

Creating an in-home display: experimental evidence and guidelines for design

Tamar Krishnamurti*, Alexander L. Davis, Gabrielle Wong-Parodi,
Jack Wang, and Casey Canfield

Department of Engineering and Public Policy
Carnegie Mellon University

November 4, 2012

Abstract

In-home electricity displays (IHDs) are digital devices that can give near-real-time information about electricity usage in the home. These devices have the potential to provide the kind of personalized feedback necessary to effect behavioral change among residential consumers. We present an approach to in-home display design that uses research on customer preferences to determine which features to experimentally test, while comparing preferences against experimental data to determine whether people have insight into what works. Using a computer-based simulated IHD, we found that the types of feedback information that consumers prefer (appliance-specific and dollar-feedback) are not as effective for learning than kWh feedback. Moreover, it appears that a simpler more generalized format of information provision has the potential to be more effective than a personalized IHD for knowledge acquisition. We discuss the need to use consumer preferences to inform IHD design but not at the cost of systematic experimental testing. (word count: xxx)

Keywords: In-home displays; Electricity Feedback; Consumer preferences; Learning

*Corresponding Author. Phone: 1-412-445-2663. Fax: Email: tamar@cmu.edu. All materials and data, including completely reproducible statistical analyses in Sweave, can be obtained at <http://hdl.handle.net/1902.1/18699>. Open lab notebook can be obtained at http://openwetware.org/wiki/IHD_simulation : Notebook/IHD_Field_Trial; Applied Energy; Energy; Energy Policy; Journal of Environmental Psychology; Energy Conservation and Management; Energy and Buildings; We thank Jay apt...This material is based upon work supported by the Department of Energy under Award Numbers DE-OE0000300 and DE-OE0000204. Disclaimer: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights, Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

To be effective, any information supplied to residential electricity customers must be adapted to their needs, especially for those who have limited knowledge of electricity. Unfortunately, what is provided to U.S. households by electric utility companies does not meet this criterion. Most people receive their monthly bill, scan it to identify the amount they owe, and then discard it. This is partly because the information on the bill is in a format that is too dense or complicated to decipher [1]. For example, usage information is typically presented only in kilowatt-hours (kWh), a unit that has been found to be opaque to customers [2]. When the bill is supplemented with inserts or other offerings, they are perceived to be a scam or gouging opportunity, and are usually ignored [3]. Without the right information that promotes understanding, customers cannot effectively conserve energy.

In the absence of available and usable information, customers will create ‘folk theories’ or mental models of how electrical devices function and use energy [4, 5]. Incorrect theories may encourage them to engage in actions that can be particularly wasteful of energy [6, 7]. Take, for example, the common folk theory about thermostats, the “valve” theory, which holds that the quantity of cooling or heating in the home is proportional to the thermostat setting. This would lead customers who want to cool their household to set their thermostat lower (e.g., to zero Celsius) hoping for faster cooling, but instead just waste energy when the air conditioner cools too much. Without easily interpretable information that corrects these folk theories, many customers would not adopt appropriate measures to save electricity, even if they wanted to.

Recognizing the problem, researchers and utility companies have tried providing customers with in-home electricity displays (IHDs) that can give near-real-time information about electricity usage. One of the earliest examples of a simple, and particularly effective, IHD was that used in the Twin Rivers study [8, 9]. In this study, participants were given a simple light that flashed blue when one could cool the home by opening the windows rather than using the air conditioner. This display yielded almost a 20% reduction in monthly electricity use over the short duration of the study. Paired with the advanced meters (‘smart-meters’) of the smart grid, more sophisticated IHDs can provide customers high resolution feedback about their electricity consumption.

Since the seminal Twin Rivers study, mounting evidence has shown that IHDs can substantially help customers curtail their electricity use. A variety of displays have been used in these field studies, including retail (e.g., the Power-Cost Monitor) and custom devices (e.g., The Residential Energy Cost Speedometer; [10]). Each IHD provides different types of feedback information (e.g., kWh use, cost of electricity, monthly spending), in different formats (e.g., graphs, tables, numbers, visual-analogs). In a recent review of these field trials, Davis et al. (2012) found four custom displays (Bluelight, [8]; RECS [10], Fitch [11] and Electricity Consumption Display [12]) to be the most effective for reducing overall consumption ($\sim 20\%$, $\sim 13\%$, $\sim 12\%$, $\sim 12\%$, respectively). Thus, it appears that custom designed IHDs can provide the right information in an easily understood manner, leading to effective reductions in electricity use.

While these findings are encouraging, the small sample sizes of these studies ($N = 20, 99, 101$, and 8 , respectively) calls into question the real-world efficacy of these displays. Casting further doubt, studies of popular retail devices, that have much larger sample sizes, such as the TED and PowerCost Monitor, show much smaller effect sizes (e.g., Norton et al, 2008). Moreover, once adjusted

for plausible bias in study design, effects are substantially reduced (Davis et al, 2012). While studies that demonstrated large versus small effects differed in many ways, an important difference was the type of IHD they used, leaving open the possibility that some displays were more effective than others by design. In addition, the existing IHD field studies vary so much in experimental methods, reporting, and features of the IHDs, it is impossible to pinpoint what made one study more successful than another.

To determine the most effective display, one would ideally test each display in the field with a large enough sample size to detect even small effects. However, this approach is both time-consuming and prohibitively expensive, forcing utilities to guess which display would suit their customers’ needs. In this paper, we present findings from a simple simulated computer-based in-home display, which allows us to test display features individually and measure their effectiveness as educational tools.

Past research on customer preferences for IHD features has used interviews, surveys, and other similar approaches (e.g., focus groups). The options participants generated or could choose from have generally fallen into five categories [13, 14]: *Units*, *Time Aggregation*, *Physical Aggregation*, *Comparators* and *Format*.¹

Overall, the available literature shows some agreement on what customers want in terms of units and comparators, while there are fewer findings on time aggregation, physical aggregation, and format. However, relying on preferences to infer how customers will behave is risky, as “what people think they want and what they actually want are not always the same” [15]. This is consistent with more basic psychological research, showing that customers are not always good at predicting what they will like, concentrating too much on changes [16] or showing bias toward their present feelings [17]. People also have been known to reject policies in prospect, but like them once implemented [18]. Thus, examining preferences alone may give a certain, but potentially incorrect, perspective of how an in-home display should be designed to be most effective.

We suggest a complementary approach that uses research on customer preferences to determine what features to experimentally test, while comparing preferences against experimental data to determine whether people have insight into what works. To date, little experimental work (with the exception of enhanced bills, [19, 20]) has been conducted. The various field studies, interviews, and surveys have neither separated specific elements of IHDs according to their effectiveness, nor measured important intermediaries of effectiveness such as learning and motivation [3, 13, 21, 22, 23].

In the following section, we present new data on customer preferences for the attributes found on the most common commercially available IHDs. We then present findings of an experimental test of two units of information (\$ vs. kWh information) and two forms of aggregation (total vs. appliance-specific), using an IHD simulation. We conclude with a discussion of the implications of our findings, as well as how IHD simulations can usefully inform the development and testing of IHDs in the field.

¹A more exhaustive review of the specific attributes that fall into these categories may be found in Appendix A.

1 Survey

1.1 Materials

First, after a thorough review of all commercially available in-home displays, we developed a list of 19 displays available for direct purchase by residential homeowners. Next, we developed a list of the most common types of electricity feedback information provided by these 19 displays and in the existing literature on IHDs.

1.2 Procedure

Participants were first asked to rate this set of feedback types with the following instructions: “Here is a list of information that might appear on an in-home display. Please rate each type of information in terms of how much you would like to have it on the display.” All information attributes were rated from 1 (not at all) to 5 (extremely), with the additional option of responding “I don’t know.”

Participants then created their own display by selecting what features they wanted from those listed above. They also responded to the following questions about the display they created, “How much would you like to have the in-home display you created on the previous page? (1 = not at all; 5 = extremely),” “How effective do you think that in-home display would help you to reduce your electricity use? (1 = very ineffective, 7 = very effective),” “How often would you look at the in-home display you created? (1 = never, 7 = daily),” and “How much do you think you would save, in dollars, on your monthly electricity bill if you had the display you created on the previous page?” The full materials are described in Appendix B.

1.3 Participants

Participants were bill-paying electricity customers ($N = 151$) recruited using the Amazon Mturk system. The average age was 31.5 years old ($SD = 10.9$), with 42% being male, and most having an income between \$51K and \$75K per year. The average electricity bill among these customers was \$107/month.

1.4 Results

In general, people prefer an IHD (mean rank = 1.61) to receiving information on a computer (mean rank = 1.93) or smart-phone (mean rank = 2.10). The average willingness-to-pay for an IHD was \$150.

