Energy Research & Social Science 8 (2015) 190–197

Contents lists available at ScienceDirect

Energy Research & Social Science

journal homepage: www.elsevier.com/locate/erss

Original research article

Energy efficiency and the misuse of programmable thermostats: The effectiveness of crowdsourcing for understanding household behavior

Marco Pritoni [a,b,*], Alan K. Meier [b], Cecilia Aragon [c], Daniel Perry [c], Therese Peffer [d]

[a] Western Cooling Efficiency Center, University of California Davis, 215 Sage Street, Suite 100, Davis, CA 95616, USA
[b] Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA
[c] Department of Human Centered Design & Engineering, University of Washington, 407A Sieg Hall, Box 352315, Seattle, WA 98195, USA
[d] California Institute for Energy and Environment, 2087 Addison Street, 2nd Floor, Berkeley, CA 94708, USA

ARTICLE INFO

Article history:
Received 30 July 2014
Received in revised form 1 June 2015
Accepted 2 June 2015

Keywords:
Survey
User behavior
Energy efficiency

ABSTRACT

Programmable thermostats are generally sold as energy-saving devices controlling heating and cooling systems, but can lead to energy waste when not operated as designed by the manufacturers. We utilized Amazon Mechanical Turk, an online crowdsourcing service, to investigate thermostat settings and behavior in households. We posted a survey and paid respondents to upload pictures of their thermostats to verify self-reported data. About 40% of programmable thermostat owners did not use programming features and 33% had programming features overridden. Respondents demonstrated numerous misconceptions about how thermostats control home energy use. Moreover, we found that 57% of households were occupied nearly all the time, limiting the potential energy savings. The study revealed flaws in self-reported data, when collected solely from traditional surveys, which raises concerns about the validity of current thermostat-related research using such data. "Ground truth" temperature data could now be available in homes with Internet-connected thermostats. Online crowdsourcing platforms emerge as valuable tools for collecting information that would be difficult or expensive to obtain through other means. Advantages over traditional surveys include low-cost, rapid design–implementation–result cycle, access to diverse population, use of multimedia. Crowdsourcing is more effective than alternative online tools due to easier recruitment process and respondents' reputation system.

1. Introduction

1.1. Surveying thermostat behavior

Maintaining thermal comfort in a home is a web of competencies and behaviors [1–3]. It requires a mental model of the technical relationship between the thermostat, the heating system, and the physical and social characteristics of the home. Equally important, maintaining thermal comfort involves a set of practices and behaviors encompassing much more than setting the temperature; these include choice of clothing, the use of windows, instructions to children, and the general perception of thermal comfort. The technical and behavioral aspects are both closely related and entangled. Installing a new thermostat is a relatively minor technical intervention; nevertheless, it causes a realignment of practices and behaviors. The first step in understanding this realignment is to examine how occupants use a new thermostat.

Programmable, digital thermostats are generally sold as "energy-saving devices" and have been purchased by more than a hundred million consumers in North America and Europe in the past two decades.[1] The reduction in energy consumption arises from the automatic use of energy-savings settings, i.e., turning off the heating and cooling system when not needed, or "setting back" the temperature in winter (or "up" in summer) during unoccupied periods and at night [4]. Simulations of residential buildings in North American climates suggest that lowering the thermostat 2–4 °C for 8 h per night saves 5–15% a year on the heating bill [5,6], although the savings depend strongly on thermostat control strategy, climate and other factors. But unlike an efficient boiler or a refrigerator, the mere installation of a (programmable)

* Corresponding author at: Western Cooling Efficiency Center, University of California Davis, 215 Sage Street, Suite 100, Davis, CA 95616, USA. Tel.: +1 530 752 0345; fax: +1 530 754 7672.
E-mail address: marco.pritoni@gmail.com (M. Pritoni).

[1] Despite the growth in sales of more advanced devices, such as Internet-connected thermostats, the number of programmable thermostats currently installed in homes is far higher (∼10 fold).

http://dx.doi.org/10.1016/j.erss.2015.06.002
2214-6296/

thermostat does not affect the house's energy use. Field studies have, not surprisingly, observed wide ranges in energy savings from programmable thermostats. Some studies found no significant difference in energy consumption between households with manual and programmable thermostats [6–9], some found an increased consumption with programmable thermostats [10,11], while a few others observed small energy savings [12–14]. The range of outcomes found in these studies may be explained by different climates and fuel use, small sample size, self-selection bias, and difficulty of establishing a baseline.

In 1995, the US Environmental Protection Agency (EPA)'s ENERGY STAR program established specifications for programmable thermostats, suggesting that homeowners could save about $180 a year with a qualifying programmable thermostat. ENERGY STAR requirements included certain features such as default comfort and energy-saving setpoint temperatures. Twenty years later, the EPA concluded that consumers were not using programmable thermostats effectively due to programming difficulties and lack of understanding of terms such as setpoint. As a result, the EPA discontinued the ENERGY STAR programmable thermostat program in December 2009 [4]. Around the same period, careful observations of thermostat behavior in Finland, the UK, Japan, and the United States have found that many people are confused and unable to operate a thermostat as expected by the manufacturers [15–19].

More reliable—and contemporary—data on indoor temperature settings are needed to support design of more effective policies, help educate consumers, and improve saving estimates. However, these data need to be collected in the context of thermal comfort preferences, household occupancy patterns, and people's interaction with the device [4]. The aim of this study is two-fold: (1) to investigate how people actually use thermostats in their homes, and (2) to assess the validity of crowdsourcing as a tool for collecting household thermostat usage behavior. Before delving in our results, the state of current research is summarized below.

