Using Public Participation to Improve MELs Energy Data Collection

Iris Cheung, Margarita Kloss, Rich Brown and Alan Meier

Environmental Energy Technologies Division

October 2013
DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal 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 its 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, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or The Regents of the University of California.

ACKNOWLEDGEMENTS

This work was supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, Building Technologies Office, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>1</td>
</tr>
<tr>
<td>1 Introduction and Background</td>
<td>1</td>
</tr>
<tr>
<td>2 Project Activities and Report Outline</td>
<td>2</td>
</tr>
<tr>
<td>3 Controlled Experiment</td>
<td>3</td>
</tr>
<tr>
<td>3.1 Device Selection</td>
<td>3</td>
</tr>
<tr>
<td>3.2 Protocol Refinement</td>
<td>3</td>
</tr>
<tr>
<td>4 Open House Experiment</td>
<td>4</td>
</tr>
<tr>
<td>4.1 Setup</td>
<td>5</td>
</tr>
<tr>
<td>4.2 Recruiting</td>
<td>6</td>
</tr>
<tr>
<td>4.3 Experiment</td>
<td>7</td>
</tr>
<tr>
<td>4.4 Web Form</td>
<td>7</td>
</tr>
<tr>
<td>5 Findings</td>
<td>8</td>
</tr>
<tr>
<td>5.1 Meter Identification</td>
<td>8</td>
</tr>
<tr>
<td>5.2 Product Identification</td>
<td>9</td>
</tr>
<tr>
<td>5.3 Mode Identification</td>
<td>9</td>
</tr>
<tr>
<td>5.4 Measurements</td>
<td>10</td>
</tr>
<tr>
<td>5.5 Measurement Time for Each Test Device</td>
<td>13</td>
</tr>
<tr>
<td>6 Crowdsourcing Approaches</td>
<td>14</td>
</tr>
<tr>
<td>6.1 Recruiting Methods</td>
<td>14</td>
</tr>
<tr>
<td>6.2 Meter Manufacturer Outreach</td>
<td>19</td>
</tr>
<tr>
<td>7 Conclusions</td>
<td>20</td>
</tr>
<tr>
<td>7.1 Methods of Recruiting Citizen Measurers (CMs)</td>
<td>21</td>
</tr>
<tr>
<td>7.2 Quality of Data Collection</td>
<td>21</td>
</tr>
<tr>
<td>7.3 Recommendations and Future Work</td>
<td>23</td>
</tr>
<tr>
<td>8 References</td>
<td>23</td>
</tr>
<tr>
<td>Appendix - Data Collection Web Form</td>
<td>25</td>
</tr>
</tbody>
</table>
ABSTRACT

Miscellaneous Electric Loads (MELs) have proliferated in the last decade, and comprise an increasing share of building energy consumption. Because of the diversity of MELs and our lack of understanding about how people use them, large-scale data collection is needed to inform meaningful energy reduction strategies. Traditional methods of data collection, however, usually incur high labor and metering equipment expenses. As an alternative, this paper investigates the feasibility of crowdsourcing data collection to satisfy at least part of the data collection needs with acceptable accuracy.

This study assessed the reliability and accuracy of crowd sourced data, by recruiting over 20 volunteers (from the 2012 Lawrence Berkeley Lab, Open House event) to test our crowdsourcing protocol. The protocol asked volunteers to perform the following tasks for three test products with increasing complexity - record power meter and product characteristics, identify all power settings available, and report the measured power.

Based on our collected data and analysis, we concluded that volunteers performed reasonably well for devices with functionalities with which they are familiar, and might not produce highly accurate field measurements for complex devices. Accuracy will likely improve when participants are measuring the power used by devices in their home which they know how to operate, by providing more specific instructions including instructional videos. When integrated with existing programs such as the Home Energy Saver tool, crowdsourcing data collection from individual homeowners has the potential to generate a substantial amount of information about MELs energy use in homes.

1 Introduction and Background

Miscellaneous Electric Loads (MELs) have proliferated in the last decade, and comprise an increasing share of building energy consumption (DOE 2012). These trends underscore the need for effective efficiency measures, and information is needed to better understand the energy use of MELs. Because of the diversity of MELs and our lack of understanding about how people use them, large-scale data collection would inform meaningful energy reduction strategies. However these studies usually incur high labor costs and metering equipment expenses. MELs are ubiquitous and likely found in every home, making crowdsourcing an appealing alternative, to obtain at least part of the data needed to understand--and eventually reduce--MELs energy consumption.

Crowdsourcing, or citizen science, has been effectively employed for many years for various scientific and nonscientific causes, such as for the Search for Extra Terrestrial Intelligence
@Home (SETI@Home), which uses Internet-connected computers to analyze radio telescope data. The Internet and mobile technology have made crowdsourcing even more powerful in mobilizing large numbers of people for data gathering. For example, bird observation data submitted by amateur and professional bird watchers were instrumental in building a bird population database, which would not have been possible without crowdsourcing. The Cornell Lab of Ornithology and National Audubon Society later launched eBird, an online database where citizen scientists can submit data to inform bird abundance and distribution at various spatial and temporal scales. Several scientific papers (Fink et al. 2013 and Kelling et al. 2013) have been published based on the eBird database, demonstrating the value of eBird as an important tool for scientific contribution.

The objective of this project was to test and refine the methodologies for collecting MELs energy data through crowdsourcing. We aimed to better understand two research questions, in order to optimize crowdsourcing data collection and inform energy reduction strategies.

- How well can members of the public collect data of interest?
- What is the best method to recruit volunteers to participate in data collection efforts?

2 Project Activities and Report Outline

This project included two main tasks. The first task involved assessing the quality of crowdsourced data by inviting volunteers to collect MELs power consumption information and verify their measurements. Rather than doing field measurements in homes, we had test subjects measure power use under controlled conditions, on a set of products for which the power modes and levels were well understood by the research team. Due to the short time-setup of the experiment, we had participants perform power measurements on test devices rather than energy data, but the goal was to use these results to refine methodologies for collecting MELs energy data. In the first phase, we recruited LBNL staff as volunteers to participate in controlled experiments, with the goal of testing and improving our initial protocol for a second deployment. In the second phase, we deployed the improved protocol and recruited volunteers at the 2012 LBNL Open House, which took place on October 13, 2012. Descriptions and findings from the two experiments are documented in Sections 3, 4, and 5 of this report.

