Title

When Smart Thermostats Are Dumb: Lessons Learned from Evaluating Eight Advanced Thermostats

Authors

Peffer, Therese E.
Pantelic, Jovan
He, Yingdong
et al.

Publication Date

2022-08-01

Peer reviewed

When smart thermostats are dumb: lessons learned from evaluating advanced thermostats

Therese Peffer, California Institute for Energy and Environment, UC Berkeley
Jovan Pantelic, Delos
Yingdong He, Center for the Built Environment, UC Berkeley
Fred Bauman, Center for the Built Environment, UC Berkeley

ABSTRACT

A decade after a certain round thermostat upended the industry, we now find dozens of so-called smart or advanced thermostats, with remote control, occupancy sensing, and multiple optimization strategies. Many studies evaluate these features in the residential sector. But how does one choose the right advanced thermostat for the application for a small commercial setting?

This paper describes the evaluation and testing of several advanced thermostats. The research team reviewed the literature, interviewed 20 facilities managers, conducted sensory accuracy tests, and performed a heuristic evaluation of several advanced thermostats currently on the market.

What did we learn by studying these thermostats? We learned which features are important to facilities managers and what support they need to improve their performance. We discovered that thermostats differed in sensor accuracy: lighted colored screens may cause self-heating, which can throw off accuracy by as much as 4F; (even a degree or two Fahrenheit can make a difference in saving energy). We studied thermal asymmetry: the sensitivity of the thermostat temperature sensor to a source of radiant heat, such as a large window, sunlight on a wall or any warm surface in the room, such as large equipment. Finally, the heuristic evaluation revealed some of these advanced thermostats fail basic usability tests.

Introduction

In the early 1980s, the Environmental Protection Agency (EPA) promoted thermostats with a clock or schedule function—“programmable” thermostats—as a means of saving energy. One can adjust the desired or target temperature, such as reduce the heating setpoint or increase the cooling setpoint, to reduce heating or cooling cycles. Larger adjustments or turning equipment off are recommended for periods when the space is unoccupied or in homes at night when occupants are asleep. Simulation studies indicate that 1 degree Fahrenheit would save 3% energy (O’Leary 2012); setbacks of 7-10ºF for 8 hours can save 10% over a year (US Department of Energy (DOE) 2021). However, early programmable thermostats were not easy to use (Peffer et al. 2011) nor did they necessarily save energy. The EPA discontinued their EnergyStar program for programmable thermostats in 2009 due to lack of savings and poor usability (Kaplan 2009). A few years later they developed the Climate Controls specification and Connected Thermostat specification. The current (2021) EnergyStar certification uses the term “Smart Thermostats” that are “independently certified to save energy” (US Department of Energy (DOE) 2022). Two other organizations have developed specifications: the Consortium for Energy Efficiency (CEE) for connected thermostats (Consortium for Energy Efficiency

(CEE) 2021) and the California Energy Commission for Occupant Controlled Smart Thermostats (California Energy Commission (CEC) 2019). These thermostats provide both energy and non-energy benefits, including improved comfort, added convenience, and lower costs due to reduced energy consumption.

The last decade has witnessed several advancements that have improved thermostat functionality and usability. Smartphones have become ubiquitous since Apple's iPhone debuted in 2007, encouraging full internet and computing capability in other embedded devices and upping the game for color touchscreen user interfaces. More and more devices are network connected, which enables remote communication and response to utility grid signals; voice recognition has enabled Alexa (introduced in 2014) and other virtual intelligent assistance as a hub for smart management. Algorithms go beyond simple timed autorecovery[1] to sophisticated machine learning and Artificial Intelligence.