1.1.1. Temperature settings

Energy simulations require reliable temperature setpoints to predict energy savings for a particular heating and cooling strategy. Information about temperature preferences and thermostat settings of US households (Table 1) have been compiled in several large energy surveys (600–30,000 participants), including the Residential Energy Consumption Survey (RECS) [20–22] and the American Home Comfort Study (AHCS) [23]. In addition, many utilities and regional organizations conduct their own surveys, notably the California Residential Appliance Saturation Survey (RASS) [24,25], and Residential Customer Characteristic Survey (RCCS) in Seattle [14]. Energy surveys conducted by the European Union collect less information about thermostat settings [9,26]. The European Statistics Agency recommends making the collection of thermostat-related data for heating and cooling a high priority [27]. All of these surveys suffer from problems associated with self-reported data. Temperature settings are especially unreliable [28]. These surveys are also expensive, so they are conducted rarely (every 3–10 years), thus failing to capture technology and behavioral trends. Several other recent studies tried to examine indoor temperatures in dwellings in UK [29] and US [30]. The authors note the need for further investigations of thermal comfort both in the UK and elsewhere due to the limited data.

1.1.2. Household context

House routines and schedules can significantly limit the effectiveness of programmable thermostats. For example, in houses that are occupied all the time people are less likely to tolerate less comfortable temperatures. Irregular household schedules also represent a challenge for programmable thermostats, which are designed mostly to enforce a fixed schedule. Large surveys do not contain detailed data about home schedules or occupancy.

1.1.3. Interface-related behavior

Recent research showed that people find programmable thermostats confusing and cumbersome to program [19]. Difficulties in planning, programming and managing the setpoints through the interface might reduce saving potentials of programmable thermostats [19]. However, these experiments were conducted in a laboratory so little is known about what people do in their own homes. Lack of data and shortcomings of previous research led us to search for alternative approaches to collecting thermostat-related information.

1.2. Online crowdsourcing services as means of data collection

Starting in about 2005, online services began performing small, defined, tasks for clients. These are tasks that machines cannot complete easily (e.g., labeling images or music, transcribing audio recordings, removing errors from data, filtering blogs for adult content, and taking surveys). Each task usually requires only a few minutes for a person to complete, although some may be as long as half an hour. The online service matches customers with workers, and arranges for payment, data compilation, and quality control. The workers are paid according to the number of tasks that they accomplish. Several services have appeared since 2005, but Amazon Mechanical Turk is by far the largest. In 2014, Amazon Mechanical Turk claimed over 500,000 workers located in 190 countries and over 200,000 tasks currently on offer to its workers [31]. Similar online crowdsourcing services have recently become available in other languages, including French, German, and Japanese. A recent survey found over 75 services available around the world.[2]

Researchers have already used online crowdsourcing services to conduct behavioral experiments and compared their results to conventional techniques. For example, Paolacci et al. [32] obtained similar results from judgment and decision-making experiments through crowdsourcing and laboratory measurements. The most important benefits are: wide and diverse population, very short design-result cycle, low cost, supportive web infrastructure, and complete anonymity (while maintaining identifiability) of workers [32–35]. Crowdsourcing data has also been used in a growing number of energy-related fields. Bazilian et al. [36] surveyed research on crowdsourcing for energy analysis, noting the usefulness of such techniques in determining energy usage patterns and for replacing expensive sensors with human computation. The Mechanical Turk has also been used in the US to survey saturation of miscellaneous appliances [37,38] and characteristics and usage patterns of ceiling fans [39]. Ford and Karlin [40] studied energy feedback with this tool. However, thermostat behavior has not been previously evaluated through crowdsourcing approaches. The present study is unique because (1) it investigates thermostat use beyond setpoints, exploring how people actually use thermostats in the context of their homes; (2) it uses pictures uploaded by participants to validate the results of the survey and obtain other insights in energy behavior; (3) it examines crowdsourcing as a tool for collecting behavioral information about energy use in homes.

2. Methods

We collected thermostat usage data using an online crowdsourcing service (Amazon Mechanical Turk). The survey was

[2] http://publish.smartsheet.com/89f4ed2c1e9e45618870f62f17054dd7.

Table 1
Availability of thermostat information from surveys.

Information	RECS 2009	AHCS 2009	RASS 2009	RCCS 2009
% of households with no thermostat/manual/programmable	Y	Y	Y	Y
% of households doing day setback/night setback (cooling season)	Y[a]	N	Y	N
% of households doing day setback/night setback (heating season)	Y[a]	N	Y	Y
% of households actually programming the thermostat	N	Y	N	N
Detailed settings (time and temperature for each setting)	P	N	P	P
Frequency of overriding/hold	N	P	N	N
Thermal comfort/overall satisfaction with Heating Ventilation Air Conditioning (HVAC) and thermostat	N	Y	N	N
Household demographics (number of people, age, etc.)	Y[c]	N	Y	Y[c]
Schedule/daily occupancy	P[b]	N	N	P[b]
House energy consumption	Y[c]	N	N	Y[c]

Y, yes; N, no; P, partial data available (usually only temperature and not time).
[a] Only for programmable thermostats.
[b] Only whether house used for business, no info whether always somebody at home.
[c] Data aggregated by thermostat type are not available.

limited to Amazon Mechanical Turk workers with programmable thermostats living in the US. The survey consisted of two parts: a questionnaire about the worker's thermostat usage and the option to upload a photo of the worker's thermostat. Table 2 summarizes the survey parameters.

One advantage of online surveys is the short time between releasing the survey and obtaining results, which allows for frequent iterations. We performed three small-sample surveys in May 2010 with 63 subjects, but adjusted wording and questions at each iteration. The number of respondents was 3, 10 and 50 in the first, second and third surveys respectively. Ten pictures of residential thermostats were collected in this preliminary phase to verify the self-reported data. The value of iteration was demonstrated in the criteria for the photographs of the thermostats. The first batch of photos was unclear, so we required subsequent uploads to meet minimum readability criteria. Results from these preliminary surveys are not included in the present analysis.