The second task of the project asked how to best recruit volunteers to participate in MELs data collection efforts. We researched a number of available recruiting schemes and assessed each method for crowdsourcing, evaluating parameters such as whether the volunteers would be representative of the general population, protection of participant confidentiality, etc. These findings are discussed in Section 6 of this report. Section 7 presents the conclusions and future work in this subject area.

---

3 Controlled Experiment

Eight individuals participated in the controlled experiments in Building 90 at LBNL. These tests allowed the initial definition and refinement of the study protocol, including:

- Selecting devices for testing,
- Developing user instructions,
- Targeting key data for collection,
- Fine-tuning the data entry web-form for user input, and
- Streamlining video recording of experiments to supplement data collection and assist in resolution of post-experiment questions.

3.1 Device Selection

The project team considered a number of electronic devices for inclusion in the protocol, including computer monitors, music docking stations, computer speakers, rechargeable power tools, and others. We designed the protocol to begin with a relatively simple device, and progressively scale up to more complex devices, so that we would be able to explore the abilities and limits of the public to discern and measure the power states. After a series of testing, we selected the following devices:

- Level I - Lamp: As the first device presented for metering, the lamp represented a simple case of two power states: on and off.
- Level II - DVD player: The DVD player had four power states: on, off, play, and tray-out. (Note: “On” refers to when the DVD player is turned on but performing no other function such as “play” or “tray-out”)
- Level III - Digital picture frame: The picture frame represented a more complex power usage scheme, and included a variable brightness setting: off, on-low, and on-high. We included the “low” versus “high” setting to test if participants could recognize that the brightness setting can be adjusted to more than one level.

In addition, we requested that participants separately measure the power consumption of any detachable power supplies (the test picture frame used in this study used a detachable power supply).

3.2 Protocol Refinement

The series of controlled experiments afforded the project team the opportunity to refine the study protocol in a stepwise manner. Researcher recommendations and participant feedback were incorporated into the subsequent experimental round. The revisions made to the protocol and experimental tools had been notably reduced by the last two subjects, as the protocol reached its final form:

- Information sheet & cards,
• Study web site,
• Meter selection,
• Participant instructions,
• Web form,
• Experimental checklist,
• Researcher script, and worksheet, etc.

Significant improvements during the controlled experiments included the following:

• Clarified and consolidated user instructions, data input, and exit survey in a single online web form;
• Fine-tuned the experimental setup to its final configuration, as shown in Figure 1 (Section 4 below describes in detail how the setup was used for data collection);
• Integrated human subjects requirements; and
• Developed an auditable system to track and record distribution of incentives to study participants.

Figure 1 - Schematic of Experimental Setup

4 Open House Experiment

As discussed previously, our initial protocol underwent several iterations, based on the feedback we collected from both researchers and participants. We continued to recruit and test volunteers and improve the protocol until it stabilized prior to the Open House. We also moved
the protocol survey from a Google Form to a commercial survey site, SurveyGizmo, to ensure data recording reliability.

4.1 Setup

The Open House setup was similar to the experimental setup shown in Figure 1. Each table was about 3'x6', and a digital camcorder was mounted on a tripod at one end to capture the entire experimental station. Video recording was active throughout the course of the experiments, and we frequently checked the camera for participant anonymity, view angle, and adequate memory. Two tables were located five feet apart and faced different directions, in order to minimize noise and disruptions between the experimental stations (See Figure 2).

A different power meter was used at each table - the Belkin Conserve Insight and the "Watts up? PRO ES" (Wattsup Pro)\(^3\). Both meters are user friendly and designed for use for the average consumer, although the Wattsup Pro is also widely used in the plug-load research community because of its data logging feature. By pressing through a few screens, the user can view readouts of the test device's power consumption in Watts, accumulated energy use in kilowatt-hours, and the annual energy cost of operating the device.

Figure 2 - Two Experimental Stations at LBNL Open House

Figure 3 shows the setup at one experimental station. Each station used a power strip with four outlets; each outlet had a separate hard on/off switch. The power strip was located at one end of the table and plugged into the power meter. Three test devices were plugged into the outlets on the power strip. When the participant asked to perform measurements on a particular test device, the researcher switched on the device at the power strip, so that the test device was

\(^{3}\) Belkin Conserve Insight - http://www.belkin.com/us/F7C005-Belkin/p/P-F7C005.jsessionid=A8EE0A5FC1E43A04D60FB154763A10F0. Wattsup Pro - https://www.wattsupmeters.com/
connected to power while the other two test devices remained off. In effect, the power meter measured only the power strip and the test device.

Figure 3 - Experimental Station Setup

The main component of the test setup was the laptop, which the participant used to access the web form. The web form had instructions to guide the participant, served as the portal for data entry, and included the exit survey. Volunteers were provided a magnifying glass to ease reading device nameplate information, which is frequently very small type.

4.2 Recruiting

The test setups were located inside a tent and adjacent to the Standby Power Exhibit. Staff on the research team regularly showcased the Standby Power Exhibit at LBNL Open Houses to highlight the standby power consumed by common household electronics. The Exhibit displays common plug loads and allows users to interact with the devices and read power measurements for different power settings. The Exhibit draws people who are interested in learning more about standby power; by co-locating the crowdsourcing experiment in the adjacent space, the recruiting was able to benefit from these synergies.