An advanced thermostat is simultaneously (1) a networked device (e.g., Internet of Things), (2) an embedded device (e.g., microcomputer capable of logging data, performing analytics, and conducting machine learning), and (3) a controller for Heating, Ventilation and Air Conditioning (HVAC). While these products are available off-the-shelf at hardware and electronic stores, these products tend to be installed as part of custom-engineered/built-up systems. While most advanced thermostats are compatible with typical small commercial HVAC equipment[2] wiring to the thermostat, ensuring that these devices communicate with all essential HVAC equipment is still an ongoing problem.

Advanced thermostats have the potential to significantly lower energy consumption by reducing the runtime of heating, cooling, and fan equipment (20-30% based on several field and simulation studies) while maintaining safe and comfortable indoor environments. These "smart control devices" have the ability to adjust zone temperature setpoints to allow the most energy efficient operation based on a variety of inputs, including occupancy, a programmed or learned schedule, or reliability signals from the utility grid, as well as other advanced learning algorithms such as occupant temperature preferences. Advanced technologies include:

- Sensing: temperature sensors outside the thermostat, occupancy sensors, carbon dioxide;
- Advanced algorithms: learning occupant schedules and preferences, learning internal space load profiles, controlling temperature based on price signals, outside temperature, or number of occupants, or optimized for comfort and cost, fault detection and diagnostics;
- Engaging user interfaces: simple to use, understand, and remember ("walk-up usable");
- Networked-systems: to the Internet, to other thermostats (e.g., others in the same zone or other zones in the building), or to other systems, such as voice-activated gateways, sensors, ceiling fans, or Building Management Systems (BMS). Networked systems allow a single manager to access and change settings in multiple zones or buildings.
- Load flexibility or improved grid reliability by responding to reliability or price signals.

Differentiating features between smart thermostat products tend to be user interfaces, optimization and learning algorithms, and sensing. Small commercial buildings use thermostats, but have different HVAC requirements and occupancy patterns, both of which affect energy consumption. While the intelligent adjustment of zone temperature setpoints (as performed by an

[1] Autorecovery is a common feature of thermostat that turns on the Heating or Cooling system a few minutes before an expected occupied mode, thus "recovering" from an energy-saving mode.
[2] The list includes Roof Top Units (RTUs), conventional compressor-based Air Conditioning, and Forced Air Units with low voltage (24V AC).

advanced thermostat) has been shown to reduce resource consumption in all climates, simulation studies indicate that the amount of savings varies by climate (Hoyt, Arens, and Zhang 2015). Advanced thermostats have their largest impact on resource consumption in buildings in which occupants and their interaction with zone thermostats play an important role in the operation of the building's HVAC system. Reducing or eliminating HVAC runtime during unoccupied (e.g., closed business) or activities that require less conditioning (e.g., sleep) can save energy. In addition, the setpoints (e.g., temperature below which heating occurs or above which cooling is triggered) affect energy consumption and thermal comfort. Moreover, the occupants' understanding and ease of use of the thermostat affects its operation in automatic and manual modes. For these reasons, occupancy patterns and occupant interaction with the thermostat are also important factors in resource consumption.

This study is part of the California Energy Commission-funded California Energy Product Evaluation (Cal-EPE) Hub project to evaluate market-ready technologies at the early adoption phase and disseminate the findings through a public website. The intention is to provide information and advice regarding features, functions and performance for people who procure thermostats for commercial and industrial settings, such as for small commercial building chains (e.g., banks, retail, restaurants), schools, hotel/motels, or college dormitories. This study does not address the average homeowner seeking a thermostat for their home.

This paper outlines the methods used in evaluating the thermostats, the results, and a brief discussion.

Methods

This section describes a brief literature review and thermostat selection, interviews with facilities managers, evaluation of sensors, and heuristic evaluation.

