# A Mouse Over a Hotspot Survey: An Exploration of Perceptions of Electricity Consumption and Patterns of Indecision

Lucas Pereira
Madeira-ITI / LARSYS and prsma.com
Funchal, Portugal
lucas.pereira@m-iti.org

Yoram Chisik
Madeira-ITI / LARSYS
Funchal, Portugal
ychisik@m-iti.org

***Abstract*—In order to promote a more effective understanding of electricity consumption in households, and leverage the potential of current and future energy conservation initiatives, we have developed an online survey in which visual icons of various electrical devices coupled with a Likert scale are used to investigate people's perception of the amount of electricity consumed by a variety of individual domestic appliances. In this paper we present initial results from a pilot study conducted with 293 participants from over 50 countries and outline our plans for future work. Our initial results suggest that despite a considerable number of respondents was able to give the correct answers, many others seem to lack an understanding of appliance electricity consumption, and thus further work should be carried in this direction.**

***Index Terms*—Electricity consumption, domestic appliances, perceptions, mental models, sustainability, eco-feedback**

## I. Introduction

The consumption of electricity has drawn a lot of attention in recent years due to the environmental impact associated with its generation. Electricity presents a particular problem to consumers and researchers alike, as unlike other utilities such as water and gas, electricity is an invisible resource with no visible form, flow or weight, thus making it hard or even impossible for people to gauge the quantity of electricity consumed by individual appliances.

Prior studies about the perceptions of electricity consumption [1], [2] have shown that people do not have a working mental model of the rate in which electricity is consumed by various electrical devices and thus rely on analogous notions drawn from other domains such as size, duration and frequency of use. Consequently, consumers frequently make quantification errors leading to under and over estimation of consumption and ineffective energy conservation efforts.

To address this question we set up an online survey aimed at exploring the way in which people gauge the electricity consumption of various domestic appliances. Our objective in this survey is twofold: First we want to build a large multicultural database of electricity consumption estimations as a means of exploring the perceptions and miss-perceptions of electricity consumption. Second, knowing people have no working mental model of electricity consumption and thus would likely be very hesitant in their answers we want to explore the hesitation patterns of the survey respondents by examining their mouse motions as they deliberate their answer and hover over the various screen elements as opposed to just examining the final ratings.

The data generated by such a study should provide useful insights into the perceptions of electricity of a large and diverse group of consumers which in turn should enable us to ponder the nature of electricity consumption and its implications for the design of eco-feedback, technologies, energy conservation programs, and new policies. In this paper we present some initial results from a pilot study we conducted and our initial plans for this study. The remainder of this paper is organized as follows: in next section we present the design rationale of this pilot study. We then present and discuss the obtained results before we conclude and outline direction for future work on this topic.

## II. Study Design

As an initial pilot study we have developed a web based system in which a series of screens each containing a single icon representing a domestic electrical appliance coupled with a label are displayed to a participant who is asked to rate the electric consumption of the device using a relative 1-10 Likert scale representing low or high consumption as opposed to using specific measurements such as kWh, which as literature suggest are not common metrics used by people [3].

The icons representing 41 appliances of different size and function have been drawn using the same visual language, as shown in figure 1. The question and labels have been translated into 18 of the most commonly spoken languages to enable us to get the broadest global perspective. The text on the screens has been kept to a minimum in order to maintain the same meaning in various languages and reduce the potential for misinterpretation. The various translations were validated in an initial pilot study with over 40 participants comprised of at least 2 native speakers of each language. *Ad hoc* follow up interviews were conducted with these participants.

The screens are displayed in a random order to minimize the potential of bias due to brand recognition or sequence

978-3-901882-99-9 


Fig. 1. Listing of the appliance icons used in this pilot study

patterns, e.g., several large-scale appliances displayed one after the other. Once the participant clicks one of the Likert scale score buttons, the response is recorded and the system moves on to the next appliance screen. Figure 2 shows an example of a rating screen translated in two different languages. Upon ratting all the appliances, the participants are asked a few demographic questions about themselves and their electricity bill.