A recruiting table was located at the entrance to the tent that contained the experimental setups (see Figure 4). We displayed two common power meters that could be used in the home to meter plug-load devices and determine annual energy consumption. A researcher staffed the table throughout the day to recruit volunteers and obtain oral consent from the participants. The oral consent script was required as part of the LBNL Human Subjects Protocol, and contained: 1) a brief description of our study, 2) the subject’s consent to use the collected data for scientific results, and 3) consent to perform videotaping of their hands and the experimental process. Recruiting was successful; the study hosted a continuous flow of participants and both experimental stations were occupied throughout the entire Open House event.
4.3 Experiment

After the volunteer consented to participate in the experiment, they were assigned to one of the two test setups. Researchers briefly explained the purpose of the test, and introduced the volunteer to the different elements of the setup, including the web form, power meter, and test devices. See Figures 5 and 6. The researchers at each table encouraged the participant to follow the instructions on the web form as closely as possible, enter the data they generated, and ask any questions that might arise. While participants proceeded with the experiment, researchers timed each step (such as how long it took them to identify product information and to meter each device). Researchers also noted any observations or problems encountered.

4.4 Web Form

The web form (Appendix A) prompted the participant to fill in the number assigned to them by the recruiting staff. This became a unique identification number, linking the web form data, video recording, and incentive award. The second set of questions directed the participant in completing the meter-related information, including meter manufacturer name and model number. The third set of questions were product-specific, and asked the participant to identify: 1) the product type, 2) model number, 3) rated power, and 4) available power settings, and subsequently to measure the power consumption at each power setting. The web form also asked if the product had a power adaptor. The web form concluded with an exit survey, which elicited participant feedback and demographic information.
5 Findings

Below we summarize our preliminary findings from the Open House experiment and describe how they can inform MELs crowdsourcing data collection in homes in the future.

5.1 Meter Identification

We asked the volunteers to identify the manufacturer name and model number of the power meter. For the Belkin Insight, meter information is only available near the meter plug, and one
would need to unplug the meter in order to read the information. To avoid confusion, we attached a fictitious label near the meter screen that read “Belkin xz123.” Overall, the volunteers did well on identifying meter manufacturer and model number, especially for the Belkin since the fictitious label was in an obvious location. For the Wattsup Pro, 3 out of 8 volunteers put “Electronic Educational Device, Inc.” as the meter manufacturer, while the rest filled in the model name in this field. There appeared to be confusion between manufacturer and model names, especially in this case as the model number was clearly listed on the meter, and the manufacturer name could only be found on the nameplate located on the back.

5.2 Product Identification

As indicated previously, the three test devices were a lamp, DVD player, and digital picture frame; the last of which had an external power adaptor. All volunteers did very well in identifying the product type and manufacturer name for all three products. For product model number, 14 of the 18 volunteers recorded the model number on the DVD player correctly, whereas one recorded a partial model number, two identified the serial number as the model number, and one could not locate it. All but one of the 18 volunteers recorded the model number on the digital picture frame correctly. Finally, 16, 14, and 6 of the 18 volunteers correctly identified and recorded the rated power for the lamp, DVD player, and digital picture frame, respectively. Many appliances list their rated voltage and current (amps) rather than power; this was the case for the digital picture frame. Since volunteers were less familiar with power expressed in volts and amps, many of them recorded the volts but did not list the amps.

5.3 Mode Identification

As shown in the web form, we asked volunteers to independently identify up to four available power settings for each device. The results are plotted in Figure 7. For the simplest device - the lamp, almost all volunteers were able to identify both the on and off modes. A majority of the volunteers also identified the on and standby modes of the DVD player correctly, but only a third detected the DVD tray-out as a separate power mode. For the digital picture frame (DPF), most volunteers recognized the on mode as one mode and the power adaptor being unplugged as a separate mode. Several volunteers discovered the brightness adjustment button, which was located on the side of the digital picture frame, but only one recognized it as a different power-consuming state and measured the power draw separately.
5.4 Measurements

Along with the ability to identify available power settings of devices, we assessed the ability of volunteers to accurately perform the corresponding power measurements. Figures 8 through 12 display the measured power reported by volunteers at selected power settings for the three test devices. As shown in Figure 8, the measured on-mode power for the lamp ranges between 15 and 17W. Some of the variation resulted from the two light bulbs with slightly different power ratings in each of the experimental stations. In addition, we also found that measurements may vary slightly depending on when the measurement is taken, due to bulb warm-up. Possible variations in power measurements when taken at different times, or caused by heating, filtering, and other processes, should be considered as large-scale crowd-sourcing protocols are developed, to ensure data quality and usability. For example, the protocol should specify how long the participant should meter different types of products, and whether they should average the readings taken over time.

For the on-mode power of the DVD player, measurements reported by volunteers were consistent throughout (Figure 9), except for one person who reported the measured on-mode power as off-mode power, and vice versa (Figure 10). Figure 11 shows the reported power measurements when the DVD player was in the tray-out mode. There are only six data points, because only six volunteers successfully identified and measured this power mode. Five of the six volunteers reported measured power between 4.5 and 5.2W, one volunteer reported 8W (the volunteer did not take the measurement until after the DVD tray closed again, therefore the power measurement reflected the on-mode power).
For the on-mode measured power of the digital picture frame (Figure 12), there are primarily two sets of data points - one at the 1.7W level and the other around 4W. This is because the brightness setting of one picture frame was set to full brightness, while the setting of the other picture frame was between high and low brightness (power draw of LCD displays varies directly with the screen brightness). Some volunteers were not actually aware that the brightness could be adjusted. In an actual crowdsourcing deployment, device state could be difficult to detect. These and similar issues should be considered when developing a large-scale crowdsourcing protocol.

Figure 8 - Measured On-mode Power for Lamp

![Figure 8](image1)

Figure 9 - Measured On-mode Power for DVD player

![Figure 9](image2)
Figure 10 - Measured Standby Power for DVD player

Figure 11 - Measured Tray-out Power for DVD player
5.5 Measurement Time for Each Test Device

The research protocol included measurements of the time required for volunteers to perform the measurement, as well as to identify meter and device information. Table 1 summarizes the length of times all 18 volunteers spent on the test device, based on how long they spent on each device and answering the questions asked on the web form; the questions included collecting device characteristics (manufacturer name, model number, and rated power), identifying available power settings, and performing corresponding power measurements.