The second round of the survey was limited to 200 respondents (collected in 7 days). Subjects were recruited through a task entitled: "Answer a short survey about your thermostat" with this description: "Complete the survey. Photograph the programmable thermostat used in your residence, and upload photo for $2 bonus." A reward of $0.40 was given for completing the survey, and a $2 bonus was offered for uploading a readable picture (increased from $1 in the preliminary phase). We collected a total of 31 pictures in this phase. Demographics were not collected in the survey, due to limitations set by the human subject protocol.

The survey was completely hosted by the crowdsourcing service, but we provided a link to a popular website (picturepush.com) to upload the pictures. The projected average hourly rate at the end of the survey was $4.00 (i.e., average completion time was 6 min[3]), plus the supplementary bonus for uploading a picture (a fairly high compensation compared to ordinary tasks).[4]

Several procedures were used to ensure reliable responses. First, workers were required to live in the US and have a quality rating greater than 90% (based on their performance of previous tasks with Amazon Mechanical Turk). Second, we asked three "nonsense questions" to guard against random answers to questions, that is, if the thermostat was attached to the ceiling, if the distance between computer and thermostat was more than one mile, and if the thermostat manufacturer was Ford or Sony (two companies that do not manufacture thermostats). Additional filtering was conducted based on completion time, clarity of photograph, and reported location (must be in the United States). We discarded a survey if it was completed in less than 90 s (the range of accepted surveys was 98–3039 s). These filters resulted in the exclusion of eight respondents (and one picture), reducing the pool to 192 subjects from 38 states (Fig. 1) and 30 pictures.

Table 2
Survey facts.

	Trial surveys	Actual survey
Period	May 2010 (2 weeks)	September 2010 (1 week)
Respondents	63	200
Pictures	10	31[a]
Drop-out	N/A	8
Reason for drop-out	Only for test	- 3 said thermostat is on the ceiling - 5 reported a non-existent brand of thermostat (1 of them was also located outside the US)
Cost	$25	$142

[a] Of those, one picture had to be removed because it showed a thermostat that is not sold in the US.

3. Results and energy considerations

We present here the major results of the final survey. Survey responses provide insights into use patterns, household routines and user understanding of terminology and operation of programmable thermostats. The full list of questions (and answers) designed to survey the type of thermostat, knowledge about its operation, and user behavior, is provided in the Appendix. The most significant outcomes are listed in Table 3. Results are compared to larger surveys, or to other available data in the literature. Responses were not weighted to correct for difference in demographics between the survey sample and the national population, since demographic information was not collected.

3.1. Programmable vs. manual thermostats

This study focused on programmable thermostats; however, 58%(±7%[5]) of the respondents reported having a *manual*

[3] $0.40 to work 6 min corresponds to an average hourly rate of $0.40/6 min × (60 min/h) = $/h 4.
[4] Amazon has subsequently added a feature that permits workers to directly upload photos or videos. This feature will greatly simplify investigations like this one.
[5] We report 95% confidence interval (normal approximation to the binomial distribution) here and in the rest of the paper. Confidence intervals represent only uncertainty due to sample size. Other sources of uncertainty such as recollection bias, estimation bias and others are not easily quantifiable, therefore they were not reported.

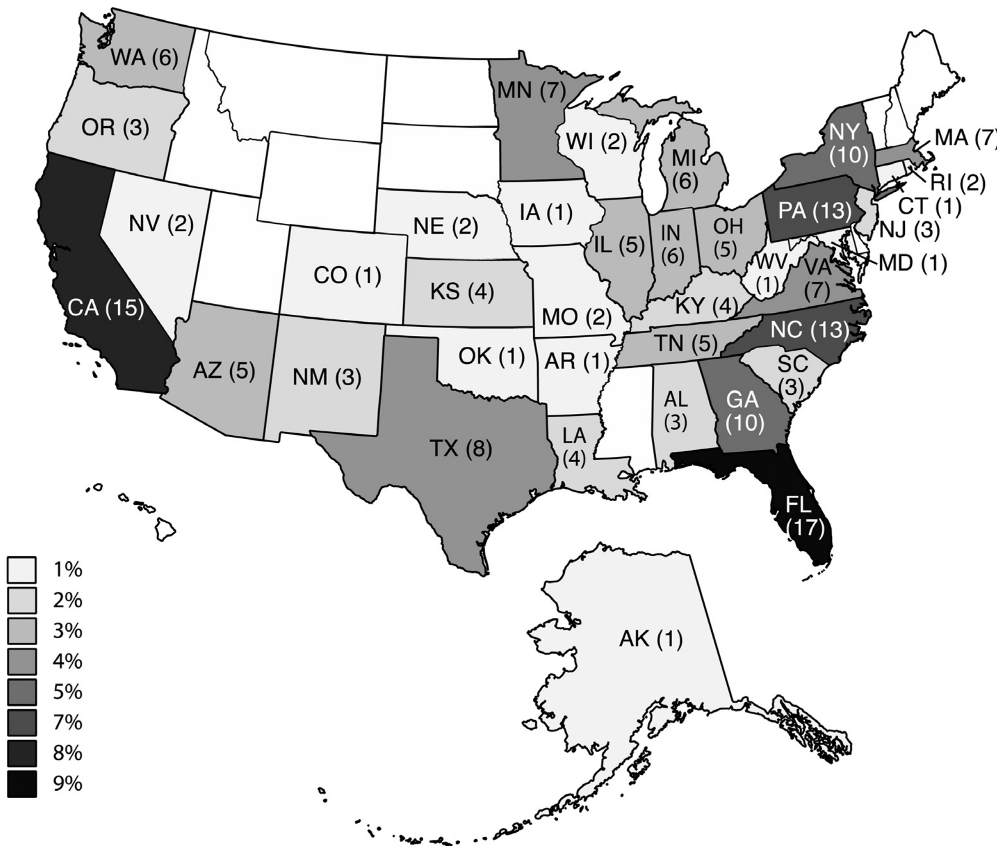

Fig. 1. Location of respondents. Actual numbers of respondents for each state are reported in parentheses.