In order to explore patterns of indecision, each time a screen with an appliance icon is displayed, the system tracks the participant mouse movements (*mouse enter*, *mouse leave* and *click*) on the main items of the survey (appliance icon, appliance label and the 10 score buttons of the Likert scale). Table I shows the mouse events for one of the respondents when rating the laptop.

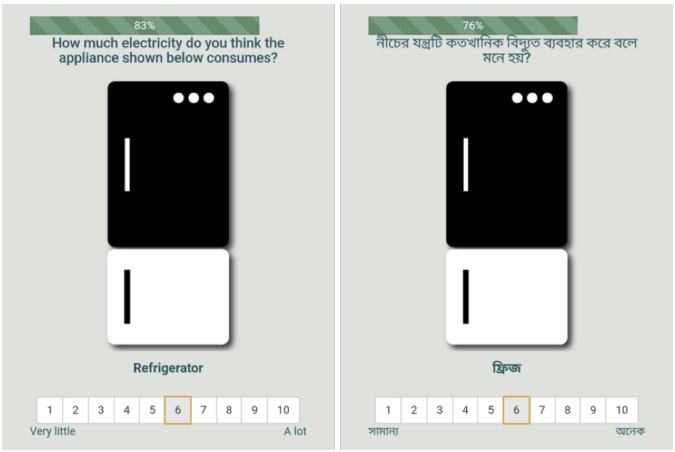


Fig. 2. An example of a rating screen (translated in two different languages) where the respondent is asked to rate the consumption of a refrigerator.

Hover events are derived from the data computationally by calculating the time difference between *mouse enter* and *mouse leave* events with a minimum of 100ms required to count a hover event, as suggested in [4]. Figure 3 - left, shows a graphical representation of the mouse events in table I. In this case, the participant hovered over score button 9 and then moved (but did not hover) over score buttons 8 and 7. Finally, the respondent hovers and click on score button 6.

Overall, the mouse interactions can vary greatly in their duration and complexity of the movements, as illustrated in figure 3. For example, for the satellite dish, three scores were hovered before the final decision. Furthermore, it is also possible to observe that the three scores were visited at least twice, and that there are two loops in score 7.

## III. Results and Discussion

The system was tested in a first study to gather an initial dataset for a test trial exploration of the data. This test lasted 12-week period, and had a total of 293 participants. The participants were drawn via posts on Facebook for which no remuneration was offered and via *jobboy.com*, a crowd-

TABLE I
List of mouse movements recorded when rating a laptop

| Rating | Time | Event Type | Elapsed Time |
|---|---|---|---|
| 9 | 14:49:29.342 | mouse leave | - |
| 8 | 14:49:29.342 | mouse enter | 33 ms |
| 8 | 14:49:29.375 | mouse leave | |
| 7 | 14:49:29.376 | mouse enter | 15 ms |
| 7 | 14:49:29.391 | mouse leave | |
| 6 | 14:49:29.392 | mouse enter | 817 ms |
| 6 | 14:49:30.209 | click | |

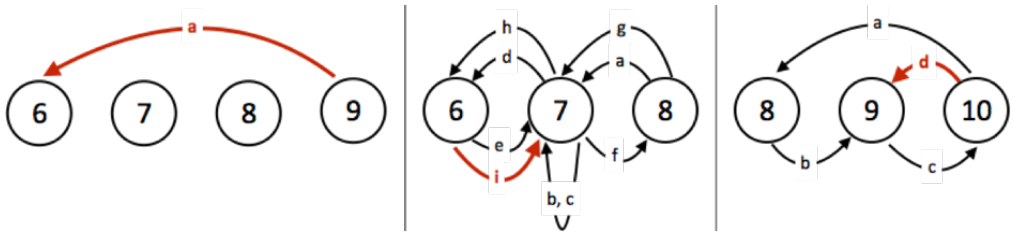


Fig. 3. Graphical representation of mouse movements when rating appliances: laptop(left), the satellite dish (center), and hair dryer (right).

sourcing web site for which a remuneration of 10 Euro cents was offered.