The participants required, on average, about the same time to measure each device (see Table 1). The lamp was the simplest device, but since it was the first device to test and meter, volunteers likely spent more time familiarizing themselves with the power meter and the test setup. There is a wide range between the minimum and maximum measured times across all three devices, showing a spectrum of abilities among volunteers. Although the volunteers were not expected to be a representative sample, the data provided valuable insight on data collection times. Given that the volunteers encountered these particular device models for the first time at the Open House, crowdsourcing participants will probably need less time for products in their own homes because they may be more familiar with the device functions. In preparation for the Open House launch, we estimated that the experiment would take volunteers between 20 and 30 minutes. For the 18 volunteers we recruited, the actual time it took to complete the experiment ranged from 11 to 33 minutes.
Table 1 - Time Needed to Collect Device Data, per Participant (Minutes)

<table>
<thead>
<tr>
<th></th>
<th>Lamp</th>
<th>DVD Player</th>
<th>Digital Picture Frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>5.0</td>
<td>5.4</td>
<td>5.5</td>
</tr>
<tr>
<td>Median</td>
<td>4.9</td>
<td>5.5</td>
<td>4.8</td>
</tr>
<tr>
<td>Min.</td>
<td>2.8</td>
<td>2.8</td>
<td>3.5</td>
</tr>
<tr>
<td>Max.</td>
<td>7.8</td>
<td>8.1</td>
<td>11.3</td>
</tr>
</tbody>
</table>

6  Crowdsourcing Approaches

6.1  Recruiting Methods

As part of the project, we conducted a survey of potential crowdsourcing recruiting schemes and evaluated each according to factors that are most relevant to crowdsourcing data collection, i.e. scale, direct and indirect costs, flexibility, and participant confidentiality. In this discussion, direct costs include those paid out to participants as compensation and direct material costs such as meter acquisitions, while indirect costs refer to advertising costs, and labor and materials costs used for the recruitment.

6.1.1 Amazon Mechanical Turk (AMT)

AMT is an online marketplace for work that requires human intelligence, where requesters post simple jobs called Human Intelligence Tasks (HITs) to be completed by workers, also known as “Turkers.” Workers select HITs to perform, and after completing and submitting them, they are paid via Amazon. This service was created in 2005 as a “crowdsourcing” web service for human computation, based on the idea that there are still many tasks human beings can perform much more effectively than machines, e.g. identifying objects in a photo or video, or researching data details. These activities usually require only a few minutes to complete, and workers are normally paid as little as a few cents for each task.

Amazon reports having more than 400,000 registered workers and 50,000-100,000 HITs at any given time (Ross et al. 2010). The minimum wage a worker is willing to work for is less than $2 per hour (Horton and Chilton 2010). In the first few years of AMT’s existence, the majority of the workers came from the United States; average household annual income was below $30K, but some people reported much higher earnings (Mason and Suri 2010). AMT was not the first source of income for most of the worker population, but the demographics are rapidly changing. In fact, young, well-educated Indians, who treat Turking as a full-time job, are becoming a progressively larger component of the Turker population since Amazon started paying workers in rupees in 2007 (Ross et al. 2010).
In the last few years, many researchers have used AMT to conduct behavioral research because it offers significant advantages compared to traditional lab experiments. The most important benefits are: wide and diverse population, very short design-result cycle, low cost, supportive web infrastructure, and complete anonymity (but identifiability) of workers (Greenblatt et al. 2013, Horton et al. 2010, Mason and Suri 2010, Paolacci et al. 2010). Paolacci et al. (2010) showed that they could reproduce judgment and decision-making experiments using AMT, with results very similar to standard lab tests, demonstrating the reliability of the workers and the validity of the experiment.

i Scale

AMT grants access to a wide population of workers with diverse demographics (Ipeirotis 2009) located both in the US and internationally. Although accuracy of demographic information is based on unverified user input, AMT survey options allow researchers to target a specific segment of the population. AMT also allows the creation of a two-part survey. In the first part, demographic data of a wide population (e.g., >1000 people) could be collected with an inexpensive survey (e.g., $0.05/person). Next, a subsample of that population with specific demographics (e.g. a sample matching the US population or a particular segment of interest) could be selected for the actual survey. However, Turkers may be skewed towards demographic groups that are younger and with lower household incomes.

ii Direct and indirect costs

Compared to other survey methods discussed in this section, AMT has low direct and indirect costs. First, the established AMT system makes the setup of a HIT and the recruitment of participants relatively straightforward. Secondly, even with a large number of participants, the direct cost is quite low as an average task on AMT pays less than $10 per hour. Finally, because AMT HIT-related infrastructure and subject payment procedures are done by Amazon, this data collection method incurs minimal administrative costs for processing payments to participants. The development and approval of an LBNL Human Subject Protocol must be completed; however this is required for all other recruitment outlets as well. Previous LBNL research projects using AMT have been approved by the LBNL Institutional Review Board (the body that approves human subject protocols).

iii Flexibility

Because of the low cost to launch a HIT, researchers could easily conduct a series of data collection stages. This would allow continuous refinement of each activity to improve data collection quality.

iv Privacy and Confidentiality
AMT has the unique feature of being able to identify subjects (for follow up questions or to avoid having multiple survey completed by the same person) but keeps the identity anonymous (Amazon provides a unique identifier but does not release personal information). The established AMT transaction setup and privacy policies would also streamline the development and approval of the LBNL Human Subjects Protocol.