1.4.1 Feedback Information Preferences

Table 1 presents the feedback information types in order of preference. A Wilcoxon signed-rank test was used to assess whether participants ranked one attribute higher than another. Information presented as “bill-to-date” and “appliance-specific” feedback were considered the most desirable IHD features, although neither was preferred over the other according to the Wilcoxon signed ranks test ($Z = 1.6$, $p = 0.11$). The least preferred way of presenting information

was comparison to a “similar household.” This attribute was rated much less desirable than any of the other attributes ($p < 0.001$).

Table 1: Feedback Information Preferences

Feedback Information Type	Mean	SD
Bill-to-date [IA1]	4.11	0.94
Appliance-Specific [IA2]	4.00	0.96
Daily Projections [IA3]	3.83	0.99
Monthly Projections [IA4]	3.80	0.96
kWh-to-date [IA5]	3.80	0.96
Daily price [IA6]	3.70	1.16
Daily peak use [IA7]	3.69	1.04
Monthly peak use [IA8]	3.68	1.08
Goal tracking [IA9]	3.61	0.96
Greenness [IA10]	3.40	1.10
Similar Household [IA11]	2.87	1.20

1.4.2 Create-Your Own

Participants’ constructions of their own IHD were almost identical to their ratings of individual feedback information types, so we omit them here. On average participants expected to like the display they’d created strongly, as seen in the mean rating being significantly above the scale midpoint of 3.0 in a one-sample t-test, $t(x) = x$, $p < .001$. They thought it would be effective way for them to reduce their electricity use with a mean judged effectiveness significantly above the scale midpoint of 4.0, $t(x) = x$, $p < .001$, and expected to look at it on average 2.3 times a week ($M = 6.11$, $SD = 1.1$, $N = 138$). They anticipated an average savings on their monthly bill of \$25 ($SD = 29$, $N = 131$), with the average monthly electricity bill being \$106 ($SD = 81$, $N = 106$).

1.5 Discussion

Bill-to-date and appliance-specific feedback stood out as the two most preferred types of information. This finding is consistent with previous research on preferences for IHD content [2, 14, 15]. There was little differentiation between other types of information, such as projections or goals. However, the more ‘gimmicky’ features, greenness and social comparisons, were strongly disliked. Participants had unreasonably high expectations for monthly savings from the IHD (about 25% of their monthly bill), given that the average monthly savings from IHDs has never been above 20% in previous research (Davis et al, 2012).

Applied energy cites XX.

Our survey suggests that people want to be able to simply look at an IHD and obtain easy-to-interpret feedback information to help them manage their monthly budget and control their current consumption. Our next study draws on these preferences to test how well people learn from what they say they want. Specifically, our study and previous research strongly suggests that people prefer bill-to-date information and appliance-specific feedback in units of \$. However, do they actually benefit from this configuration, or is it possible to get the same

benefit using the simpler to implement kWh and bill-to-date configuration? The experiments described in the next section try to answer this question using a simulated in-home display.

2 Simulated In-Home Display Experiment

2.1 Methods

Participants interacted with a computer-based in-home display simulation that allowed them to turn the appliances on and off, change the settings on various thermostats, and alter the length of use for each in 30 minute increments in the house of a fictional family, “The Smiths.” Feedback information on electricity use was updated according to these manipulations and presented in a tabular format, as seen in Figure 1.

We selected the 11 most commonly owned electricity consuming home appliances in the U.S. for participants to manipulate the use of (ref): (1) air conditioner, (2) water heater, (3) indoor lights, (4) outdoor lights, (5) refrigerator, (6) freezer, (7) oven, (8) microwave, (9) television, (10) washing machine, and (11) dryer. We used standard estimates for the real-time electricity consumption values, as well as used the standard current estimate for the average cost of electricity in the US (ref). These values for the 11 appliances were used as inputs for our simulated IHD, and thus were reflected in the values seen by our participants.

2.2 Experimental Conditions

Participants were provided the same electricity knowledge test, before and after interacting with the simulated IHD. There were six in-home display simulations that each provided participants with different information when they manipulated the appliances. The specific conditions were designed to test the effectiveness on knowledge acquisition of the information that people reported wanting to see (appliance-specific feedback in \$ units) vs. the information that would be easiest for a utility to provide (aggregate kWh usage). Details of each condition can be seen in Table 2. Knowledge acquisition was assessed on ten out of eleven appliances.²

We also included a control condition simulating more generic educational materials in which, customers are simply told to use less electricity, or are provided with information on how much electricity appliances use in general [6]. In our *passive learning* control condition, participants were given the answers to the electricity knowledge pre-test and then asked to provide those answers in the post-test. This type of passive, as opposed to discovery, learning can come from a variety of sources such as “educational campaigns, advertisements, advisory services and news media” [2].

Both theory and practice suggest that that passive learning is not sufficient for behavior change, although it may be more effective in cases when behav-

²Although participants were able to manipulate and received feedback on the electricity consumption of 11 appliances, pre-testing showed that participants felt much more comfortable with 10-item rankings. The difference in usage between indoor and outdoor lights were marginal compared to the other appliances, therefore outdoor lights were removed from the ranking tasks.

Appliance	Status	Hours	
 Air Conditioner	 78.0 ° F 	 0.00 	Reset
 Water Heater	 120.0° F 	 0.00 	Reset
 Indoor Lights	<input type="checkbox"/> OFF	 0.00 	Reset
 Outdoor Lights	<input type="checkbox"/> OFF	 0.00 	Reset
 Refrigerator	 38.0 ° F 	 0.00 	Reset
 Freezer	 -15.0° F 	 0.00 	Reset
 Oven	<input type="checkbox"/> OFF	 0.00 	Reset
 Microwave	<input type="checkbox"/> OFF	 0.00 	Reset
 Television	<input type="checkbox"/> OFF	 0.00 	Reset
 Washing Machine	<input type="checkbox"/> OFF	 0.00 	Reset
 Dryer	<input type="checkbox"/> OFF	 0.00 	Reset
		Total kWh: 0.0000	Reset All

Figure 1: Screenshot of the simulated in-home display.

ior change is convenient and cheap (low barriers) [18]. However, here passive learning serves as a control by establishing whether participants can process the knowledge areas that are being tested given that aspects of this topic may be something that participants have had little or no prior exposure to.

2.3 Participants

Participants were bill-paying electricity customers ($N = 191$) recruited using the Amazon MTurk system. Forty-two percent were male, most participants were between 22 and 34 years old. Their income ranged from X to XK per year, with most participants having an income of x per year. The average electricity bill among these customers was \$138/month.

2.4 Measures

Learning from the simulation was measured by assessing participants' rankings of the ten appliances in terms of how much kWh the appliance used in a 10-minute period as well as their rankings of these in terms of their contribution to the monthly bill. For example, participants were asked to "Imagine each appliance listed below is used for exactly the same amount of time (10 minutes). Rank the ten appliances below by how much electricity they use from 1 (the most) to 10 (the least)." The specific text of these questions may be found in

Table 2: Description of feedback information for the 6 conditions and 1 control group.³

Group Condition	Feedback Information
kWh only	Participants were shown kWh feedback information used for running the appliances at the settings they made for the duration they specified.
\$ only	Participants were shown cost (\$) feedback information used for running the appliances at the settings they made for the duration they specified.
\$ and kWh	Participants were shown both kWh and cost (\$) feedback information used for running the appliances at the settings they made for the duration they specified.
\$ by appliance	Participants were shown kWh feedback information in total, as well as for each appliance, used for running the appliances at the settings they made for the duration they specified.
kWh by appliance	Participants were shown cost (\$) feedback information in total, as well as for each appliance, used for running the appliances at the settings they made for the duration they specified.
\$ and kWh by appliance	Participants were shown kWh and cost (\$) feedback information in total, as well as for each appliance, used for running the appliances at the settings they made for the duration they specified.
Passive Learning	Participants were given the answers to the electricity knowledge pre-test and then asked to provide those answers in the post-test.

Appendix C. Additional measures and results not discussed for brevity are in Appendix D.

2.5 Procedure

Participants completed the ranking questions (pre-test) prior to viewing the IHD simulation. They were then randomized to one of the seven conditions previously mentioned. Interaction with the simulation (and viewing the passive learning information) lasted for as long as they wanted. After interacting with the simulation, they completed the same ranking questions (post-test). Lastly, they completed demographic questions.

2.6 Results

2.6.1 Level of Interaction

Participants spent a median of 34 minutes ($SE = 3.2$ minutes) interacting with the IHD simulation. Time spent interacting with sim did not differ by conditions (all ts less than 1.6).