The respondents were predominantly English speaking (83.7%), males (74%) between the ages of 18 and 40 (89%) from 50 different countries with a majority from the Indian Sub-continent: Bangladesh, India, Nepal and Pakistan (45.2%). In Europe, Portugal was the most represented country (17.6%).

We conducted an analysis of the collected data in order to get a feel for it, and as a means of formulating a solid base of hypothesis for a future large-scale study. The results are presented below.

### A. Appliance Consumption Ratings

The consumption rates reported by each participant were compared to a baseline scale that was created based on the average consumption rates for domestic appliances provided in the Non-Intrusive Load Monitoring (NILM) Wiki web site[1]. To do this, we sorted the baseline data according to the average consumption rate of each device and then divided them into 10 non-equidistant groups in order to correspond with the 10 point Likert scale used in the survey.

Figure 4 presents a summary of the obtained ratings. The appliances appear sorted from high to low consumption according to our baseline scale. There we show the number of times each rank was selected, highlighting the mode. We also show the quartile information $1^{st}$, $2^{nd}$, and $3^{rd}$ quartiles, as well as the Inter-Quartile Range - IQR). Finally, we show the expected rank (ER) according to our scale. The median ($2^{nd}$ quartile) and the ER are color-coded according to a gradient scale ranging from green (1 – very little electricity consumption) to red (10 – a lot of electricity consumption).

In order to understand how close the respondents perceptions are from the baseline scale, we ran the *Spearman* rank-order correlation between the median ratings and the Expected Result (ER). There was a strong, positive correlation between the median scores and the expected results, which was statistical significant ($\rho(39) = .817$, $P = 3.46^{-11}$).

This high correlation suggests that the central tendency is to have responses closer to our scale. However, it also highlights the fact that many respondents rated the appliances in a totally different way, which can also be observed from the high values of the IQR (3 or 4).

Another interesting observation is that all the options in the scale were selected at least once for each appliance (i.e., the range of the responses = 9). While it is not possible to infer the reason for this, we believe that one possibility is that the respondents had different interpretations of the questions. For example, some people may have interpreted the same way we did, i.e., instantaneous consumption, but others may have had other interpretations, like for example, how much power is actually spent for each time the appliance is used.

### B. Mouse movements

Our exploration of mouse movements focused on analyzing apparent hesitation on the part of the participants in choosing the consumption score for each appliance. Since we hypothesize participants will lack a clear mental model of the how