6.1.2 Google Advertising

AdWords is a targeted advertising service on Google. An advertiser can specify keywords of interest and Google users who search using those keywords in the Google search engine, or who view related content on web sites that display Google ads, would see the purchased advertisements on the web page they view. AdWords allow customers to appeal directly to an audience who is already interested in their advertising topic. The AdWord advertisement could be linked to a web site run by the crowdsourcing research team, with a form available for participants to sign up for the data collection.

i Scale

Google AdWords is ubiquitous in the Internet, so this approach potentially has immense scale. The limiting factor is selecting keywords that will reach web browsers who are looking for information on similar topics and are likely to click on the advertisement. People searching for content on energy efficiency (or related terms) may also be the people likely to take part in the crowdsourcing data collection, but few of them might see the ad or be likely to follow the link to find out more. Effectiveness of a campaign using AdWords is difficult to determine (an entire industry is devoted to this question) and would also depend on the length of time the AdWords advertisement displays. Once a campaign is launched, AdWords generates reports to determine the advertisement’s performance on a site-by-site basis; researchers could then identify well-performing sites and further focus the advertising efforts.

ii Direct and Indirect Costs

The direct cost of AdWords is the cost of advertisement, which is only charged when the user clicks on the ad. The minimum cost is one cent per click on the advertisement, which amounts to $100 per 10,000 potential participants viewing the recruiting web site. The total cost of the recruiting campaign would then depend on the rate at which web site visitors actually sign up to participate in the project. If 1% of the visitors convert to participants, the cost of the advertising campaign would be $1 per participant. Indirect costs include the design and setup of the advertising campaign. Researchers would also need to build a web interface to which the advertisement would link, displaying information about the crowdsourcing project and allowing participants to sign up.

iii Flexibility
AdWords continues to generate performance data throughout the course of the campaign, and keywords and/or specified advertising sites can be revised and fine-tuned as needed to improve success. The advertisement can be stopped at any time.

iv Privacy and Confidentiality

Privacy and confidentiality are not directly a concern with AdWords, since it is only an advertising avenue for recruiting participants. Privacy and anonymity would be an issue, however, when the volunteers conduct energy measurements and submit data for research purposes.

6.1.3 Search for Extra-terrestrial Intelligence at Home (SETI@Home)

SETI@Home is an internet based public computing project, in which citizen scientists can lend the spare processing cycles of their personal computer to analyze potential evidence of radio transmissions from extra-terrestrial intelligence. The software platform is hosted by the Space Sciences Laboratory at UC Berkeley, using observational data collected by the Arecibo radio telescope. To date, the project has not confirmed any extra-terrestrial intelligence, but has identified a number of target areas for observations (http://setiathome.berkeley.edu/).

SETI@Home claims to be the most successful public participation science project in history, with more than 6 million volunteers participating over the project's first 10 years, through 2011 (Korpela et al. 2011). Given this large user base, the SETI@Home project has considered other uses for their software infrastructure, among them collecting data on energy use of home appliances. Staff from LBNL and SETI@Home discussed the feasibility of distributing power meters to its volunteers to collect power data in their homes. Numerous logistical obstacles appeared, such as standardizing meters (and paying for them) and data confidentiality. A pilot project would be needed in order to determine if these obstacles could be overcome.

6.1.4 Social Media and Green Blogs

Social media is quickly becoming an effective marketing tool for many products and services; notable examples include Facebook, Twitter, Google Plus, and LinkedIn. Since green blogs and social media represent similar recruiting outlets, they are discussed in this section. Many green blogs now receive regular readerships and are hosted by a wide range of organizations, including non-profit organizations, student clubs, environmentally-conscious individuals, and others; examples include The Stanford Energy Club, treehugger.com, and greendrinks.com.

i Scale

Social media has the potential to reach out to a large population, as users share data collection calls on their widespread networks. The recruiting call may also be announced under relevant

sub groups within social media, targeting users who have a higher potential for participation. However, social media like Facebook was shown to have a short attenuation period, such that the power of the message peaks and also recedes quickly (Robson 2012). Users who actually respond to the data collection call may be the ones who are familiar with the topic or have a personal interest; otherwise, they may just ignore the advertisement as the information stream refreshes and turns over very quickly in most social media.

Green blogs on the other hand, may reach out to a smaller group of users than social media, and the subscriber base may be more dedicated to environmental issues. This could result in a higher participation rate. Since there are often multiple blog entries throughout the course of the day, green blogs have a high information refresh rate, similar to social media. One concern with recruiting through green blogs or other special interest groups is that their energy usage patterns may not be representative of the national population as a whole.

**ii Direct and indirect costs**

There is almost no direct cost for posting advertisements on social media and green blogs, after the data collection call has been drafted and circulated around different user groups. However, compared to AMT for example, the indirect costs would be higher in mounting a campaign, and setting up communication and data exchange interfaces, because the research team would need to directly communicate with all participants once the initial contact was made.

**iii Flexibility**

Since there is no direct cost in sending out data collection calls, social media and green blogs are quite flexible in launching multiple data calls including test runs.

**iv Privacy and Confidentiality**

Privacy and confidentiality are a major concern when it comes to launching a campaign on social media and green blogs. We have to protect privacy and maintain anonymity of the participants who submit energy data, but we cannot accomplish this directly on social media and green blogs since identity is not protected in these outlets. We could, however, use social media and green blogs only for advertising, and direct interested participants to another data submission web interface, through which they could anonymously submit data and information.

**6.1.5 The ENERGY STAR Web Site**

The ENERGY STAR web site already receives regular traffic from consumers who are interested in purchasing energy efficient products, and it would seem like a viable avenue to recruit potential crowdsourcing participants. Because the web site currently does not allow for product data submission from consumers, it would serve as an advertising medium, similar to social media and Google Advertising, which then directs participants to another data collection web interface that protects privacy and confidentiality during data submission. We are not aware
that Energy Star has ever directly solicited information from its web site users, or referred them to research programs, so a policy change in the program might be needed.

i  Scale

Users who visit the ENERGY STAR website are likely consumers who are conscious of energy use and interested in purchasing energy efficient products. Thus, potential participants recruited from the ENERGY STAR website may be more knowledgeable about the energy use of electronic products. This could ensure higher data quality, as they may be more familiar with typical product energy usage and operating modes, however, these qualities may also skew the data since this user group may use more efficient products for testing and have different usage patterns.

ii  Direct and indirect costs

There is little direct cost associated with using the ENERGY STAR website as a crowdsourcing recruiting platform. Indirect costs include: 1) seeking the approval of the ENERGY STAR team, and 2) recruiting setup.

iii  Flexibility

Since ENERGY STAR is an official government website, we could only post contents that are finalized rather than in the testing stage. In this regard, the ENERGY STAR website is less flexible and is best suited for a finalized recruiting launch.

iv  Privacy and Confidentiality

Similar to social media and green blogs, to protect privacy and confidentiality, we may only be able to use the ENERGY STAR website as an advertising platform, and direct interested participants to a separate web interface.