thermostat. The remaining respondents (42% ± 7%) said they have a *programmable* thermostat that "can automatically change the temperature at different times of the day". These proportions are consistent with other US surveys, as shown in Table 4, even without adjusting for demographic bias (testing that the % of programmable thermostats in RECS, AHCS and this study are identical leads to $\chi^2 = 1.45$, $p = 0.52$). However, authors of large surveys noted that the response varied depending on how the question was framed—if asked "can you set it so that the temperature setting automatically changes at the times of the day or night that you choose?" the households reporting having a programmable thermostat dropped by nearly half compared to the question "do you have a programmable thermostat?" [41,42].

Our survey asked several questions to confirm the accuracy of the answer to the first question about the type of thermostat. These questions focused on the use and availability of programming features. For example, 4% (±3%) of those who reported having a programmable thermostat in the first question later responded that they do not have programming features (which means the thermostat is manual). These two answers are obviously incompatible. Analogously, 3% (±2%) of the subjects who affirmed having a manual thermostat later reported they use programming features. Moreover, out of 30 pictures collected, we verified that one thermostat reported as programmable had a digital display, but it was manual in operation, and one that was defined manual

Table 3
Most significant results of the survey.

Result	% and CI[5]
Respondents with programmable thermostats	42% ± 7%
Respondents with programmable thermostats that do not use programming features (self-reported)	39% ± 7%
they don't know where the settings are	*14%* ± 5%
they don't know how to change them	*25%* ± 6%
Households with somebody at home nearly all the time	57% ± 7%
Thermostats with (self-reported) time incorrect	15% ± 5%
Respondents who mistook the target temperature setting with the current room temperature in the programmable thermostats (from picture)	>20% ± 14%
Thermostats in HOLD (from picture)	33% ± 17%
of which respondents had claimed he/she used the programs	*50%* ± 18%
Thermostats not selected by the users	65% ± 7%
Users who do not know where the instruction manual is	70% ± 6%
Energy myths	
"Setting the thermostat at higher temperature will heat the house faster"	39% ± 7%
"Thermostat controls directly the temperature of the air coming out of the HVAC system"	62% ± 7%
"Turning the temperature down when the house is unoccupied or at night is more energy consuming than keeping it always at the same temperature"	36% ± 7%

Table 4
Manual vs. programmable thermostats in recent surveys.

Survey	Population	Sample size (households)	Manual[a]	Programmable
RECS 2009	US	4000	57% ± 2%	43% ± 2%
AHCS 2009	US	30,000	58% ± 1%	42% ± 1%
RASS 2009	CA	25,000	47% ± 1%	53% ± 1%
RCCS 2009	Seattle	600	49% ± 4%	51% ± 4%
AMT 2010	Turkers, US	192	58% ± 7%	42% ± 7%

[a] When surveys indicated a different values for heating controls and cooling controls, only the first one is reported.

Fig. 2. Thermostat in hold mode (arrow) in a household reporting the use of programming features.

was actually a mechanical setback thermostat, which allows programming. These findings may indicate unfamiliarity with the terminology, misunderstanding of the question or not being aware of the thermostat's capabilities.

3.2. *Use of programs/hold mode*

We asked the respondents with programmable thermostats how they used the programming features: 14% (±5%) of the subjects reported they "do not know where the settings are", and 25% (±6%) answered that they "know where the settings are but [they] do not know how to change them". Thus, about 40% (±7%) of programmable thermostat users do not operate their thermostat as expected by the manufacturers.

The photographs offered another important insight into the operating problems: about one third of the thermostats were in "permanent hold" mode; this mode interrupts all the programs and turns the programmable thermostat into a manual thermostat (Fig. 2). In addition, 50% (±18%) of the photos of the thermostats in hold mode was uploaded by those who claimed to use programmed settings. This further undermines the reliability of self-reported data. Large energy surveys do not collect this type of information, but our findings are in line with other previous studies reporting half of thermostats in hold mode [19,43,44].

In separate work, we found that occupant interaction with programmable thermostats is difficult to describe and categorize, but has important consequences for energy use [45]. Some users "set the device and forget", while others continually fiddle with it. Some people use programmable thermostats as if they were manual thermostats, i.e., holding the temperature at desired values and manually turning the unit on and off. Other people set the programs just once per season. Sometimes users do not fully comprehend their actions and whether their thermostat is programmed or not.

These considerations confirm that the installation (and presence) of a programmable thermostat does not assure that it will be programmed to automatically set back the temperature. They also underscore the need for improved user interfaces and education if programmable thermostats are to be expected to reliably save energy. Outcault et al. [46], in a recent paper, also confirmed that significant training and support is needed to effectively operate (advanced) programmable thermostats.

3.3. *Who sets the thermostat?*

Not surprisingly, about 50% (±7%) of the respondents reported that more than one person interacted with the thermostat. Some studies report that conflicts among the occupants regarding the thermostat are a possible cause of increased energy use [47–49].

3.4. *Is the clock setting correct?*

If the thermostat is set to the wrong time and date, the heating and cooling will not occur at the preferred times. This will certainly lead to occupant frustration and probably to higher than necessary energy use. We asked respondents to report if the time displayed on the thermostat accurately reflected the actual time. About 15% (±5%) indicated that the time was not correctly set, and half of those were one hour or more off.[6] We suspect that the percentage of clocks with the wrong time could be even higher, especially in zones that use daylight saving time.