| Appliance Name | Ratings | | | | | | | | | | Quartiles | | | | ER |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 25% | 50% | 75% | IQR | |
| Air conditioner | 7 | 5 | 6 | 16 | 19 | 21 | 43 | 49 | 46 | **81** | 6 | 8 | 10 | 4 | 10 |
| Clothes dryer | 7 | 5 | 12 | 19 | 37 | 37 | 44 | 42 | **48** | 42 | 5 | 7 | 9 | 4 | 10 |
| Water heater | 11 | 11 | 17 | 18 | 24 | 23 | 47 | 52 | 31 | **59** | 5 | 7 | 9 | 4 | 10 |
| Dishwasher | 17 | 9 | 18 | 27 | **45** | **45** | 44 | 33 | 28 | 27 | 5 | 6 | 8 | 3 | 9 |
| Heater | 9 | 13 | 7 | 14 | 16 | 18 | 36 | 47 | 55 | **78** | 6 | 8 | 10 | 4 | 9 |
| Oven | 7 | 9 | 14 | 18 | 26 | 30 | 32 | 55 | 42 | **60** | 5 | 8 | 9 | 4 | 9 |
| Stove | 11 | 7 | 23 | 23 | 25 | 33 | 37 | **51** | 39 | 44 | 5 | 7 | 9 | 4 | 9 |
| Hair dryer | 16 | 33 | **50** | 39 | 27 | 35 | 34 | 27 | 16 | 16 | 3 | 5 | 7 | 4 | 8 |
| Iron | 11 | 13 | 28 | 23 | 31 | **39** | **39** | **39** | 36 | 34 | 4 | 7 | 8 | 4 | 8 |
| Microwave | 6 | 12 | 20 | 29 | 30 | 41 | **52** | 40 | 26 | 37 | 5 | 7 | 8 | 3 | 8 |
| Toaster | 9 | 27 | 40 | 37 | **46** | 40 | 24 | **38** | 10 | 22 | 3 | 5 | 7 | 4 | 8 |
| Coffee maker | 13 | 28 | 41 | 45 | **52** | 38 | 29 | 20 | 9 | 18 | 3 | 5 | 7 | 4 | 7 |
| Popcorn machine | 17 | 30 | **48** | 40 | 47 | 30 | 29 | 18 | 20 | 14 | 3 | 5 | 7 | 4 | 7 |
| Refrigerator - Large | 7 | 8 | 10 | 16 | 25 | 31 | 36 | 53 | 44 | **63** | 6 | 8 | 9 | 3 | 7 |
| Vaccum cleaner | 12 | 21 | 27 | 42 | 39 | 36 | **47** | 37 | 18 | 14 | 4 | 6 | 7 | 3 | 7 |
| Washing machine | 8 | 5 | 14 | 19 | 18 | 40 | 48 | **53** | 37 | 51 | 6 | 7 | 9 | 3 | 7 |
| Blender | 17 | 26 | 49 | 51 | 35 | **52** | 25 | 20 | 9 | 9 | 3 | 5 | 6 | 3 | 6 |
| Electric blanket | 15 | 23 | 33 | 31 | **49** | 33 | 35 | 33 | 21 | 20 | 4 | 5 | 8 | 4 | 6 |
| Freezer | 6 | 3 | 14 | 22 | 30 | 24 | 45 | **64** | 41 | 44 | 5 | 8 | 9 | 4 | 6 |
| Mixer - Table | 16 | 24 | 47 | **57** | 42 | 39 | 32 | 12 | 9 | 15 | 3 | 5 | 6 | 3 | 6 |
| Refrigerator - Medium | 7 | 7 | 8 | 16 | 25 | 42 | 43 | 45 | 45 | **55** | 6 | 7 | 9 | 3 | 6 |
| Desk computer | 12 | 19 | 38 | **51** | 44 | 29 | 35 | 23 | 21 | 21 | 4 | 5 | 7 | 3 | 5 |
| Mixer - Hand | 17 | 30 | 53 | **54** | 43 | 39 | 22 | 16 | 4 | 15 | 3 | 4 | 6 | 3 | 5 |
| Playstation | 38 | **46** | 44 | 45 | 34 | 24 | 27 | 9 | 12 | 14 | 2 | 4 | 6 | 4 | 5 |
| Refrigerator - Small | 6 | 6 | 11 | 20 | 23 | 39 | 46 | 47 | 39 | **56** | 6 | 7 | 9 | 3 | 5 |
| TV - CRT | 12 | 22 | 44 | 39 | **52** | 41 | 30 | 20 | 16 | 17 | 3 | 5 | 7 | 4 | 5 |
| Aquarium | 49 | **56** | 46 | 45 | 24 | 20 | 19 | 16 | 8 | 10 | 2 | 3 | 5 | 3 | 4 |
| Curling iron | 27 | 36 | 36 | **42** | 33 | 36 | 33 | 19 | 17 | 14 | 3 | 5 | 7 | 4 | 4 |
| Lightbulb - ICL | 39 | **54** | **55** | 38 | 23 | 29 | 12 | 12 | 18 | 13 | 2 | 3 | 6 | 4 | 4 |
| TV - Flatscreen | 13 | 25 | 37 | **43** | 42 | 41 | 30 | 23 | 17 | 22 | 3 | 5 | 7 | 4 | 4 |
| Ceilling fan | 19 | 30 | 45 | **53** | 38 | 29 | 27 | 23 | 16 | 13 | 3 | 4 | 7 | 4 | 3 |
| DVD player | 22 | **67** | 65 | 38 | 43 | 14 | 13 | 10 | 10 | 11 | 2 | 3 | 5 | 3 | 3 |
| Laptop | 13 | 43 | **49** | 42 | 42 | 29 | 23 | 20 | 15 | 17 | 3 | 4 | 7 | 4 | 3 |
| Stereo system | 23 | 38 | **61** | 59 | 38 | 21 | 20 | 12 | 6 | 15 | 3 | 4 | 6 | 3 | 3 |
| Cable TV converter | 41 | **57** | 52 | 35 | 42 | 17 | 17 | 15 | 7 | 10 | 2 | 3 | 5 | 3 | 2 |
| Desk fan | 23 | 40 | **50** | 49 | 45 | 25 | 22 | 12 | 12 | 15 | 3 | 4 | 6 | 3 | 2 |
| Lightbulb - CFL | **71** | 47 | 62 | 34 | 22 | 15 | 11 | 9 | 10 | 12 | 2 | 3 | 5 | 3 | 2 |
| Satelite dish | 26 | 23 | **44** | 43 | 37 | 22 | 26 | **30** | 20 | 22 | 3 | 5 | 7 | 4 | 2 |
| Cell phone | **87** | 68 | 34 | 26 | 22 | 16 | 6 | 8 | 11 | 15 | 1 | 2 | 5 | 4 | 1 |
| IPod - MP3 Player | **109** | 60 | 44 | 23 | 16 | 14 | 5 | 4 | 12 | 6 | 1 | 2 | 4 | 3 | 1 |
| Router | **67** | 55 | 60 | 30 | 26 | 7 | 16 | 9 | 11 | 12 | 2 | 3 | 5 | 3 | 1 |