2.6.2 Appliance Rank Deviations: kWh Use

Figures 2 and 3 show pre-post differences in rankings of appliances by hourly energy consumption, aggregating across conditions. As can be seen from the histograms, with red lines indicating the true ranking of the appliance and dashed black lines indicating statistically significant shifts toward the correct ranking at post-test, deviations from the true rank improved for AC, Microwave, and Fridge ($Z = 2.6$, $p = 0.01$; $Z = 4.8$, $p < 0.001$, and $Z = 3.1$, $p = 0.0021$, respectively). There were no overall differences for the dryer, oven, water heater, washer, freezer, TV, and indoor lights, ($Z = 1.6$, $p = 0.11$; $Z = 0.043$, $p = 0.97$; $Z = 1.1$, $p = 0.27$; $Z = 1.4$, $p = 0.16$; $Z = 0.7$, $p = 0.48$; $Z = 1.9$, $p = 0.054$; and $Z = 0.63$, $p = 0.53$, respectively).

Participants were very accurate in ranking the highest and lowest use appliances, the AC and indoor lights, both before and after interacting with the simulated IHD. Those appliances that participants were initially least sure about (as illustrated by the scattered nature of their initial rankings), the water heater and the oven, did not improve after interacting with the simulated IHD. The most striking improvement was for the microwave, an appliance that participants had certain but incorrect initial beliefs about.

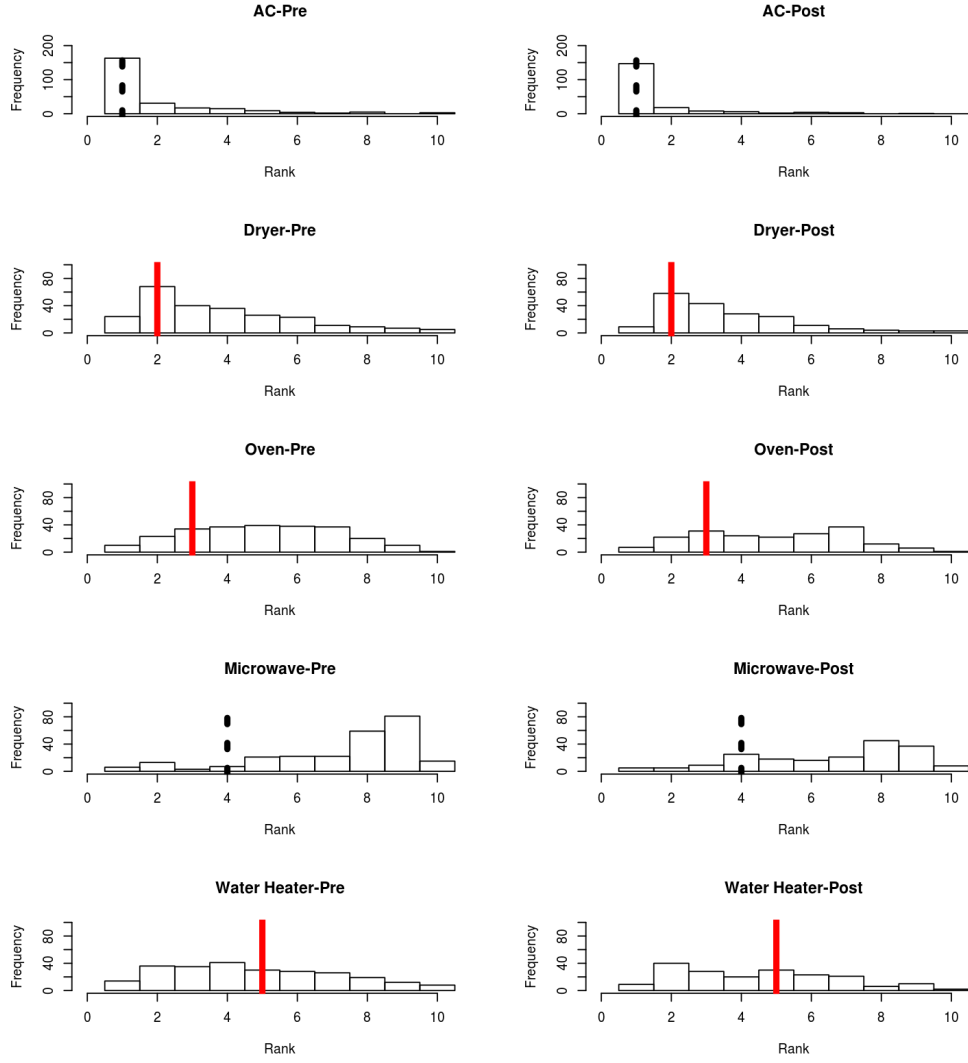


Figure 2: Histogram of appliance use rankings for each appliance in order of true ranking. Lines indicate true rank, with black dashed lines indicating statistically significant pre-post differences.

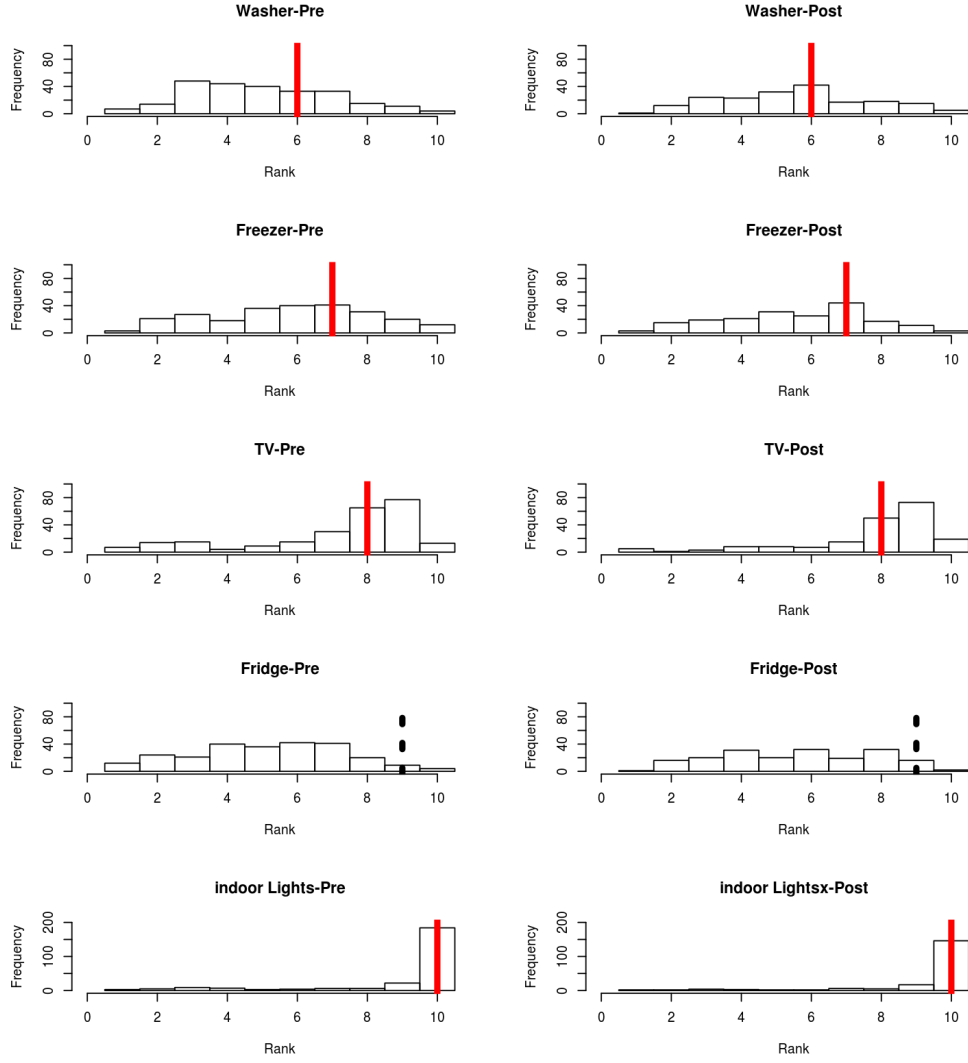


Figure 3: Histogram of appliance use rankings for each appliance in order of true ranking. Lines indicate true rank, with black dashed lines indicating statistically significant pre-post differences.

2.6.3 Appliance Rank Deviations: Cost

Figures 4 and 5 show pre-post differences in rankings of appliances by monthly contribution to electricity bill, aggregating across conditions. Results for cost rankings are similar to those for kWh use. Participants improved from their initial rankings across conditions for the washer and freezer ($Z = 3.4$, $p < 0.001$ and $Z = 3.5$, $p < 0.001$, respectively). Their rankings at post-test were worse than their pre-test rankings for the oven, water heater, and microwave ($Z = 0.28$, $p = 0.78$; $Z = 1.6$, $p = 0.12$; $Z = 1.1$, $p = 0.29$, respectively). For the AC, lights, TV, Dryer, and Fridge, participants were initially quite accurate in their rankings and either remained accurate or slightly improved ($Z = 0.029$, $p = 0.98$; $Z = 0.59$, $p = 0.56$; $Z = 1.6$, $p = 0.12$; $Z = 1.1$, $p = 0.29$; $Z = 0.62$, $p = 0.54$, respectively).