3.5. *Who selected the thermostat?*

About 40% (±7%) of the respondents with programmable thermostats said the device was not selected by a member of the family. These numbers show the importance of other agents, such as contractors, builders, landlords and previous owners in choosing the technology used in people's homes. These agents, might, for instance, choose thermostats on the basis of contracts with manufacturers or to minimize cost, rather than selecting a device that is easy to use and matches the family needs.

3.6. *Household schedule/occupancy*

More than half of the respondents (57% ± 7%) indicated that somebody is at home nearly all of the time. This is not a unique characteristic of the respondents since other surveys found similar occupancy patterns: RECS reports 57% (±2%) of households with somebody at home all day [50]. The extended periods of occupancy also indicate a limit to the potential energy savings from regular temperature setbacks during the daytime and unoccupied periods.

3.7. *Average self-reported temperature*

We asked respondents to report the actual current indoor temperature indicated by their devices, so that we could compare this with the indoor temperature shown in the uploaded picture. However, we discovered some unexpected problems: some thermostats display only the actual indoor temperature, some display both actual and target temperature (either at the same time or alternating over time), others do not display the actual temperature all the time (for instance, when in hold). We suspect that this ambiguity parallels operational uncertainty by the users. In about half of the homes, the thermostat settings (revealed in the uploaded photos) matched the reported temperatures. In about a quarter of the homes the temperature settings displayed on the thermostats differed from reported settings and about one third of the respondents uploaded photos that were too blurry to determine the setting.

3.8. *Availability of the thermostat's operating manual*

The majority (70% ± 6%) of respondents reported that they did not know where the instruction manual of the thermostat was, although most of them also stated they did not need it to operate the thermostat. 67% (±8%) of those who reported to be unable to program their thermostats did not know the location of the manual.

[6] About 10% reported the thermostat did not display the time. Initially, we thought this was unlikely, but the analysis of the pictures revealed a model that used the space on the display screen where the time digits would normally be seen to display "hold". Further, in a mechanical setback thermostat, a cover often hides the clock. In some cases, we were not able to validate the self-reported time setting with the actual data because of blurry pictures, small fonts and the unavailability of the actual time the picture was taken.

3.9. Occupants' understanding of heating and cooling system operation

We included true/false questions (see Appendix) to learn if the respondents understood how their home thermostat and HVAC system operates. Lack of understanding of these operations can lead to improper use and wasted energy. For instance, one might think that keeping a constant (comfort) temperature in the house saves energy, but this is, in most of the cases, a misconception.

First, we asked the respondent if setting the thermostat at a higher temperature would heat the house faster. (The answer is no, for a traditional US HVAC system, with some exceptions that we could easily identify.) This question traces back to a 1986 study conducted by Kempton [51] uncovering folk theories about heating controls. People equate the operation of a thermostat with a valve (i.e., like a gas pedal, the more you press it, the gasoline – or in this case heat – flows) and not as a thermostatic device (i.e., a switch controlling a single rate of heat delivery until the desired temperature is achieved). In our survey, 36% (±7%) of the respondents subscribed to the (technically incorrect) "valve theory".

Second, the respondents were asked if they believed that the thermostat directly sets the temperature of the air coming out of the HVAC system. (The correct answer for most US systems is no.) In fact, 62% (±7%) of the respondents answered affirmatively, confirming the mistaken belief that the heating control is a valve or a gas pedal rather than a switch.

Finally, respondents were asked if lowering the thermostat setting at night or when people were not at home used more energy than keeping the temperature constant. (The correct answer is no in most cases, because of reduced heat losses when allowing the temperature to drift to low values, as shown in simulations[7] [5,6]). About one third of the respondents thought that turning down the thermostat at night or when people are not at home used more energy than keeping the house at the same temperature all the time. Thus, over one third of the respondents did not understand the relationship between inside temperatures and heat loss.

3.10. Methodology results

The responses to this survey paint a remarkable picture of a technology that is widely misunderstood by its users (and probably mis-designed by its manufacturers). We showed that crowdsourcing can be used to obtain both qualitative and quantitative data about use of thermostats in homes. How reliable is this data? Can this tool replace traditional surveys? As shown above, responses to survey questions in this study agree with large energy surveys. However, a direct validation is challenging, because of the small sample of this study, and because large surveys dedicate only a few questions to thermostats. More interestingly, by combining survey and images we found inconsistency in self-reported data, which cast doubts on the current approach of using surveys for energy efficiency research. A broader discussion of the merit of crowdsourcing in comparison with other methods is presented below.

4. Discussion

The traditional approach to collecting data about buildings, the occupants, and their operating habits relies on a survey (written or telephone) or a physical visit by a trained auditor. When used to retrieve temperature setpoints and other quantitative thermostat data these methods have several limitations. Most importantly, they frequently rely on self-reported data. Different studies found that self-reported thermostat settings are inaccurate. The comparison between pictures and reported data in our survey confirms these observations. People under-report their temperature settings [7,52] to conform to social norms or to please the researcher introducing social desirability bias. Finally, settings can be hard to remember (and thus hard to accurately report), especially if the thermostat is adjusted rarely (recollection bias) [53]. Users also may find it difficult to retrieve the program inputs in order to answer a survey [4,45]. Even when reported correctly, setpoints can change seasonally or over time. Large surveys are expensive, thus they are conducted rarely (3–10 years) and fail to capture these changes.