Literature review

While many studies address residential thermostats, we found fewer studies that include thermostats in commercial settings (Johnson, Peffer, and Woolley 2012; Peffer et al. 2019; Outcault et al. 2014). One study found that hotels used thermostats with occupancy sensors (sometimes requiring the keycard to operate the thermostat) to reduce HVAC energy consumption when unoccupied (Inncom 2010). A study in New York installed advanced thermostats in dormitories; the lock out function, which limits the range of allowable temperature setpoints, reduced energy consumption (previously, some students were enjoying tropical temperatures on cold winter days) (Telkonet 2011). Another study installed networked thermostats in dormitory rooms; while the thermostats failed to save energy due to the configuration with the BMS, the facilities managers appreciated the time savings in the convenience of the remote access in changing setpoints, especially for holiday breaks (Woolley et al. 2014). Another project installed advanced thermostats in 18 small commercial buildings throughout California; the schedule alone saved energy and the network feature allowed reducing peak load during critical peak events (Peffer et al. 2019). These findings informed the identification of features and functions and their use cases in the commercial setting.

Advanced thermostat functions – evaluation considerations

For the evaluation of advanced thermostats used in the commercial sector, we anticipated different needs than the residential sector. For example, commercial buildings more typically have multiple thermostats controlling the same or different zones, thus grouping or other coordination is a key function. The thermostat remote interface via a web browser or smartphone is more likely to be used by a facilities manager and requires more functions than the physical thermostat's interface installed on a wall; the wall interface is more likely to have a lock-out feature that prevents much occupant interaction. Usability is a key factor in how facilities managers use the advanced features. Fault detection and diagnostics may be more useful in commercial settings than in residential. Finally, reducing demand charges impacts commercial buildings more so than residential. The following is a list of advanced thermostat functions:

- Response to utility signals (e.g., PG&E's Peak Day Pricing or SCE's Critical Peak Pricing) by changing target temperatures to reduce energy
- Hour-by-hour schedule customization
- Occupancy sensing through sensors or geofencing through mobile phones to turn down or off heating and cooling when no one is present
- Reporting (feedback) and notification of problems (e.g., zone not meeting target temperature)
- Energy reporting
- Remote control and programming (e.g., through voice-control, smart mobile device or web browser).
- Networked to other devices (e.g., security system, voice/speaker gateway, ceiling fans)
- Improved usability (colors, touchscreens, icons, improved flow of installation, setup and programming through wizards, and error recovery).
- Advanced algorithms such as optimization (e.g., temperature recovery, adjusting temperature setpoints, HVAC) or learning of desired temperatures or schedule (occupancy, behavior, programming).
- Automatic software updates
- Data collection that can create opportunities for targeted efforts for utilities like energy assessments, HVAC upgrades, or tune-up programs
- Minimize electric resistance heat runtimes for heat pumps with electric resistance backup systems

Selecting thermostats

In our determination of which thermostats to study, we selected thermostats that were popular or from well-established companies as well as some new technologies. The final selection was not intended to be exhaustive, but representative of the market. We spoke with facilities managers and our Technical Advisory Committee (TAC) and conducted web searches for market share of devices. The final selection of thermostats is listed in Table 1.

Table 1: Selected thermostats for evaluation

Manufacturer	Product Name	Model Number
Resideo Technologies (under license from Honeywell Home)	T9 Smart Thermostat	RCHT9 10WF
TRANE	PivotTM Smart Thermostat	XL824
Emerson	Sensi Touch Smart Thermostat	ST75
Johnson Controls Inc	GLAS Thermostat	SiO2-10000

Carrier	ConnectTM Wi-Fi Thermostat	33CONNECTSTAT
Google	Nest Learning Thermostat 3rd Gen	T3007ES
Pelican	TS250H	
Ecobee	Smart Thermostat	EB-STATES-01
Network Thermostat	RP32HE-Wi-Fi	X5H-WIFI-B

The most effective evaluation method would have been to install and test different advanced thermostat products in an operational building, but these tests are expensive and time-consuming. We opted instead to evaluate the thermostats using three methods: interviewing facilities managers and energy service providers (referred to as energy managers) familiar with one or more advanced thermostat, evaluating the temperature sensors on the advanced thermostat in a laboratory setting, and conducting a heuristic evaluation of the thermostat user interfaces.