6.1.6 Building America and other existing partnership programs

With the exception of AMT, the major disadvantage of the methods discussed above is lack of a system in place to protect anonymity of the study participants and to issue payments. The Building America program has thousands of research homes built over the years, for which homeowners have already agreed to participate in long-term energy studies. Adding MELs crowdsourcing as another component to these ongoing study sites might be feasible, and would save on administrative and data collection costs. Thus, recruiting participants through the Building America program may be an effective approach, especially during the early testing and deployment phases of MELs crowdsourcing data collection.

6.2 Meter Manufacturer Outreach
Another method of recruiting crowdsourcing participants is through collaborations with power meter manufacturers, which could be made possible by setting up a data collection scheme between meter manufacturers and consumers. For example, when purchasing a power meter, the meter manufacturer could direct potential citizen measurers (CMs) to a data depository website, where CMs would submit energy data collected from devices used in homes. Meter manufacturers would benefit from increased meter sales through information sharing among meter-purchasing CMs; we could also advertise meter sales through one of our crowdsourcing advertising schemes discussed above.

To test out the feasibility of this method, we reached out to contacts at Belkin, the maker of metering devices including the Belkin Conserve meters. Our discussions focused on developing a partnership that would mutually benefit meter sales and promote crowdsourcing data collection. One possible scheme would be for us to direct CMs to purchase meters from the Belkin website, and upon submitting proper MELs meta data and measurements, we would provide a rebate to cover a part or the entire cost of the meters. However, the Belkin representative we spoke with could not perceive a business opportunity on their part for such partnerships.

In recent years, there has been a steady increase of power meters available in the market targeted for device measurements in homes, ranging from the Kill-a-watt and Wattsup series to wireless metering systems. Some of these plug-load meters, such as the Modlet (themodlet.com) or Energy Hub (shop.energyhub.com), are capable of sending measurements to online databases. As the communication technology matures for commercially-available wireless power metering systems, outreach to manufacturers should be resumed again to explore the mobile data-submission option.

A related consideration to crowdsourcing recruiting is how meters can be distributed to volunteers for device measurements. We imagined several scenarios, ranging from gifting the meters to the volunteers to piggy-backing crowdsourcing on direct meter sales, such as the Belkin scenario described above. Alternatively, we could rely on volunteers who have already bought meters of their own, which would select for certain subsets of the population. Meter costs could range from zero (when piggybacking on a meter retailer or when the CM already has a meter) to the full cost of the meter ($50 to $200).

## 7 Conclusions

The proliferation of consumer electronics in the past decade, combined with the success of efficiency policies on white goods, has resulted in the rapidly increasing share of MELs energy consumption in buildings. MELs consist of a wide variety of product types including consumer electronics, which generally have a higher turnover rate and short product lifetime. As a result, traditional methods of data collection prove to be expensive, in developing up-to-date and accurate assessments of MELs energy consumption. As an alternative, this report investigates
the feasibility of crowdsourcing data collection to satisfy part of the data collection needs with acceptable accuracy. Below are the steps that must be assessed and demonstrated before crowdsourcing can be accepted as a viable data-gathering technique. This investigation focused on the first and second steps.

1) A wide population of participants can be recruited, and suitable meters can be distributed
2) Participants can reliably collect product characteristics and measurements
3) Participants can reliably input and upload data

7.1 Methods of Recruiting Citizen Measurers (CMs)
We investigated a number of platforms to potentially recruit Citizen Measurers (CMs) to perform the measurements, which included Amazon Mechanical Turk, Google AdWords, SETI@Home, social media and blogs, Energy Star, and Building America. In general, we found that there were tradeoffs between scale, representativeness, costs, privacy, and confidentiality.

Relying on special groups (such as environmental groups or university energy groups) would probably yield more carefully controlled data, and is more likely to provide richer contextual information about the household and the occupants, but these groups could be demographically limited and may skew data to more efficient device types and usage patterns. Homes in the Building America program, which already have a working agreement set up in place, would be an attractive target for crowdsourcing data.

Distribution of meters to the CMs is another consideration. We imagined several scenarios, ranging from gifting the meters to the CMs to piggy-backing crowdsourcing on direct meter sales. Alternatively, we could rely on CMs who have already bought meters of their own, which would select for certain subsets of the population.

Protecting the privacy and confidentiality of the CMs is another important issue. Except for AMT, with all other recruiting outlets it is difficult if not impossible to maintain anonymity of the CMs during data submission, CM payment, and subsequent data analysis. Administering payments to the CMs, if required, could become another obstacle. Some strategies, such as AMT, require payment to the CMs. Fortunately, Amazon also handles the billing so that administrative aspects are minimized. We conclude that, AMT and similar services, which have an established billing system and mechanism in protecting the anonymity of their users, are attractive venues for acquiring measurement data. Other than DOE and its contractors, Non-governmental organizations (NGOs), such as the Home Energy Magazine and other groups, are also candidates to carry out the data collection and analysis.

7.2 Quality of Data Collection
Crowdsourcing as a data collection strategy will not be viable if the CMs provide data that are filled with errors and gaps, so a major part of our research aimed to assess the quality of data collected by CMs. We recruited 18 CMs at the 2012 LBNL Open House; using a power meter
we assigned them, each participant recorded nameplate data of the power meter and test devices, identified the power modes of three devices, and measured and reported the power consumption for each mode. For this experiment, we examined the following four aspects of CM performance:

1. Ability to accurately record product characteristics data
2. Ability to reliably detect different power settings
3. Ability to accurately measure (and record) power consumption
4. Impact of increased familiarity with equipment and tasks

The first three aspects are directly related to data collection and entry; these are the data required to understand appliance energy use and to put the measurements in a broader context. Note that the sample size of 18 participants in the experiment is relatively small, and the results are not representative of the general population. Nonetheless, the results could inform the focus of future data-collection efforts and helps improve our crowdsourcing protocol.