Policymakers, program designers, and modelers would like to know temperature setpoints that are representative of the whole population, and capture periodic, exceptional and annual changes. One promising new approach relies on data from Internet-connected thermostats. Internet-connected thermostats are offered by various providers (the Nest [54] is probably the best known, but they are also offered by utilities, cable TV providers, and burglar alarm companies). The network connection allows users to control settings on a website or with a smart phone. We estimate that already more than four million homes in North America are equipped with Internet-connected thermostats and their usage is expected to grow rapidly. Similar services have recently appeared in other countries, such as France, UK and Germany [55–58]. The thermostat providers typically log thermostat settings, indoor temperatures and furnace (and air conditioner) cycle times. Some units also capture humidity, outside temperature, and other parameters. Thus, the Internet-connected thermostat offers a detailed view of heating and cooling operations in a large number of homes. These logs represent the only real "ground truth" for thermostat and temperature data. However, despite the ease of data collection, connected homes[8] are not (and are not going to be for many years) representative of the national building stock or its occupants (because of the service's opt-in feature). Further, the vendors of Internet-connected thermostats have been reluctant to share this information with researchers and policymakers. Finally, as anticipated above, setpoints are essential, but not sufficient to estimate savings in homes. Additional information, such as thermal comfort, occupancy patterns as well as motivation and type of interaction with the device need to be understood. Connected thermostats do not provide that information.

The richest techniques to understand these qualitative aspects are interviews and direct observation of the household. Kempton [51] pioneered this type of research almost 30 years ago. Recent work has explored in detail how people interact with programmable (and advanced) thermostats [31,45,60]. These techniques are expensive and can only reach a very small fraction of the entire population (due to self-selection and geographical location of the researchers). We propose, as an alternative way of obtaining both setpoints and other qualitative information, the use of online crowdsourcing. Our research method shares the same advantages of online surveys, being inexpensive, rapid, and able to collect data from a geographically diverse population. Furthermore, online crowdsourcing does not limit the kind of information that could be collected compared to traditional surveys based on telephone or mail. We collected pictures uploaded by Turkers, to crosscheck survey results with the real settings of the thermostat. In another

[7] The result depends on thermal mass and insulation of the house, and capacity of the HVAC system, but in most of the cases the setback strategy will save energy as documented in the literature. The ability to automatically setback the temperature at certain times is the main reason why properly programmed thermostats save energy over manual thermostats kept at constant temperature.

[8] Internet-connected thermostats can be considered nodes of the new smart-grid, the new electric network of the 21st century that uses information technology to deliver electricity efficiently, reliably, and securely. [59] NIST. (2015). *Smart Grid: A Beginner's Guide*. Available: http://www.nist.gov/smartgrid/beginnersguide.cfm.

investigation using Amazon Mechanical Turk, Ford and Karlin [40] tested eco-feedback interfaces and different ways of displaying feedback information. It is now relatively easy to ask Turkers to upload videos or pictures of daily routines that have impact on energy use. Crowdsourcing can also engage the respondents in more dynamic interactions with the researcher. For example, Amazon Mechanical Turk has the unique feature of being able to identify subjects (using a numeric ID). This identification feature allows for follow-up questions but keeps the respondent's identity anonymous. Thus, researchers can ask open-ended questions and use the survey as a form of remote simplified interview with the subject. For instance, the researcher could obtain pictures of seasonal settings from the same Turkers, thus creating a longitudinal study.

Rapid turnaround allows conducting pilot studies and clarifying ambiguous questions, as well as fine-tuning the survey design. Researchers can quickly adapt the survey to new findings and post it again. They can request the previous subjects to complete the new survey or exclude them from the pool. This feature is particularly valuable when investigating complex problems such as the relationship between behavior and energy use. We took advantage of this feature, repeating the survey three times, before the final run.

The crowdsourced respondents are not representative of the national population. Crowdsourcing platforms tend to underrepresent elderly, low-income and low-education people, introducing sampling bias [33,61]. Other studies showed that cell weighting can be used to calibrate and adjust the results [39], but some segments of the population will most likely never be adequately captured through crowdsourcing. In our survey we did not collect the demographics of the sample, but our results match reasonably well data collected in other studies even without weighting.

Crowdsourcing platforms simplify recruitment, solving a significant problem with other online survey tools. On the other hand, relying on a pool of people paid for their work raises problem of reliability, truthfulness and intrinsic motivation. Paolacci [62] showed that Turkers are both intrinsically and extrinsically motivated and reliable, and Rand [63] and Shapiro et al. [64] consider them honest, consistent and conscientious. The reputation system, which gives the option to select only Turkers with high score (based on their previous work), is another significant advantage compared to traditional online surveys.

Finally, the Mechanical Turk offers researchers toolsets to easily manage the submitted data and perform simple evaluations. It is now possible for the workers to directly upload photos or videos (rather than uploading to a separate site as was done in this paper). This provides a useful tool for tracking how users adjust the settings on a thermostat or, for example, collecting images of building features.

Online crowdsourcing still shares many of the drawbacks of other surveys based on self-reporting. Reliability and truthfulness of the survey pool cannot be tested, especially when some tasks involve off-line (non-verifiable) components. The respondents can still be confused with the terminology or meaning of the questions. Personal interviews would still be more helpful for these kinds of exploratory studies.

5. Conclusions and future work

The Amazon Mechanical Turk is a valuable crowdsourcing tool for collecting information that would be difficult or expensive to obtain through other means. Online crowdsourcing platforms have several advantages over traditional (postal or telephone) surveys. For example, a survey can be easily repeated in winter and summer and thus capture seasonal variation, and the pairing of a survey with a photograph of the thermostat offers an unprecedented mean of verification. Easier recruitment process and reputation system make this tool more effective than alternative online tools. Limitations include non-representativeness and concerns about reliability of the population. This study was limited to the United States, but it could be easily duplicated in other countries offering similar online crowdsourcing services. Knowledge and understanding of users' behavior in operating thermostats is highly valuable to policymakers, energy efficiency researchers, manufacturers, as well as behavior change practitioners in utilities and non-governmental organizations worldwide.