Contrary to the findings for kWh use, the post-simulation rankings for the microwave were more inaccurate than the pre-simulation rankings. However, unlike their precise but incorrect initial beliefs about how many kWhs it used in 10 minutes, participants had imprecise beliefs about how much the microwave would cost them in a month. Participants learned that the microwave used much more energy than they expected, but they seemed to incorrectly extrapolate this greater energy use to monthly cost, not taking into account how infrequently the microwave is used compared to other appliances. More detailed analyses can be found in Appendix F.

2.6.4 Appliance Rank Deviations by Specific Condition

Following the aggregate analysis, we then looked at how the rankings of each appliance changed for each treatment condition individually. As can be seen in Table 3 and X, a random effects model [24, 25] with different intercepts for each appliance examines the effect of each treatment condition aggregating across appliances, weighting them appropriately. Overall, more improvements in ranking accuracy occur in conditions in which kWh rather than \$ units or appliance-specific information are provided. Moreover, participants appear unable to extrapolate from monthly cost feedback to monthly kWh use. The appliance specific information does not appear to contribute to improvements in ranking accuracy. Detailed analyses can be found in Appendix E.

Table 3: Hierarchical linear model with varying-intercepts for appliance and participant.

Condition	Improvement	t-Statistic	p-value
kWh	0.12	0.75	0.3
AS	-0.095	-0.73	0.31
\$	-0.18	-1.1	0.22
Passive	0.65	2.5	0.019

2.7 Discussion

Participants were able to learn how much energy their appliances used from the simulation. Learning was particularly pronounced for appliances that they had preconceived but incorrect notions about, such as the microwave, which

used much more energy than they expected. Learning was more modest or non-existent for appliances that they knew little about (the water heater and the oven) beforehand. Participants also demonstrated high accuracy in identifying which appliances used the most (air conditioner) and least (indoor lights) out of the ten.

Although the effects were small, participants learned more about how much energy the appliances used when provided with kWh feedback, regardless of whether they were also provided feedback in dollars or by appliance. Thus, participants may have difficulty focusing their attention when provided with a large tabular display that provides kWh information on each appliance. This type of ‘information overload’ has been found elsewhere, especially with respect to the well-established literature on working-memory (cite). This literature suggests that people have the capacity to consider roughly three ‘chunks’ of information at one time. The ineffectiveness of dollar feedback indicates that participants had difficulty translating feedback provided in monetary units to energy units. These two results are important because people overwhelmingly prefer bill-to-date (in dollars) and appliance-specific feedback, but we find no evidence of their effectiveness in learning.

In contrast to participants’ ability to learn the energy usage of each appliance, participants were largely unable to learn how much each appliance cost in a month. The finding that participants who became more accurate in their knowledge of how much energy the microwave used, but less accurate in their knowledge of how much it costs per month provides a window into the learning process. It seems as though participants were basing their monthly cost estimates on their energy ranking estimates, failing to take into account how often each appliance is used. This explains the opposing effects for the microwave, as learning that it uses more energy would also lead one to overestimate how much it costs per month. This finding suggests that people need projected monthly costs for their appliances, as they have difficulty extrapolating from current cost and energy use to monthly cost. We found that these projected monthly costs are a feature that participants find desirable, but it is not at the top of their list (it ranked fourth below bill-to-date, appliance-specific feedback, and daily projections).

Perhaps the most surprising finding from the experiment was the consistent success of the passive learning condition. This condition merely provided participants with the answers to the knowledge test, which they were able to remember and subsequently regurgitate. In the real world it is much cheaper and more feasible to just provide a flyer with information about appliance-specific energy consumption, monthly cost, and how people should adjust their behavior to save energy. This cheap and easy solution may not make consumers more knowledgeable about their own electricity consumption, but may induce as much or more behavioral change as an IHD, using a much simpler method. However, the external validity of this finding is questionable for two reasons. First, participants likely did not develop the working knowledge of their appliances needed to extrapolate to new appliances, and thus any benefit from deeper comprehension and associated feelings of mastery and control over ones environment will not emerge. Second, people routinely receive flyers and inserts, which they dispose of without viewing, so this method of providing information may not be useful in the real world because there is saturation in this medium.

3 General Discussion

– a paragraph of set-up, reminding readers of the motivation, why you chose the task, and how you innovated in with it. – a reprise of the results, including your contributions to the literature. – a paragraph on limits and needed research – 2-3 paragraphs (max) on practical implications if these results hold (after noting the differences between the experiments and the world).

Maybe we can say something about how important the extremes (particularly on the heavy energy use side are). Perhaps we can also say something about how much getting the ranking right for the appliances in the middle matters.

We need to be clear about what we mean by behavior change. I think what our results allow us to say is that we are testing the configuration of attributes that leads to: enhanced understanding about the relative costs and energy use of appliances enhanced understanding of the most effective ways to save energy.

Paragraph of set-up, reminding readers of the motivation, why you chose the task, and how you innovated in with it. â a reprise of the results, including your contributions to the literature. â a paragraph on limits and needed research â 2-3 paragraphs (max) on practical implications if these results hold (after noting the differences between the experiments and the world).

Overall, the results suggest that people need to be provided with only an aggregate summary of their current kWh consumption, a projected aggregate or disaggregate monthly bill, or a flyer that will not be discarded that provides key information.

People seem to have the extremes down. The ones they have stronger beliefs about they seem to learn and the ones they’re initially confused on, they don’t learn.

Two explanations: 1) Familiarity frequency of use * look at data and see if their familiarity matches with our results 2) Familiarity via utility/campaigns * look at the correlation in rank by function (food functions, clothes function etc) Heuristic is same function and same use. It would make you really wrong with the dryer or washer if you are right about one and not the other yet you link them together. 3) Mental model of how it works

3.1 Future directions

: mental models of ‘diffuse’ appliances

First, we look at several appliances that participants seemed to disagree in their rankings about: oven, water heater, washer, freezer, and fridge. We also look at mental models of the microwave, that was consistently underestimated in rank. We want to know how they think these appliances work, and why they had such disagreement on their rankings (maybe do group discussions?).

Maybe people didn’t get the point of it, it was boring or whatever. Thus, this is a test of effectiveness rather than efficacy, which is good, as these are problems that will probably happen in the real world (baruch’s quote).

Follow-up studies and stuff for discussion section:

a) Intervention where we get them to explain how it works to get them to learn. Control intervention that says look at same appliance or think about same appliance without âexplainingâ

b) Provide them with new appliances and if they really understood the previous list of appliances, they should be able to extrapolate. If they don’t they

should look just like passive learning people. Passive learning people should be really crappy on extrapolating from one condition to the new appliances.