Interviewing energy managers

In order to understand how well these advanced thermostats met the needs of energy managers, we developed an interview protocol of over ten questions. The key questions were: 1) familiarity with each thermostat; 2) the most common means of accessing the thermostats (web browser or smartphone); 3) functions most often used; 4) which features enabled energy savings; 5) whether energy reports generated from the thermostats were read by the manager; 6) which notifications were most useful; 7) which features they wished the thermostat had; and 8) whether cost was an important criterion. We worked with our networks and the TAC to find 20 people to interview by phone and video conference.

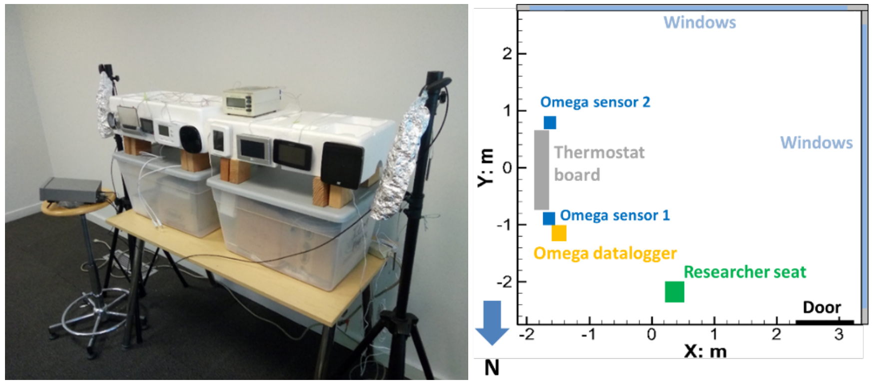
Testing sensors

We measured the standby power consumption, and, since even a degree temperature difference can save energy, we conducted three tests of the temperature sensor for each thermostat. One test measured thermal symmetry: comparing the thermostat temperature sensor measurement compared to a laboratory measurement in a controlled laboratory setting at both 68 F (20C) and 75F (24 C). The second test assessed thermal asymmetry: the accuracy of the thermostat temperature measurement when exposed to a simulated heated window surface for different surface temperatures and locations. The third test evaluated the accuracy of temperature sensors over time.

All tests were conducted in the Center for the Built Environment (CBE) comfort chamber at UC Berkeley in California. The chamber layout is shown in Figure 1:

- The thermostats were fixed on foam boards which were placed near the center of the east wall of the CBE chamber, four feet above the floor.
- Two high accuracy Omega sensors, shielded by aluminum foil to prevent effects from radiant surfaces, were used to measure the air temperature near the thermostats.
- For the thermal asymmetry testing, four heated panels, which formed a warm surface (width 58 inches, height 22.5 inches) simulating a warm window, were placed in front of the thermostats, 39 inches above the floor.

Figure 1. Chamber layout for thermal **symmetry** testing. Left: Thermostats attached to foam board; Right: schematic of floor plan showing locations of devices.



For the thermal symmetry testing, the test procedure was as follows:

- Thermostats were powered-off for more than 30 minutes to reduce the potential effects of the heat generated by thermostat elements before the test;
- Thermostats were turned and kept on for 20 minutes (for startup and stabilization);
- Thermostats were kept on for another 60 minutes and the display temperatures were recorded every 5 minutes.