The results of our study show, from many perspectives, that people have a poor understanding of programmable thermostats and an imperfect grasp of how heating systems operate. We identified significant failures in the people–technology interactions, probably resulting in wasted energy use. About 40% of programmable thermostat owners did not use the energy-saving programming features. In about one third of the homes, the occupants disabled or overrode the programming features. Responses suggest thermostats have confusing user interfaces that leads to operation different from what manufacturers expected. The installation of a new thermostat affects a wide range of practices, many of which were not anticipated by the thermostat designers or energy researchers. The study also suggests weaknesses of self-reported data when collected exclusively in the form of a survey, which raises concerns about the validity of current thermostat-related research using such data. "Ground truth" temperature data is now possible with Internet-connected thermostats, but these still represent a very small portion of the home control systems.

Future work should aim at extending this experiment to a larger sample (a shortcoming of the present study is the small sample size, with 192 survey respondents and 30 pictures), testing the same method in other countries, and conducting longitudinal studies collecting diary logs, pictures and possibly videos.

Acknowledgments

This work was supported by the Office of Energy Efficiency and Renewable Energy, Building Technologies Program of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. The authors wish to thank Gari Kloss for her tremendous work in the management and coordination of this study, and Dr. Eleonora Grandi for helpful discussions, review and help drafting the manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.erss.2015.06.002

References

[1] Shove E, Pantzar M, Watson M. The dynamics of social practice: everyday life and how it changes. Sage; 2012.
[2] Stern PC. Individual and household interactions with energy systems: toward integrated understanding. Energy Res Soc Sci 2014;1:41–8.
[3] Sovacool BK. What are we doing here? Analyzing fifteen years of energy scholarship and proposing a social science research agenda. Energy Res Soc Sci 2014;1:1–29.
[4] Peffer T, Pritoni M, Meier A, Aragon C, Perry D. How people use thermostats in homes: a review. Build Environ 2011;46(12):2529–41, http://dx.doi.org/10.1016/j.buildenv.2011.06.002. ISSN 0360-1323.
[5] Plourde A. Programmable thermostats as means of generating energy savings: some pros and cons. In: Canadian Building Energy End-Use Data and Analysis Centre, Technical Report CBEEDAC 2003-RP-01; 2003.
[6] Haiad C, Peterson J, Reeves P, Hirsch J. Programmable thermostats installed into residential buildings: predicting energy savings using occupant behavior & simulation. Southern California Edison; 2004.