**Product Characteristics** The CMs were able to accurately record the product type (i.e. lamp, DVD player, and digital picture frame). They were less successful in identifying the product model number; the overall error rate was 14% (22% for the DVD player and 6% for the digital picture frame). In recording the product rated power, some nameplates expressed their electrical characteristics in terms of volts and amps instead of watts. This confused some CMs, and the success rates were 90%, 80%, and 30% for the lamp, DVD player, and digital picture frame, respectively.

**Mode Detection** The CMs successfully identified the on, standby, and off modes with high success, ranging from 80-100% (see Figure 7). They made more errors as the devices become more complex and modes required taking special actions, such as the DVD tray in/out mode (success rate of 33%), and the brightness setting of the digital picture frame (success rate of 6%).

**Power Consumption** The CMs measured power consumption relatively successfully for the modes that they identified accurately. For example, only 1 of 18 measurements was erroneous and that was because one CM had transposed the standby and active power measurements.

**Product Familiarity** We observed that CMs were able to measure and record product data more rapidly as they progressed through the tasks, showing that people were learning and becoming more familiar with the power meter and the measurement protocol. We also noticed that CMs took a significant amount of time to study the products to understand the labeling and power modes. We suspect that measurement times and error rates will decline when CMs measure products in their own home, with which they are already familiar and have used regularly. For the measurement of complex products, we would need to develop more specific instructions (with e.g. instructional video) to guide them through product functionalities and power settings.
7.3 Recommendations and Future Work

Based on the findings of this project, we conclude that crowdsourcing data collection may not produce highly accurate field measurements, but has the potential to be a useful and inexpensive way to gather MELs data for simpler devices. Our assessment focused on power measurements, but energy data would be of greater interest to inform energy efficiency policies. Measuring energy use offers simplifications and new complexities. For example, identification and measurement of specific modes would not be a concern. On the other hand, engagement with the CMs will need to continue for longer periods, deployed meters would be required to have storage ability or be set up to transmit data to a central location, and finally we need a deeper understanding of how behavior may impact collected data quality.

With the CM recruiting logistics, we conclude that the selected methods should take into account payment to CMs, distribution of meters, and confidentiality, and that data collection is most efficient when conducted through projects/programs that already have a privacy and confidentiality and contract agreement in place, such as the Building America program homes.

This project served as a good first-step to assess the feasibility of MELs data collection through crowdsourcing and identified important next steps in order to refine a working protocol. For next steps, we recommend extending the crowdsourced testing to ongoing Building America projects, as an efficient way to further improve the data collection protocol developed as part of this project (adding video instructions) and to test energy measurements. We also recommend adding a data submission component to the Home Energy Saver™ (HES) tool, such that users could collect and upload MELs energy data for the devices used in their homes, following a version of our crowdsourcing protocol. The collected data can be used to improve MELs energy estimates of the HES tool, and will inform purchasing guidance for MELs devices. Finally, wireless power meters that enable device control through mobile and remote applications, such as the Modlet, should also be tested in this phase of the project.

8 References


http://www.eecs.berkeley.edu/Pubs/TechRpts/2012/EECS-2012-195.html.


Appendix - Data Collection Web Form
Crowdsourcing Experiment - Data Collection Form

Page One

1. Enter the number labeled on your table.*

2. Enter Meter’s manufacturer name:* 

3. Enter Meter’s model number, if available:

Step 1a - Enter Meter Information

Step 1b - Check Meter Display

Make sure meter display is showing measurements in Watts (W).

Note: All power measurements will be recorded in Watts (W). In some cases you may need to press a few buttons and switch display to show power in Watts, instead of voltage or other available units.

Step 2 - Device #1: Product Information

Product information can usually be found on the front, side, and/or back of the device.

4. Enter Product type:

For example, notebook computer, toaster, etc.

5. Enter Product’s manufacturer name, if available:
6. Enter Product’s rated power, as labeled on product:

Record rated power in Watts (W). If rated power in Watts is not available, record the rated input voltage (V) and amperge (A). *

7. Power Setting 1 ----- Select the closest power setting, or select Other and enter power setting description.*

- Off
- On
- Active
- Play
- Idle
- Sleep
- Standby
- Warm up
- Low
- Medium
- High
8. For Power Setting 1, enter corresponding measurement in Watts (W).

10. For Power Setting 2, enter corresponding measurement in Watts (W).

11. Power Setting 3 — Select the closest power setting, or select Other and enter power setting description.
12. --- For Power Setting 3, enter corresponding measurement in Watts (W).

13. Power Setting 4 ----- Select the closest power setting, or select Other and enter power setting description.

   - Off
   - On
   - Active
   - Play
   - Idle
   - Sleep
   - Standby
   - Warm up
   - Low
   - Medium
   - High
   - Other

14. --- For Power Setting 4, enter corresponding measurement in Watts (W).
15. If product has a power adapter, disconnect power adapter from product. Then measure and enter the power use of the power adapter alone.

This is an example of a power adapter. It may also be referred to as power supply, AC adapter, power brick, or other names.

Copy of Page One

For Device #2, repeat steps 2 and 3.

Device #2: Product Information

16. Enter Product type:

For example, notebook computer, toaster, etc.

17. Enter Product's manufacturer name, if available:
18. Enter Product’s model number, if available:

19. Enter Product’s rated power, as labeled on product:

Record rated power in Watts (W). If rated power in Watts is not available, record the rated input voltage (V) and amperage (A). *

Device #2: Measure and Enter Power Use

- Electronics usually have two or more power settings. Each power setting is a state with distinct functionality and respective power consumption.
- With product plugged into power meter, experiment with product and determine the available power settings.
- Use your best judgement to determine how many power settings each device has.
- Wait until power measurement stabilizes (it can take anywhere from 10 to 30 seconds), then take measurement.
- If the reading continues to fluctuate by a small amount after stabilizing, just record the closest average value.

20. Power Setting 1 ----- Select the closest power setting, or select Other and enter power setting description.*

- Off
- On
- Active
- Play
- Idle
- Sleep
- Standby
- Warm up
21. --- For Power Setting 1, enter corresponding measurement in Watts (W).