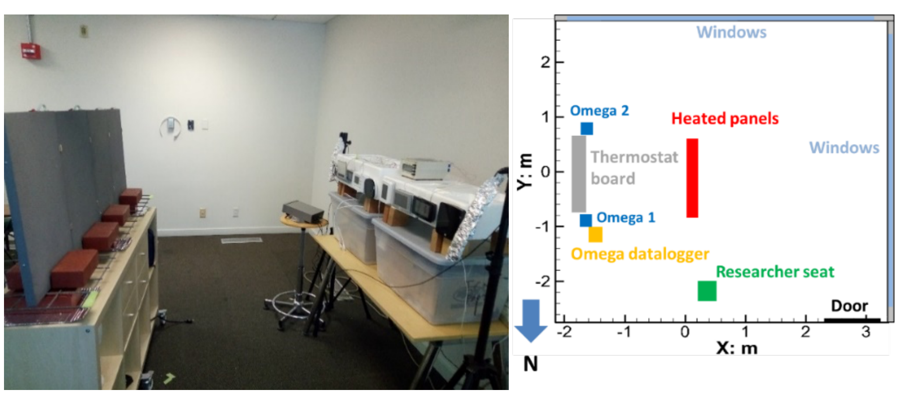
We repeated the thermal symmetry testing with thermostat calibration.

For the thermal asymmetry testing (Figure 2), the test procedure was as follows:

- Thermostats were powered-off for more than 30 minutes;
- Thermostats were turned and kept on for 10 minutes (for startup and stabilization);
- Thermostats were kept on and exposed to the warm surface for another 10 minutes; and
- Thermostats were kept on for another 60 minutes and the display temperatures were recorded every 5 minutes.

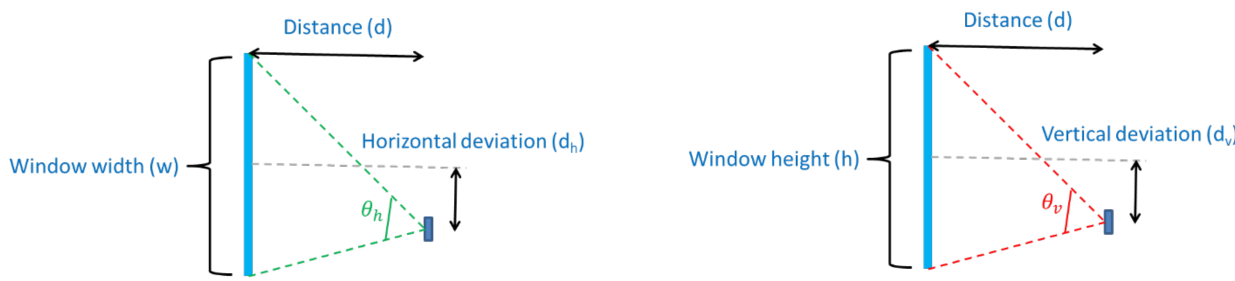
In the thermal asymmetry testing, the thermostat display showed the temperature deviation caused by warm surfaces (representing windows) at 86F (30C), 97F (36C), and 107.6F (42C).

Figure 2. Chamber layout for thermal **asymmetry** testing. Left: Radiant heated panel across from thermostats; Right: schematic of floor plan showing locations of devices.



The simplified view factor is used to qualify the positional relationship between the simulated warm window and thermostats, and it originates from the method for calculating the sky view factor of an indoor human body that was proposed in the study (He et al. 2021). The simplified view factor (f_{view}) is calculated according to the window height (h), window width (w), vertical deviation (d_v), horizontal deviation (d_h), and the distance between the thermostat and the window (d). Since the simplified view factor is not related to the shape, size, or materials of the thermostat, the calculation method can be applied for different thermostats.

Figure 3. The calculation of thermostat view factor. Left: plan of view factor seen by object; Right: elevation of view factor seen by object.



Conducting a heuristic evaluation

Several studies—including by the US EPA—have shown that usability of thermostats affects energy savings. While we planned for usability testing and a heuristic evaluation—our own assessment of usability using recognized usability guidelines (or heuristics)— we were unable to access actual users due to COVID.