[7] Nevius M, Pigg S. Programmable thermostats that go berserk: taking a social perspective on space heating in Wisconsin. In: Proceedings of the 2000 ACEEE Summer Study on Energy Efficiency in Buildings, vol. 8. 2000. p. 233–44.
[8] Cross D, Judd D. Automatic setback thermostats: measure persistence and customer behavior, Chicago; 1997.
[9] Shipworth M, Firth SK, Gentry MI, Wright AJ, Shipworth DT, Lomas KJ. Central heating thermostat settings and timing: building demographics. Build Res Inf 2010;38:50–69.
[10] Sachs H. Programmable thermostats. Washington, DC: American Council for an Energy Efficient Economy; 2004 https://www.energystar.gov/ia/partners/prod_development/revisions/downloads/thermostats/ACEEE.pdf?0544-2a1e
[11] Bouchelle MP, Parker DS, Anello MT. Factors influencing space heat and heat pump efficiency from a large-scale residential monitoring study. In: Proceedings of the 2000 ACEEE Summer Study on Buildings and Energy. 2000.
[12] Analytics RLW. Validating the impact of programmable thermostats. Middletown, CT: GasNetworks; 2007 http://www.efi.org/docs/cee_thermostats.pdf
[13] Michaud N, Megdal L, Baillargeon P, Acocella C. Billing analysis & environment that "Re-Sets" savings for programmable thermostats in new homes. In: Presented at the IEPEC. 2009.
[14] Tachibana D. Residential customer characteristics survey 2009; 2010 http://www.seattle.gov/light/Conserve/Reports/Evaluation_15.pdf
[15] Matsunami H. Overview of the Research Institute of Behavior Observation. In: Presented at the Osaka Gas Meeting. 2012.
[16] Karjalainen S, Koistinen O. User problems with individual temperature control in offices. Build Environ 2007;42:2880–7.
[17] Rathouse K, Young B. Market transformation programme – domestic heating: use of controls; 2004. Available at: http://efficient-products.defra.gov.uk/ReferenceLibrary/Domestic_Heating_Controls_RPDH15.pdf
[18] Rathouse K, Young B. RPDH15: use of domestic heating controls. Watford, UK: Building Research Establishment; 2004.
[19] Meier A, Aragon C, Peffer T, Perry D, Pritoni M. Usability of residential thermostats: preliminary investigations. Build Environ 2011;46.
[20] Energy Information Administration (EIA). Residential energy consumption survey: preliminary housing characteristics tables. Energy Information Administration; 2005.
[21] Energy Information Administration (EIA). Residential energy consumption survey, housing characteristics tables, table HC15.7: air-conditioning usage; 2005.
[22] Energy Information Administration (EIA). Residential energy consumption survey, housing characteristics tables, table HC14.5: space heating usage; 2005.
[23] Decision Analyst. 2008 American Home Comfort Survey. Arlington: Decision Analyst; 2008.
[24] California Energy Commission (CEC). California Statewide Residential Appliance Saturation Study. Sacramento: California Energy Commission; 2004, 300-00-004.
[25] California Energy Commission (CEC). 2009 California Statewide Residential Appliance Saturation Study. California Energy Commission; 2010.
[26] Mavrogianni A, Johnson F, Ucci M, Marmot A, Wardle J, Oreszczyn T, et al. Historic variations in winter indoor domestic temperatures and potential implications for body weight gain. Indoor Built Environ 2013;22(2):360–75, http://dx.doi.org/10.1177/1420326X11425966.
[27] Eurostat EC. Manual for statistics on energy consumption in households; 2013.
[28] Kempton W, Krabacher S. Thermostat management: intensive interviewing used to interpret instrumentation data. In: Kempton W, Neiman M, editors. Energy efficiency: perspectives on individual behavior. Berkeley: American Council for an Energy Efficient Economy; 1987. p. 245–62.
[29] Vadodaria K, Loveday DL, Haines V. Measured winter and spring-time indoor temperatures in UK homes over the period 1969–2010: a review and synthesis. Energy Policy 2014:252–62.
[30] Roberts D, Kerylyn L. Variability in measured space temperatures in 60 homes. National Renewable Energy Laboratory (NREL); 2013.
[31] Walker SL, Lowery D, Theobald K. Low-carbon retrofits in social housing: interaction with occupant behaviour. Energy Res Soc Sci 2014;2:102–14.
[32] Paolacci G, Chandler J, Ipeirotis P. Running experiments on Amazon Mechanical Turk. Judgm Decis Mak 2010;5.
[33] Mason W, Suri S. Conducting behavioral research on Amazon's Mechanical Turk; 2010. Available at: http://ssrn.com/abstract=1691163
[34] Horton J, Rand D, Zeckhauser R. The online laboratory: conducting experiments in a real labor market. NBER Working Paper, vol. w15691; 2010.
[35] Paolacci G, Warglien M. http://experimentalturk.wordpress.com/about/. 2009.
[36] Bazilian M, Rice A, Rotich J, Howells M, DeCarolis J, Macmillan S. Open source software and crowdsourcing for energy analysis. Energy Policy 2012:149–53.
[37] Yang H-C, Donovan S, Young S, Greenblatt J, Desroches L-B. Assessment of household appliance surveys collected with Amazon Mechanical Turk. Energy Effic 2015:1–13, 2015/03/04.
[38] Greenblatt JB. U.S. residential miscellaneous refrigeration products: results from Amazon Mechanical Turk Surveys. LBNL; 2013.
[39] Kantner CLS. Ceiling fan and ceiling fan light kit use in the U.S. results of a survey on Amazon Mechanical Turk. Lawrence Berkeley National Laboratory; 2013.
[40] Ford R, Karlin B. Graphical displays in eco-feedback: a cognitive approach. In: Proceedings of the HCII 2013 Conference. 2013.
[41] Peffer TE. California DREAMing: the design of residential demand responsive technology with people in mind. Berkeley, CA: University of California Berkeley; 2009.
[42] Meier AK, Aragon C, Peffer TE, Pritoni M. Thermostat interface and usability: a survey, report no. 4182E. Berkeley: Lawrence Berkeley National Laboratory; 2010.
[43] Archacki R. Carrier thermostat mode summary: summer 2003; 2003 (Personal correspondence to Gaymond Yee).
[44] Lopes JS, Agnew P. FPL residential thermostat load control pilot project evaluation. In: Proceedings of ACEEE Summer Study on Energy Efficiency in Buildings, vol. 2. 2010. p. 184–92.
[45] Peffer T, Perry D, Pritoni M, Aragon C, Meier A. Facilitating energy savings with programmable thermostats: evaluation and guidelines for the thermostat user interface. Ergonomics 2013;56(3):463–79, http://dx.doi.org/10.1080/00140139.2012.718370.
[46] Outcault S, Barriga C, Heinemeier K, Markley J, Berman D. Thermostats can't fix this: case studies on advanced thermostat field tests. In: ACEEE summer study 2014. 2014.
[47] Parker D, Barkaszi S, Sherwin J, Richardson C. Central air conditioner usage patterns in low-income housing in a hot and humid climate: influences on energy use and peak demand; 1996.
[48] McCalley L, Midden C. Goal conflict and user experience: moderators to the use of the clock thermostat as a device to support conservation behavior. In: Proceedings from the 2004 ACEEE Summer Study on Buildings and Energy, vol. 7. 2004. p. 251–9.
[49] Karjalainen S. Gender differences in thermal comfort and use of thermostats in everyday thermal environments. Build Environ 2007;42:1594–603.
[50] Energy Information Administration (EIA). Residential energy consumption survey, housing demographics; 2009. Available at: http://www.eia.gov/consumption/residential/data/2009/-undefined
[51] Kempton W. Two theories of home heat control. Cogn Sci 1986;10:75–90.
[52] Vine E, Barnes BK. Monitored indoor temperatures and reported thermostat settings: how different are they? Energy 1989;14:299–308.
[53] Lutz J, Wilcox BA. Comparison of self reported and measured thermostat behavior in new California houses. In: Proceedings of the 1990 ACEEE Summer Study on Energy Efficiency in Buildings, vol. 2. 1990. p. 91–100.
[54] Lab N. Nest learning thermostat website; 2015. Available at: https://nest.com/
[55] TADO. TADO Cooling Intelligente Klimasteuerung; 2014. Available at: http://www.tado.com/de/
[56] Netatmo. Netatmo website; 2014. Available at: http://www.netatmo.com/
[57] QIVIVO. QIVIVO website; 2014. Available at: http://www.qivivo.com/
[58] Hive Active Heating. Hive Active Heating website; 2014. Available at: https://http://www.hivehome.com/
[59] NIST. Smart grid: a beginner's guide; 2015. Available at: http://www.nist.gov/smartgrid/beginnersguide.cfm
[60] Yang R, Newman M. Learning from a learning thermostat: lessons for intelligent systems for the home. In: Presented at the UbiComp '13. 2013.
[61] Ipeirotis PG. Demographics of Mechanical Turk; 2010.
[62] Paolacci G, Chandler J. Inside the Turk: understanding Mechanical Turk as a participant pool. Curr Dir Psychol Sci 2014.
[63] Rand DG. The promise of Mechanical Turk: how online labor markets can help theorists run behavioral experiments. J Theor Biol 2012:172–9.
[64] Shapiro DN, Chandler J, Mueller PA. Using Mechanical Turk to study clinical populations. Clin Psychol Sci 2013;1:213–20.