By looking at several cues at once while accounting for correlations, we can say with confidence that despite the many, many features to choose from, the size of an appliance matters to people.

People do rely on the superficial cues about energy that they have access to. It is important to understand which of these people most rely on, so that we can more deeply understand how people understand and choose to use energy. Good energy choices can be encouraged in a variety of way, including but not limited to top-down policies, market-based incentives, extensive educational programs, home energy audits, and new home technologies. For example, in-home smart devices that give real-time feedback on energy-use can encourage energy conservation (Darby, 2006; Delmas et al., 2013). But implementing effective climate policies is politically difficult (Dietz, Ostrom, & Stern, 2003), home audits require time and resources that make scaling up nearly impossible, and new in-home energy technologies may be years away from mainstream use.

By understanding, and eventually changing either the cues people have access to, or their perceptions, we hope to encourage better ways of communicating energy information and making possible good and widely usable energy consumption habits.

References

Attari, S. Z., DeKay, M. L., Davidson, C. I., & de Bruin, W. B. (2010). Public perceptions of energy consumption and savings. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(37), 16054–16059.

Brunswik, E. (1956). Perception and the Representative Design of Psychological Experiments. University of California Press.

Cowen, L., & Gatersleben, B. (2017). Testing for the size heuristic in householders' perceptions of energy consumption. Journal of Environmental Psychology, 54, 103–115.

Darby, S. The effectiveness of feedback on energy consumption: A Review for DEFRA of the Literature on Metering, Billing and direct Displays. 24 (2006).

Delmas, M. A., Fischlein, M. & Asensio, O. I. Information strategies and energy conservation behavior: A meta-analysis of experimental studies from 1975 to 2012. *Energy Policy* **61**, 729–739 (2013).

Dietz, T., Ostrom, E. & Stern, P. C. The struggle to govern the commons. *Science* **302**, 1907–1912 (2003).

He, H. A., Greenberg, S., and Huang, E. M. One size does not fit all: applying the transtheoretical model to energy feedback technology design. In Proc. CHI '10, ACM (2010), 927–936.

Heller, Konstantinos, Borchers. Counter Entropy: Visualizing Power Consumption in an Energy+ House. In CHI '13: Extended Abstracts of the 2013 ACM annual conference on Human Factors in Computing Systems, April 2013.

Kahneman, D., & Frederick, S. (2002). Representativeness Revisited: Attribute Substitution in Intuitive Judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), Heuristics and Biases (1st ed., pp. 49–81). Cambridge University Press.

Lesic, V., Bruin, W. B. de, Davis, M. C., Krishnamurti, T., & Azevedo, I. M. L. (2018). Consumers' perceptions of energy use and energy savings: A literature review. Environmental Research Letters, 13(3), 033004.

Marghetis, T., Attari, S., and Landy, D., (Under Review). Correcting misperceptions of home energy use.

Schley, D. R., & DeKay, M. L. (2015). Cognitive accessibility in judgments of household energy consumption. Journal of Environmental Psychology, 43, 30–41.

Appendix: Features

1. How **big** is each appliance?
2. How **long** is each appliance typically used?
3. How much **light** does each appliance produce?
4. How much does each appliance **heat** itself or its environment?
5. How **loud** is each appliance?
6. How much **water** does each appliance use?
7. How much does each appliance **cool** itself or its environment?
8. How big is the **motor** of each appliance?
9. How much does each appliance **heat water**?
10. How complex is the **software** each appliance runs?
11. How **electronic** is each appliance?
12. How **mechanical** is each appliance?
13. How much does each appliance **move** itself or its environment?
14. How **frequently** do you use each appliance?