Instead, we relied on our previous work: a heuristic evaluation conducted after a usability test was able to match violations of usability guidelines with poor usability and support for these heuristics with better usability (Peffer et al. 2013). A usability study of five thermostats measured performance on several tasks ranging from simple (turning on heat, set time of day, identify current temperature) to more complex (set target temperature, see future setting, set thermostat for away period). The researchers developed a series of metrics that correlated with usability (Perry et al. 2011):

- **Success**: Whether or not task was successfully completed.
- **Path length** participant took to perform given task. This is compared to the ideal path length (e.g., the minimum number of steps (press button, scroll through menu, click up arrow) to accomplish task. If changing the target temperature or setpoint can be accomplished with three clicks, that represents the ideal path length; the number of clicks that a participant took will be compared to that number.
- **Time** necessary to complete a task
- **Self-evaluation**: Survey where perceived difficulty of the task and user self-confidence in achieving success was evaluated

The heuristic evaluation surveyed user interface, ergonomics, web design, controls, and other literature to develop a set of heuristics. Then the heuristics were applied to each thermostat and usability performance. The most applicable heuristics are listed in Table 2.

Table 2: Useful heuristics for thermostat usability (Peffer et al. 2013)

Heuristic	Source
Visibility of available options on the home screen: Include all important and often used actions at the home level; consider no covers or clearly provide affordances	(Polson and Lewis 1990; Norman 2002)
A wide and shallow decision tree	(Shneiderman 1988)
Navigation cues	(Peffer et al. 2013)
Clear hierarchy of display	(Peffer et al. 2013)
Consistency and standards: Use a graphic tabular form to view the temperature setpoints for the time of day and day of week, Use plain English wherever possible (no abbreviations) and standard icons	(J Nielsen 1994; Norman 2002; Shneiderman and Plaisant 2009)
Natural mappings: Use clear affordances	(Norman 2002)
Error prevention and recovery: When possible, include confirmation prompts (e.g., do you want to save?), or some other means of confirming when something is edited or changed	(Polson and Lewis 1990; Jakob Nielsen 1993; Norman 2002; Shneiderman and Plaisant 2009)
Feedback from controls	(Karjalainen 2008; Bordass, Leaman, and Bunn 2007; Shneiderman and Plaisant 2009)

While a heuristic evaluation is not a substitute for usability testing with a variety of end users, the evaluation provides some insight. We outlined several use cases applicable to the commercial sector (e.g., grouped thermostats), developed tasks (Table 3), and then evaluated these tasks for each thermostat using the heuristics. The tasks included:

Table 3: List of tasks for evaluation

Task	Task description
1	Turn on/off heating
2	Set time and day
3	Identify current temperature and Change temperature
4	Identify setting for future day/time
5	Set thermostat for away period
6	Advanced functionality: grouping thermostat, turn on/off Machine Learning, Demand Response functionality