Fig. 4. A summary of the obtained ratings for each appliance.

[1]NILM Wiki, http://wiki.nilm.eu/appliance.html

Average clicks by score

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Avg. | 19 | 24 | 31 | 29 | 29 | 26 | 26 | 23 | 18 | 24 |
| St. Dev. | 19 | 16 | 14 | 10 | 8 | 8 | 10 | 13 | 11 | 15 |

Average hovers by score

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Avg. | 27 | 49 | 66 | 72 | 72 | 68 | 64 | 54 | 41 | 31 |
| St. Dev. | 24 | 29 | 27 | 19 | 13 | 13 | 21 | 25 | 23 | 19 |

Average clicks / hovers ratio by score

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Avg. | 69% | 48% | 45% | 40% | 40% | 38% | 40% | 42% | 45% | 75% |
| St. Dev. | 13% | 11% | 9% | 6% | 7% | 8% | 7% | 8% | 9% | 10% |

Fig. 5. Clicks, Hovers and Clicks/Hovers ratio, averaged by the 41 appliances.

much electricity is consumed by each device we assume they will hesitate before choosing a score.

For this analysis we looked at the ratio between hovers and clicks [4]. A hover is an event in which the cursor enters an area of the screen occupied by one of the relevant system elements (appliance icon, appliance label and the 10 rating buttons) and stays for a minimum of 100 ms [4].

Figure 5 shows the average clicks, hovers and ratio clicks/hovers for each rating score. As it can be observed, there is a clear tendency for more un-clicked hovers in the middle scores. In principle this can be understood in a sense that when in doubt about the answers the respondents would hover the mouse in the middle-scores. However, while this can suggest doubt and lack of confidence in the answers, part of it can also be explained by the central tendency bias that is known to affect Likert-scales [5].