3.2 Conclusion

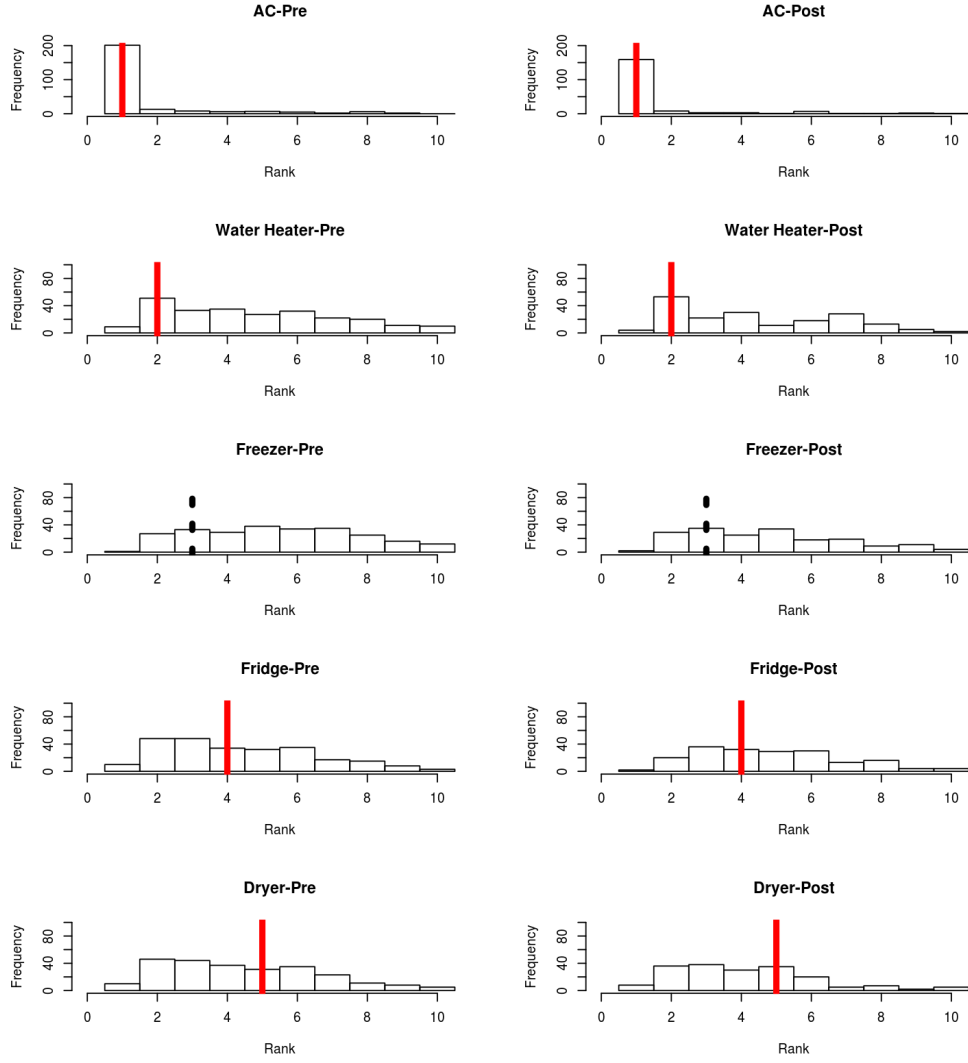


Figure 4: Histogram of appliance monthly cost rankings for each appliance in order of true ranking. Lines indicate true rank, with black dashed lines indicating statistically significant pre-post differences.

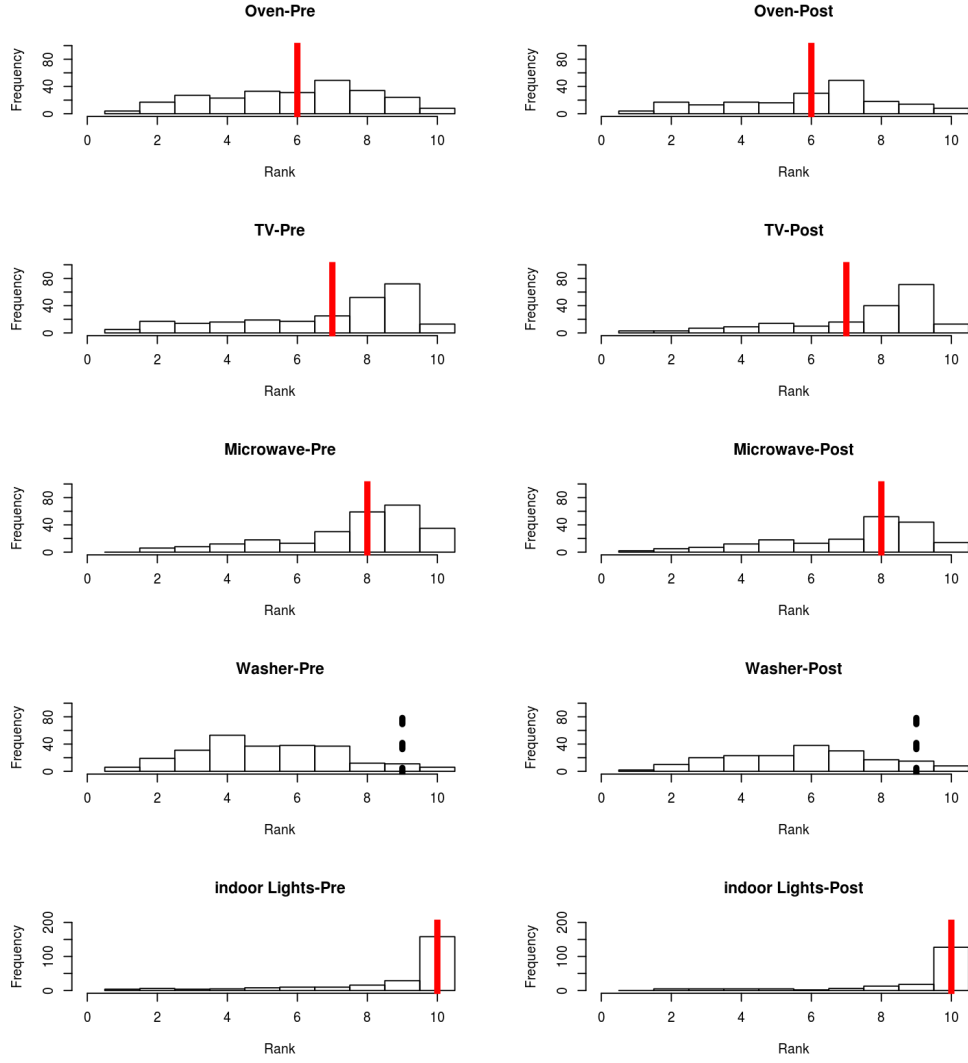


Figure 5: Histogram of appliance monthly cost rankings for each appliance in order of true ranking. Lines indicate true rank, with black dashed lines indicating statistically significant pre-post differences.

A Supplemental Review of Customer Preferences

A.1 Units

The information displayed on an IHD can come in different units, such as current cost (\$), cost/day, power (W or kW), energy (kWh), and carbon emissions (CO₂ tons). In general, people prefer the cost of electricity above all other possible ways to display electricity use [14, 15]. This is consistent with customers wanting simple information in units that they already understand. A number of studies have found that people want to see their costs either as current rate of expenditures (in \$/day) or cumulative cost in \$ per billing period [2, 15]. A few studies have found that people want to be able to switch from the cost of electricity to kWh when looking at the IHD, with one study finding that people prefer energy (kWh) over power (W or kW) [14].

A.2 Time Aggregation

The displayed information can come in increments of years, quarters, months, weeks, days, hours, real-time or other periodic cycles. Unlike preferences for units, there appears to be no consensus regarding people's preferences for how time should be aggregated. Some, for example, prefer to see their electricity consumption on an hourly basis [15], while others prefer to see it on a quarterly basis, compared to some reference point like the previous quarter [3]. Still others prefer to see their electricity use displayed as daily load curves [26] rather than ten-day curves [27]. However, while there is no unanimous preference for time-period, people generally want to be able to switch time periods with the press of a single button [2, 15]. While people are used to getting information monthly, when they receive their bill, more frequent information may be helpful [28].

A.3 Physical Aggregation

While we know of no research on whether people prefer electricity use information by room, by specific household member, or for the whole house, a recent study found that people strongly prefer appliance-specific information [14], and they want to see this information in dollar units [15].

A.4 Comparators

Comparisons typically examined have been to one-self (historically), to other customers (social comparisons), or to targets (goals). The most frequent finding is that people want to compare their current use to their own use at some point in the past [3, 14]. Moreover, people want to compare their personal electricity use to a self-set goal or target [2, 15]. In contrast, nearly all people express a strong rejection of social comparisons [3, 15, 29], wherein they see their electricity use compared to some other group of customers, such as their neighbors. Indeed, there is little evidence suggesting social comparisons motivate people to reduce their household electricity use [21].

A.5 Format

Displayed information can be formatted as a chart, picture, table, numerically, as text, or as a combination of audio and visual feedback. Like time aggregation, there is a lot of variability in preferences for format.

Visual analogs. Several studies have shown that people like visual-analog ‘speedometers’ that show cumulative use or rate of use [15]. This partially explains why many people rate the GEO minim, which uses this type of format in its display, highly ([2]; green energy options; <http://www.greenenergyoptions.co.uk/>).

Graphical displays. Customers usually prefer graphical displays in the form of bar graphs, although some people like bell-curves as well [22]. Karjalainen [14] developed eight paper-based IHD prototypes and evaluated them using preference assessments and think-aloud protocols. People understood bar charts, pie charts, and numerical tables easily. However, the researchers found that a tabular display was not only, “instantly understood by everyone,” but also preferred the most, as 7/15 participants ranked it then best of the eight prototypes.

Text, diagrams and colors. There is some evidence that combinations of texts and diagrams may enhance each other [22]. However, gimmicky pictures (e.g., stacks of coins, different sized houses) are strongly disliked [3]. Paetz [29] found that colors were useful for showing price information, as one participant wrote in their diary that “With a quick glance I was able to check the current price. It is quick and easy to comprehend with the three colors.”

A.6 Other

A few other attributes have been proposed as potentially proving effective in enhancing people’s understanding and use of information: (a) Motivation, including incentives, tariffs, and goals; (b) interactivity, allowing the user to determine what appears on the display dynamically; (c) grouping of information by room, type of appliance, subgroups, or hierarchically; (d) oracles, advisors, or intelligent tutors, that provide people tips, suggestions, answers, or even act on behalf of the customer.