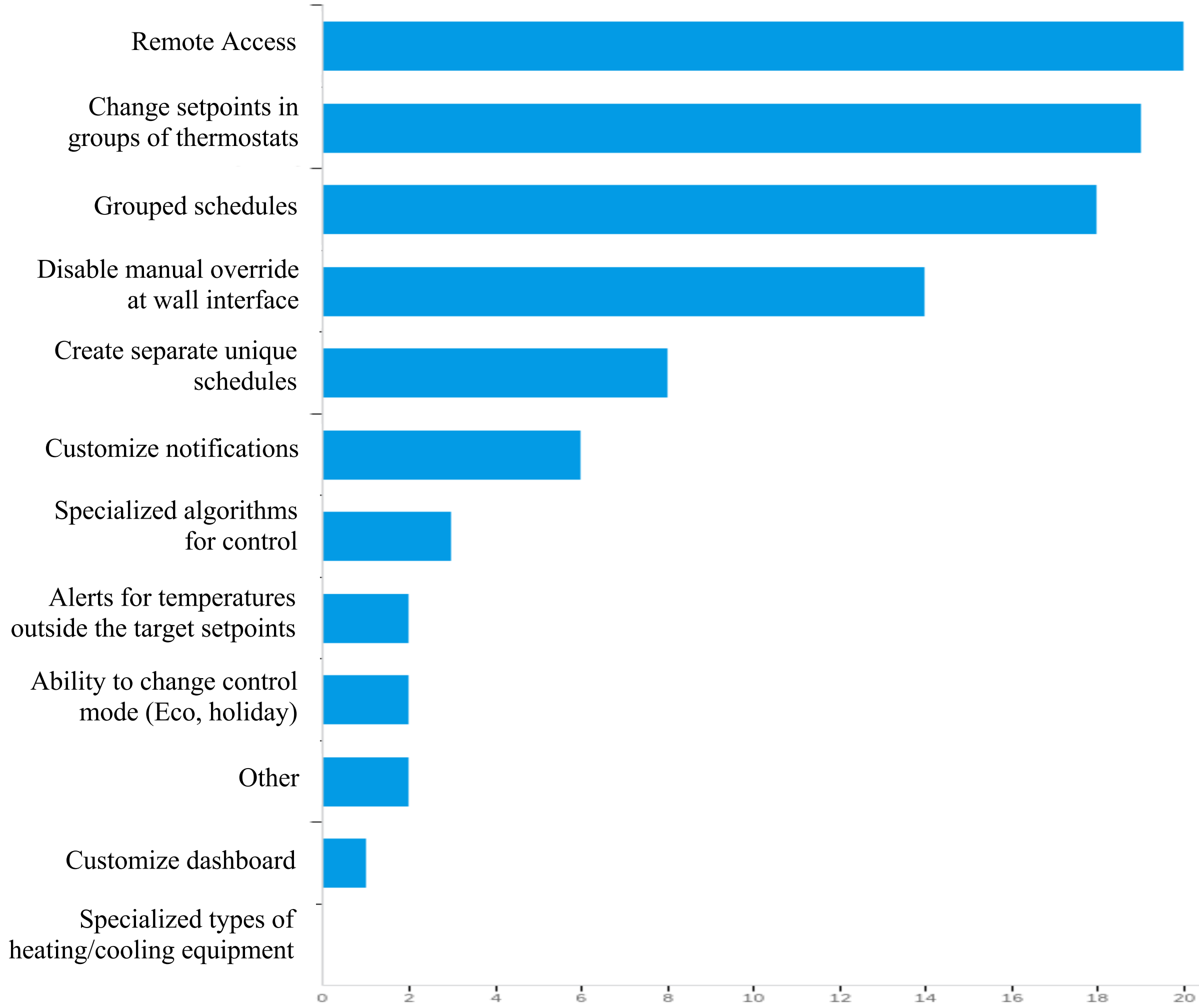
Results

This section provides the results from the interviews with facilities managers, evaluation of sensors, and heuristic evaluation.