22. Power Setting 2 ----- Select the closest power setting, or select Other and enter power setting description.

23. --- For Power Setting 2, enter corresponding measurement in Watts (W).

24. Power Setting 3 ----- Select the closest power setting, or select Other and enter power setting description.

Options:
- Off
- On
- Active
- Play
- Idle
- Sleep
- Standby
- Warm up
- Low
- Medium
- High
- Other
25. --- For Power Setting 3, enter corresponding measurement in Watts (W).

26. Power Setting 4 ---- Select the closest power setting, or select Other and enter power setting description.

- Off
- On
- Active
- Play
- Idle
- Sleep
- Standby
- Warm up
- Low
- Medium
- High
- Other
27. --- For Power Setting 4, enter corresponding measurement in Watts (W).

28. If the product has a power adaptor, disconnect the power adaptor from the product. Then measure and enter the power use of the power adaptor alone.

This is an example of how a power adaptor looks like. It may also be referred to as power supply, AC adaptor, the brick, and other names.

**Copy of Copy of Page One**

For Device #3, repeat steps 2 and 3.

Device #3: Product Information

29. Enter Product type:

For example, notebook computer, toaster, etc.
30. Enter Product's manufacturer name, if available:


31. Enter Product's model number, if available:


32. Enter Product's rated power, as labeled on product:

Record rated power in Watts (W). If rated power in Watts is not available, record the rated input voltage (V) and amperge (A).*


Device #3: Measure and Enter Power Use

- Electronics usually have two or more power settings. Each power setting is a state with distinct functionality and respective power consumption.

- With product plugged into power meter, experiment with product and determine the available power settings.

- Use your best judgement to determine how many power settings each device has.

- Wait until power measurement stabilizes (it can take anywhere from 10 to 30 seconds), then take measurement.

- If the reading continues to fluctuate by a small amount after stabilizing, just record the closest average value.

33. Power Setting 1 ------ Select the closest power setting, or select Other and enter power setting description.*

- Off
- On
- Active
- Play
- Idle
34. --- For Power Setting 1, enter corresponding measurement in Watts (W).

35. Power Setting 2 ----- Select the closest power setting, or select Other and enter power setting description.

36. --- For Power Setting 2, enter corresponding measurement in Watts (W).
37. Power Setting 3 ----- Select the closest power setting, or select Other and enter power setting description.

- Off
- On
- Active
- Play
- Idle
- Sleep
- Standby
- Warm up
- Low
- Medium
- High
- Other

38. --- For Power Setting 3, enter corresponding measurement in Watts (W).

39. Power Setting 4 ----- Select the closest power setting, or select Other and enter power setting description.

- Off
- On
- Active
- Play
- Idle
- Sleep
- Standby
- Warm up
40. --- For Power Setting 4, enter corresponding measurement in Watts (W).

41. If product has a power adaptor, disconnect power adaptor from product. Then measure and enter the power use of the power adaptor alone.

This is an example of how a power adaptor looks like. It may also be referred to as power supply, AC adaptor, the brick, and other names.

New Page

Exit Survey - Feedback

Your answers to the exit survey are optional.
42. Please rank in order of importance

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>We enjoy having electronics and appliances with a lot of features and services.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>It is very important to us to save money on our energy bills.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>We think it is important to limit our use of energy to save resources and because of environmental effects.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

43. How would you rate the following tasks?

<table>
<thead>
<tr>
<th>Task</th>
<th>Very easy</th>
<th>Relatively easy</th>
<th>Neutral</th>
<th>Relatively difficult</th>
<th>Very difficult</th>
</tr>
</thead>
<tbody>
<tr>
<td>Locate the product manufacturer and model number.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Read the product manufacturer and model number.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identify the power settings of the product.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learn how to use the power meter.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measure device power consumption.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enter the measured data into the web form.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

44. Were the instructions easy to follow?

<table>
<thead>
<tr>
<th>Instructions</th>
<th>Very easy</th>
<th>Relatively easy</th>
<th>Neutral</th>
<th>Relatively difficult</th>
<th>Very Difficult</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

45. What was the most important thing that you learned from this exercise? (please check all that apply)

- [ ] How to use meters to measure power.
- [ ] Learning about product power settings.
- [ ] Increased knowledge of product power consumption and energy use.
46. Did you enjoy participating in this study?

- Strongly agree
- Agree
- Neutral
- Disagree
- Strongly disagree

47. How likely would you participate in similar activities again?

- Very likely
- Likely
- Neutral
- Unlikely
- Very unlikely

48. How likely would you participate in similar activities again?

- Very likely
- Likely
- Neutral
- Unlikely
- Very unlikely

49. Would you recommend this to a friend?

- Very likely
- Likely
- Neutral
- Unlikely
- Very unlikely

New Page

Exit Survey - Demographics

50. What year were you born in?

51. What is the highest degree or level of schooling that you have completed? If currently enrolled, mark highest degree received.

- 8th Grade or below
- Some high school
- High school or the equivalent (for example: GED)
- Some college credit
52. Are you Spanish/Hispanic/Latino?

- No, not Spanish/Hispanic/Latino
- Yes, Mexican, Mexican American, or Chicano
- Yes, Puerto Rican
- Yes, Cuban
- Yes, other Spanish/Hispanic/Latino

53. If not Spanish/Hispanic/Latino, what is your race? Mark one or more races to indicate what you consider yourself to be.

- White
- Black, African American, or Negro
- American Indian or Alaska Native
- Asian Indian
- Chinese
- Filipino
- Japanese
- Korean
- Vietnamese
- Other Asian
- Native Hawaiian
- Guamanian or Chamorro
- Samoan
54. What is the primary language you use at home?

- English
- Other

55. Please indicate the range in which your annual income falls:

- 0 - $24,999
- $25,000 - 49,999
- $50,000 - 74,999
- $75,000 - 99,999
- 100,000 - 149,999
- 150,000 and greater

Thank You!

Thank you for taking our survey. Your response is very important to us.