B Survey Measures

Here is a list of information that might appear on an in-home display.
Please rate each type of information in terms of how much you would like to have it on the display.

1. How many dollars you've spent on electricity so far this month. [IA1]
2. How much electricity each of your appliances uses. [IA2]
3. How many dollars you will spend each day, if you keep using electricity at the same rate. [IA3]
4. How much electricity you will use by the end of the month, based on how much you've used so far. [IA4]
5. How many Kilowatt hours of electricity you've used so far this month. [IA5]
6. The daily price of electricity—if you are on a program where electricity prices differ each day. [IA6]
7. The time of day you use the most electricity. [IA7]
8. When you used the most electricity over the last month. [IA8]
9. How well you're meeting a goal that you've set for how much electricity you will use. [IA9]
10. How “green” your electricity use is each month. [IA10]
11. How your electricity use compares with that of a neighbor or a similar household. [IA11]

All special features were also rated on a 1-5 scale (not at all to extremely) or “I don't know.” Code for each measure is in brackets (e.g.,[SF1]). They were given the following instructions:

Below is an extensive list of special features that can be added to the in-home display. Please rate each feature on how much you would like it to be on a display.

1. You can set the in-home display to automatically control appliances in your home, such as your air conditioner. [SF1]
2. It allows you to pre-set a spending limit and alerts you when you are approaching or passed this limit. [SF2]
3. It warns you if a power outage (blackout) is likely. [SF3]
4. It helps you set a goal of how much electricity to conserve and tracks whether you're approaching your goal. [SF4]
5. It has graphics that change color when electricity prices are more or less expensive than usual. [SF5]

6. You can set it to provide you with personal alerts about how much electricity you've used or how much money you've spent on electricity that month. [SF6]
7. It provides movies and educational tutorials that explain how to use the display. [SF7]
8. It has a screensaver that is pleasing to look at when you are not using the display to access information. [SF8]
9. It provides you with audio feedback as well as visual feedback (for example, it might talk to you or sound alarms). [SF9]
10. You can upload and view digital photographs on the display. [SF10]
11. You can read blogs and access the internet on the display screen. [SF11]
12. You can play games on the display. [SF12]
13. You can access social networking sites (e.g., Facebook) on the display. [SF13]

B.1 Special Features

The rankings of the special features can be seen in Table 4. None of the top six features (automatic appliance control, spending alerts, blackout notice, set and track goals, changing colors, and monthly alerts) differed in important ways. The second tier of features were educational tutorials, audio and visual feedback, and a pleasing screensaver. The least preferred features were ones that would make the display multi-purpose, such as photographs, use of blogs and the internet, playing games, and Facebook. These were all rated significantly lower than the lowest of the second tier of features (audio/visual feedback), $V = 3489$, $p < 0.001$.

Table 4: Special Feature Preferences

Item	Mean	SD
Automatic appliance control [SF1]	3.81	1.09
Spending alerts [SF2]	3.75	1.04
Blackout notice [SF3]	3.75	1.15
Set and track goals [SF4]	3.71	1.08
Changing color [SF5]	3.63	1.12
Monthly alerts [SF6]	3.60	1.07
Tutorials [SF7]	2.97	1.29
Screensaver [SF9]	2.83	1.34
Audio/Visual [SF10]	2.89	1.34
Photographs [SF10]	2.16	1.37
Blogs and Internet [SF11]	2.07	1.33
Games [SF12]	1.84	1.21
Social Networking [SF13]	1.84	1.25

B.2 Retail Preferences

Participants then rated 19 retail devices on aesthetics and expected usefulness. To make the task easier, participants only rated a subsample of five of them. All ratings were also rated on a 1-5 scale (not at all to extremely) or “I don’t know.” Code for each measure is in brackets (e.g.,[RL1]). They were given the following instructions:

Here are five commercially available in-home displays . Please rate each one in terms of (a) how much you like how it looks and (b) how useful it would be to you. How much do you like the way this display looks?

Here is an example of the description (the rest are in appendix 1):

Eco-Meter (Click image below for larger picture)

- Easy-to-read with large numbers and LED back-lighting.
- Designed to sit on your counter-top.
- Uses LED lights to alert you about important events (for ex., a period of high cost electricity).
- Displays the cost of the electricity youâve been using.
- >More about Eco-Meter

Participants then rated 19 retail devices on aesthetics and expected usefulness. To make the task easier, participants only rated a subsample of five of them. All ratings were also rated on a 1-5 scale (not at all to extremely) or “I don’t know.” Code for each measure is in brackets (e.g. [RL1]). They were given the following instructions:

Here are five commercially available in-home displays . Please rate each one in terms of (a) how much you like how it looks and (b) how useful it would be to you. Overall, how useful would this display be?

The devices were the EnergyHub HomeBase, OpenFrame, Control4 EC100, HEC, Silverstat 7,

Participants then created their own display by selecting what features they wanted from those listed above. Their constructions were almost identical to their ratings, so we omit them here. They also responded to the following questions about the display they created

- How much would you like to have the in-home display you created on the previous page? (not at all, slightly, moderately, strongly, extremely)
- How effective do you think that in-home display would help you to reduce your electricity use?” (very ineffective, ineffective, somewhat ineffective, neither effective nor ineffective, somewhat effective, effective, very effective)
- How often would you look at the in-home display you created? (Never, less than once a month, once a month, 2-3 times a month, once a week, 2-3 times a week, daily).

Table 5: Retail Preferences for Aesthetics and Usefulness

Item	Aesthetics		Usefulness		N
	Mean	SD	Mean	SD	
HomeBase	4.17	0.79	3.67	0.92	35
OpenFrame	3.96	1.06	3.64	1.15	27
EC100	3.90	0.84	3.76	0.96	101
HEC	3.77	1.12	3.66	1.26	31
Silverstat 7	3.77	1.09	3.30	0.95	13
PowerCost Monitor	3.76	0.93	3.82	1.07	37
Insight	3.69	1.00	4.15	0.77	29
Envi	3.61	0.99	3.85	0.77	28
Elite	3.54	0.93	3.82	0.80	24
Cent-A-Meter	3.49	0.77	3.50	0.77	37
Powertab	3.43	1.27	3.14	0.69	7
TED5000	3.37	1.05	3.70	0.89	113
GEO Minim	3.33	1.18	3.26	1.20	30
E2	3.30	0.67	3.18	1.17	10
Ratesaver	3.25	0.75	3.25	0.75	12
Intellifocus	2.97	0.98	3.11	0.93	86
Eco-Meter	2.96	1.00	3.12	0.97	26
Emu	2.92	1.00	2.83	0.72	12
Eco-Eye	2.50	1.27	2.80	1.14	10

- How much do you think you would save, in dollars, on your monthly electricity bill if you had the display you created on the previous page?

They liked the display they created, on average expected it to like it strongly (M=3.96, SD=0.84, N=138), they thought it would be effective (M=5.89, SD=1.09; N=139), expected to look at it on average 2-3 times a week (M=6.11, SD=1.1, N=138), expected significant monthly savings on their bill (M=\$36.2, SD=\$40.1, N=106; 95% CI [28, 44]).

The average willingness to pay for an IHD was \$154 (95%CI [2,306]) (SD=870, N=131), average WTP for IHD on a recurring monthly bill was \$18 (SD=39, N=121), average expected savings on the monthly bill was \$25 (SD=29, N=131), the average monthly electricity bill was \$106 (SD=81, N=106), and they reported using 893 kWh on average per month (SD=1134, N=20), but most either didn't answer the question or reported they didn't know.

C Sim. Measures

The electricity knowledge tests measured six main dependent variables to capture learning from the simulation. The first measure assessed participants' rankings of the ten appliances in terms of how much kWh the appliance used in a 10-minute period:

[jack] What was the text for the use?

Next, participants ranked the ten appliances by their contribution to the monthly bill:

Imagine the average summer monthly cost for a family of two adults and two children for each appliance listed below. Rank the ten appliances below by how much they would contribute to the monthly bill from 1 (the most) to 10 (the least).

The third question asked them to identify the correct unit of electrical energy:

What are the units of electrical energy called?

- Kilowatt (kW)
- Kilowatt-hours (kWh)
- British Thermal Units (BTU)
- Volts (V)
- Horsepower (HP)

Participants were then asked to identify the correct method of calculating energy:

Which of the following is equal to the amount of energy consumed by an electrical appliance?

- Power rating multiplied by the cost of electricity.
- Power rating added to the cost of electricity.
- Power rating multiplied by the time it's used.
- Power rating divided by the time it's used.
- Power rating added to the time it's used.