## IV. Conclusion and Future Work

The initial analysis of the data seems to confirm our hypothesis that people lack a mental model for electricity consumption and thus the consumption scores they assign to various appliances is fraught with hesitation and misconceptions.

The data and post survey interviews conducted with some of the respondents raised questions related to interpretation of the duration of use and regional variation in familiarity with various domestic appliances. For example, toasters and hair dryers are high consumers but are only used for short periods of time, refrigerators are generally big but are controlled by a thermostat and thus do not consume electricity all the time. The current survey does not clearly address these conceptual scenarios.

The survey is aimed at a wide global audience but displays the same set of iconified appliances to all respondents ignoring regional variations in exposure to different kinds of appliances. For example, Danish respondents remarked that popcorn-machines are American contraptions and although they can imagine their purpose they are not part of the Danish sphere of domestic appliance and thus they have no clue about them. Similar remarks were given by Korean respondents with regards to clothes dryers, while others remarked their region is not served by cable television and thus they have no knowledge of cable converters.

We deliberately displayed the same set of appliances to all respondents partly for the sake of consistency but primarily because we wanted to see how participants would respond to an unfamiliar device. We hypothesized that unfamiliar devices would lead to high levels of hesitation and over/under estimation of consumption which seems to be confirmed by the data, but raises questions of generalizability and thus we need to further examine the question of regional variation and how it affects the perceptions of the participants.

Our next steps include a large-scale study in which we plan to get a larger and more diverse sample and address the issues uncovered in the pilot. We plan to ask a more diverse set of questions to look at various concepts of consumption related to devices e.g., people ranked a hair dryer as a low consumer because they think of it as a small device used for a short period of time. Hence, the questions should address both the concept of power required to use a device (measured in Watts) and overall energy consumption (measured in kWh) to get a better sense of the concepts people have for the energy needed to power a device and the overall impact it has on their daily consumption of electricity.

Also, rather than use a single relative Likert scale as we did in the pilot study (for the sake of simplicity and consistency) we intend to use a mix of relative and absolute scales and include a measure of confidence to explore how confident participants are with their own responses.

In addition, if we are able to get sufficient level of support for the various languages involved we would like to add open ended questions and conduct more extensive follow up interviews with some of the participants.

## Acknowledgements

This research was supported by the FCT doctoral grant SFRH/DB/77856/2011 and project UID/EEA/50009/2013.

## References

[1] W. Kempton and L. Montgomery, "Folk quantification of energy," *Energy*, vol. 7, no. 10, pp. 817–827, 1982.

[2] Y. Chisik, "An image of electricity: Towards an understanding of how people perceive electricity," in *Proceedings of the 13th IFIP TC 13 International Conference on Human-computer Interaction - Volume Part IV*, ser. INTERACT'11. Berlin, Heidelberg: Springer-Verlag, 2011, pp. 100–117. [Online]. Available: http://dl.acm.org/citation.cfm?id=2042283.2042295

[3] G. Wallenborn, M. Orsini, and J. Vanhaverbeke, "Household appropriation of electricity monitors," *International Journal of Consumer Studies*, vol. 35, no. 2, pp. 146–152, Mar. 2011. [Online]. Available: http://onlinelibrary.wiley.com/doi/10.1111/j.1470-6431.2010.00985.x/abstract

[4] J. Huang, R. W. White, and S. Dumais, "No Clicks, No Problem: Using Cursor Movements to Understand and Improve Search," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '11. New York, NY, USA: ACM, 2011, pp. 1225–1234. [Online]. Available: http://doi.acm.org/10.1145/1978942.1979125

[5] R. Garland, "The mid-point on a rating scale: Is it desirable," *Marketing Bulletin*, pp. 66–70, 1991.