Participants then estimated the cost of a kWh (we used \$0.13 as an approximation):

"The cost of a kWh varies a little from state to state and from hour to hour. What is the average cost of a kWh for a residential U.S. customer?"

Finally, participants estimated the monthly kWh use for households.

About how many kWh does a typical U.S. household use in a summer month?

D Additional Measures and Results

As seen in Table 6, participants were not more able to infer how energy is calculated in any of the treatments. Most (136) participants did not change their answer. Some (34) changed from the wrong answer to the right answer, whereas others changed from the right answer to the wrong answer (22). There was no detectable variation by condition, but aggregating across all conditions there was improvement.

Condition	Correct	No Change	Incorrect	t-value
\$ and kWh by appliance	6	17	3	0.42
\$ and kWh aggregate	4	18	4	0
\$ by appliance	3	21	7	-0.63
kWh by appliance	5	17	1	0.64
Aggregate \$ only	3	23	2	0.19
Aggregate kWh only	3	17	3	0
Passive Learning	10	23	2	1.4
Total	34	136	22	4.7

Table 6: Identification of how energy is calculated by condition. T-values comparing proportion switched to correct versus incorrect for each condition are based on posterior simulations from a beta(0,0) prior.

D.0.1 Units of Energy

As seen in Table 7, in all conditions participants were more likely to change their answer from an incorrect unit of energy to the correct one (kWh) than from correct to incorrect. However, none of the differences were statistically significant. Aggregating across all conditions, participants were more likely to change their answer in the correct direction. However, there is no control group of participants who did not receive any feedback (i.e., merely completed the pre and post questionnaires), to compare them against. Thus, this positive change could be an effect of learning from the questions, that frequently asked about kWh, rather than the feedback.

Condition	Correct	No Change	Incorrect	t-value
\$ and kWh by appliance	5	19	2	0.48
\$ and kWh aggregate	7	16	3	0.54
\$ by appliance	3	26	2	0.21
kWh by appliance	10	9	4	0.67
Aggregate \$ only	3	24	1	0.43
Aggregate kWh only	11	10	2	1.1
Passive Learning	9	25	1	1.6
Total	48	129	15	13

Table 7: Identification of the correct unit of energy (kWh) by condition. T-values comparing proportion switched to correct versus incorrect for each condition are based on posterior simulations from a beta(0,0) prior.

D.0.2 Cost of a kWh

As seen in Table 8, participants in all groups were closer to the true kWh cost of \$0.13 in the post-treatment period, except for those who received \$ and kWh aggregate feedback. The only statistically significant improvement was for participants in the passive learning condition, the median of which got the true kWh cost exactly correct.

Condition	Median kWh Cost			
	Pre	Post	V	p-value
\$ and kWh by appliance	1.3	1.12	9.5	0.26
\$ and kWh aggregate	1	1.2	8	0.18
\$ by appliance	1	0.33	9	0.23
kWh by appliance	0.3	0.25	8	1
Aggregate \$ only	0.55	0.3	7	0.14
Aggregate kWh only	1.115	1.05	11	0.67
Passive Learning	0.75	0.13	10	0.025
Total	1	0.35	436	0.0031

Table 8: Median estimates of the cost of a kWh before and after treatment. T-values are from Wilcoxon paired rank deviations from the true values.

D.0.3 Monthly kWh

Table 9 shows estimates of monthly kWh use.

Condition	Median kWh Use			
	Pre	Post	V	p-value
\$ and kWh by appliance	450	500	126	0.02
\$ and kWh aggregate	600	300	74	0.46
\$ by appliance	200	322	38	0.62
kWh by appliance	600	600	50	0.57
Aggregate \$ only	920	800	16	0.26
Aggregate kWh only	300	320	74	0.76
Passive Learning	290	958	2	6.8e-07
Total	2000	300	2506	0.0048

Table 9: Median estimates monthly kWh use before and after treatment. T-values are from Wilcoxon paired rank deviations from the true values

E Detailed kWh rank analyses

Table 10 summarizes the effect of each treatment condition on each appliance. AC benefited for kWh by appliance for understanding use, $V = 15$, $p = 0.054$, or total only with kWh use, $V = 28$, $p = 0.021$. Dryer slightly benefitted with total kWh and \$, but not otherwise $V = 100$, $p = 0.096$. Microwave differed in kWh by appliance $V = 128$, $p = 0.06$, total \$ only $V = 154$, $p = 0.064$, total kWh only $V = 80$, $p = 0.015$, and passive learning $V = 424$, $p = 7.5e-06$. Heater benefited from kWh by appliance $V = 83$, $p = 0.055$, and passive learning $V = 222$, $p = 0.01$. TV benefitted from both kWh and \$ by appliance $V = 106$, $p = 0.043$, and passive learning $V = 188$, $p = 0.038$. The refrigerator benefitted from total kWh only $V = 147$, $p = 0.034$, and passive learning $V = 240$, $p = 0.0019$. Indoor lighting was helped only by passive learning $V = 38$, $p = 0.067$. The washing machine did not benefit in any condition. There was no learning across any condition for the freezer.

Appliance	\$, kWh, AS	\$, kWh	\$, AS	kWh, AS	\$	kWh	Passive	Total
AC				+		+		2
Dryer		+						1
Microwave				+	+	+	+	4
Oven					-			(-1)
Water Heater				+		-	+	2 (-1)
TV	+						+	2
Fridge						+	+	2
Indoor Lights							+	1
Washer								0
Freezer								0
Total	1	1	0	3	1 (-1)	3 (-1)	5	

Table 10: Helped (+) or harmed (-) rankings of energy use for each appliance. AS is appliance specific feedback.

Unexpectedly, the accuracy of ranking of two appliances were actually harmed by the treatment. Oven was *harmed* by total \$ only, but no other conditions, $V = 64$, $p = 0.042$. Heater was *harmed* by total kWh only $V = 20$, $p = 0.012$.

Table 11: Air Conditioner

Condition	Pre Rank Dev	Post Rank Dev	Statistic	p-value
\$ only	1.2	0.86	41	0.5
kWh only	1.2	0.39	28	0.021
\$ and kWh	0.75	1.1	26	0.59
\$ app.	0.8	0.5	48	0.19
kWh app.	1	0.27	15	0.054
\$ and kWh app.	0.97	0.38	37	0.34
Passive	0.81	0.39	40	0.22

Table 12: Dryer

Condition	Pre Rank Dev	Post Rank Dev	Statistic	p-value
\$ only	2.1	2.1	126	0.43
kWh only	2.1	1.3	66	0.39
\$ and kWh	2.3	1.3	100	0.096
\$ app.	2.3	2.1	74	0.79
kWh app.	1.2	1.2	55	0.79
\$ and kWh app.	2.4	2.8	80	0.9
Passive	1.7	1.1	178	0.43

Table 13: Oven

Condition	Pre Rank Dev	Post Rank Dev	Statistic	p-value
\$ only	1.9	2.7	64	0.042
kWh only	2.9	2.8	133	0.12
\$ and kWh	2.2	1.8	118	0.65
\$ app.	2.7	2.7	111	0.84
kWh app.	2.6	2.3	55	0.53
\$ and kWh app.	2.3	2.2	66	0.65
Passive	1.9	1.8	249	0.99

Table 14: Microwave

Condition	Pre Rank Dev	Post Rank Dev	Statistic	p-value
\$ only	3.5	3.2	154	0.064
kWh only	3.6	3	80	0.015
\$ and kWh	3.8	3.2	98	0.31
\$ app.	3.5	3.4	99	0.57
kWh app.	3.5	2.7	128	0.06
\$ and kWh app.	3.9	4.2	37	0.19
Passive	3.8	1.5	424	7.5e-06

Table 15: Heater

Condition	Pre Rank Dev	Post Rank Dev	Statistic	p-value
\$ only	2.1	2.1	60	0.67
kWh only	1.8	2.5	20	0.012
\$ and kWh	1.9	1.8	70	0.27
\$ app.	2.1	2.1	97	0.77
kWh app.	2.2	1.6	83	0.055
\$ and kWh app.	1.9	2.2	84	0.95
Passive	2.1	1.3	222	0.01

Table 16: Washer

Condition	Pre Rank Dev	Post Rank Dev	Statistic	p-value
\$ only	2	1.8	84	0.96
kWh only	1.7	1.6	40	0.72
\$ and kWh	1.9	1.8	92	0.44
\$ app.	2.1	1.8	126	0.43
kWh app.	1.9	1.6	81	0.51
\$ and kWh app.	1.9	1.5	76	0.36
Passive	1.8	1.7	182	0.36

Table 17: Freezer

Condition	Pre Rank Dev	Post Rank Dev	Statistic	p-value
\$ only	1.8	2.3	83	0.41
kWh only	2.2	1.8	111	0.27
\$ and kWh	1.9	1.5	102	0.8
\$ app.	2.1	2.2	73	0.59
kWh app.	2.2	1.9	86	0.37
\$ and kWh app.	2.2	2.3	70	0.51
Passive	1.8	1.5	110	0.28

Table 18: TV

Condition	Pre Rank Dev	Post Rank Dev	Statistic	p-value
\$ only	1.6	1.5	100	0.5
kWh only	1.8	1.5	58	0.77
\$ and kWh	1.7	1.5	56	0.84
\$ app.	1.9	1.8	95	1
kWh app.	1.2	1.1	29	0.75
\$ and kWh app.	2.1	1.4	106	0.043
Passive	1.5	0.64	188	0.038

Table 19: Fridge

Condition	Pre Rank Dev	Post Rank Dev	Statistic	p-value
\$ only	3.8	3.7	140	0.19
kWh only	4.2	3.4	147	0.034
\$ and kWh	4.4	3.8	150	0.23
\$ app.	3.3	3.5	140	0.95
kWh app.	3.5	3.5	58	0.64
\$ and kWh app.	4.2	4.1	95	0.69
Passive	3.8	2.2	240	0.0019

Table 20: Indoor Lights

Condition	Pre Rank Dev	Post Rank Dev	Statistic	p-value
\$ only	0.72	0.69	27	0.63
kWh only	1.3	0.48	18	0.55
\$ and kWh	1.4	0.92	38	0.94
\$ app.	1.4	1.2	54	0.92
kWh app.	0.56	0.41	10	1
\$ and kWh app.	1	1.3	16	0.44
Passive	0.78	0.24	38	0.067

F Detailed cost rank analyses

As can be seen in Table 21, for the oven, participants were harmed by total kWh and \$ $V = 83$, $p = 0.093$, harmed by kWh by appliance $V = 10$, $p = 0.041$, but helped by passive learning $V = 300$, $p = 0.007$. Heater benefitted in \$ and kWh by appliance $V = 118$, $p = 0.05$, and passive learning $V = 286$, $p = 0.02$. For the washing machine benefitted from kWh by appliance $V = 120$, $p = 0.007$, total kWh only $V = 107$, $p = 0.044$, and passive learning $V = 372$, $p = 0.043$. The freezer only benefitted from passive learning $V = 405$, $p = 0.0078$. TV benefitted from kWh and \$ by appliance $V = 118$, $p = 0.047$, and \$ only by appliance $V = 126$, $p = 0.077$. Indoor lights were harmed by \$ and kWh by appliance $V = 6$, $p = 0.057$, but helped by passive learning $V = 69$, $p = 0.019$. There were no significant benefits for cost for AC, dryer, microwave, refrigerator.

Appliance	\$, kWh, AS	\$, kWh	\$, AS	kWh, AS	\$	kWh	Passive	Total
AC								
Dryer								
Microwave								
Oven		-		-			+	
Water Heater	+						+	
TV	+				+			
Fridge								
Indoor Lights	-						+	
Washer				+		+	+	
Freezer							+	
Total	2 (-1)	(-1)		0	1 (-1)	1	5	?

Table 21: Helped (+) or harmed (-) rankings of energy cost for each appliance. AS is appliance specific feedback.

Thus, in terms of learning the cost each appliance contributes to the monthly bill, only passive learning was effective, with some indication that kWh and \$ appliance specific information also help.

References

- [1] W. Kempton and L.L. Layne. The consumer’s energy analysis environment. *Energy Policy*, 22(10):857–866, 1994. 2
- [2] S. Darby. Smart metering: what potential for householder engagement? *Building Research & Information*, 38(5):442–457, 2010. 2, 5, 6, 19, 20
- [3] S. Roberts, H. Humphries, and V. Hyldon. Consumer preferences for improving energy consumption feedback. *Report to Ofgem, Centre for Sustainable Energy*, 2004. 2, 3, 19, 20
- [4] L. Rozenblit and F. Keil. The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science*, 26(5):521–562, 2002. 2
- [5] F.C. Keil. Folkscience: Coarse interpretations of a complex reality. *Trends in Cognitive Sciences*, 7(8):368–373, 2003. 2

- [6] G. Wood and M. Newborough. Dynamic energy-consumption indicators for domestic appliances: environment, behaviour and design. *Energy and Buildings*, 35(8):821–841, 2003. 2, 6
- [7] T. Krishnamurti, D. Schwartz, A. Davis, B. Fischhoff, W.B. de Bruin, L. Lave, and J. Wang. Preparing for smart grid technologies: A behavioral decision research approach to understanding consumer expectations about smart meters. *Energy Policy*, 2011. 2
- [8] C. Seligman, J.M. Darley, and L.J. Becker. Behavioral approaches to residential energy conservation. *Energy and Buildings*, 1(3):325–337, 1978. 2
- [9] C. Seligman, LS Becker, and J.M. Darley. Encouraging residential energy conservation through feedback. *Advances in environmental psychology*, 3:93–113, 1981. 2
- [10] J.K. Dobson and J.D.A. Griffin. Conservation effect of immediate electricity cost. feedback on residential consumption behavior. *Proceedings of the 7th ACEEE Summer Study on Energy Efficiency in Buildings*, 1992. 2
- [11] L. McClelland and S.W. Cook. Energy conservation effects of continuous in-home feedback in all-electric homes. *Journal of Environmental Systems*, 9(2):169–173, 1979. 2
- [12] T.J. Yun. Investigating the impact of a minimalist in-home energy consumption display. In *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems*, pages 4417–4422. ACM, 2009. 2
- [13] G. Wood and M. Newborough. Energy-use information transfer for intelligent homes: Enabling energy conservation with central and local displays. *Energy and Buildings*, 39(4):495–503, 2007. 3
- [14] S. Karjalainen. Consumer preferences for feedback on household electricity consumption. *Energy and Buildings*, 43(2):458–467, 2011. 3, 5, 19, 20
- [15] W. Anderson and V. White. Exploring consumer preferences for home energy display functionality. *Centre for Sustainable Energy, Tech. Rep*, 2009. 3, 5, 19, 20
- [16] T.D. Wilson and D.T. Gilbert. Affective forecasting. *Advances in experimental social psychology*, 35:345–411, 2003. 3
- [17] G. Loewenstein and J.S. Lerner. The role of affect in decision making. *Handbook of affective science*, 619:642, 2003. 3
- [18] L. Steg and C. Vlek. Encouraging pro-environmental behaviour: An integrative review and research agenda. *Journal of Environmental Psychology*, 29(3):309–317, 2009. 3, 7
- [19] C. Egan. Graphical displays and comparative energy information: what do people understand and prefer. *Summer Study of the European Council for an Energy Efficient Economy*, (2-12), 1999. 3

- [20] Hoivik Wilhite, H. A., JG. Olsen, and R. Ling. Advances in the use of consumption feedback information in energy billing: the experiences of a norwegian energy utility. In *Proceedings, European Council for an Energy-Efficient Economy*, 1999. 3
- [21] C. Fischer. Feedback on household electricity consumption: a tool for saving energy? *Energy Efficiency*, 1(1):79–104, 2008. 3, 19
- [22] S. Roberts and W. Baker. Towards effective energy information. improving consumer feedback on energy consumption. *Centre for Sustainable Energy*, 2003. 3, 20
- [23] B. Neenan, J. Robinson, and RN Boisvert. Residential electricity use feedback: A research synthesis and economic framework. *Retrieved October*, 26:2011, 2009. 3
- [24] A. Gelman and J. Hill. *Data analysis using regression and multi-level/hierarchical models*, volume 1. Cambridge University Press New York, 2007. 13
- [25] A. Gelman, Y.S. Su, M. Yajima, J. Hill, M.G. Pittau, J. Kerman, and T. Zheng. arm: Data analysis using regression and multilevel/hierarchical models. *R package version*, pages 1–3, 2010. 13
- [26] T. Ueno, R. Inada, O. Saeki, and K. Tsuji. Effectiveness of displaying energy consumption data in residential houses. analysis on how the residents respond. *Proceedings, European Council for an Energy-efficient Economy, paper*, 6, 2005. 19
- [27] D. Allen and K. Janda. The effects of household characteristics and energy use consciousness on the effectiveness of real-time energy use feedback: a pilot study. In *Proceedings of the ACEEE Summer Study on Energy Efficiency in Buildings*, pages 7–1, 2006. 19
- [28] H. Wilhite and R. Ling. Measured energy savings from a more informative energy bill. *Energy and buildings*, 22(2):145–155, 1995. 19
- [29] A.G. Paetz, B. Becker, W. Fichtner, and H. Schmeck. Shifting electricity demand with smart home technologies—an experimental study on user acceptance. In *30th USAEE/IAEE North American Conference Online Proceedings*, 2011. 19, 20