Facilities managers and energy service provider interviews

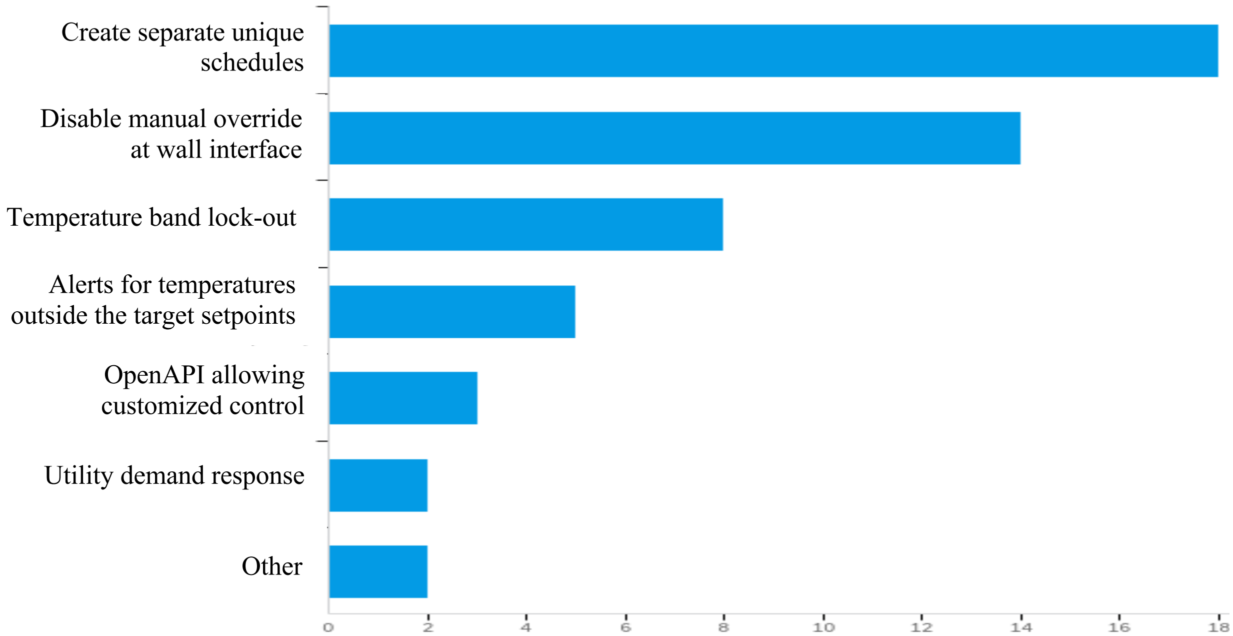
The interviews were conducted via Zoom between July-September 2020. In terms of how the facility managers interacted with advanced thermostats, we learned that both web browser and mobile phone access were equally used and valued. In response to questions on what features of advanced thermostats were commonly used (Figure 4), popular responses were: Remote access, Changing setpoints, Setting up schedule, and Disabling manual override.

Figure 4: Responses to “what features of advanced thermostats do you use?”



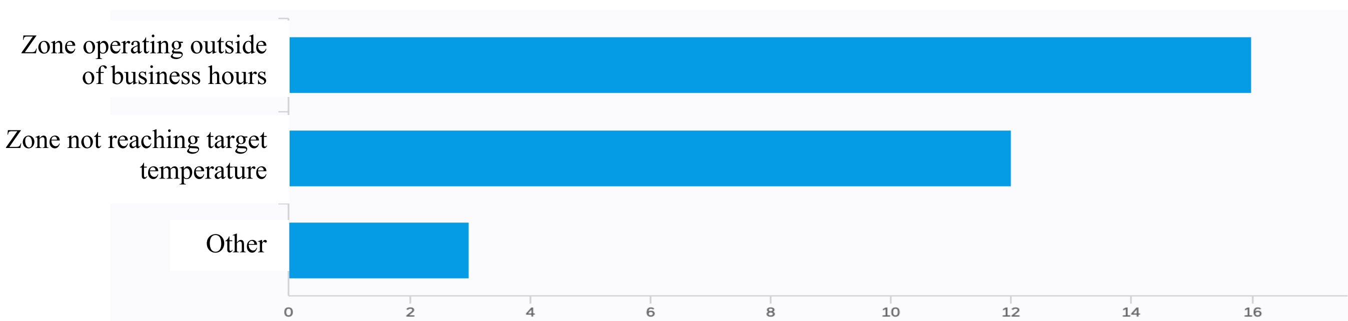
In response to the question “what feature of the advanced thermostat helps you save energy?”, the most common response was: “Create unique schedules”, followed by “Disable manual override of setpoints at the wall interface”, and “Temperature band lock-out.” Surprisingly, only three used the Open API of the thermostat for access and only two used utility demand response functions. See Figure 5.

Figure 5: Responses to “what feature helps you save energy?”



An impressive 95% of respondents said they use the energy reports. A large majority (80%) replied that cost was a factor in choosing thermostats. When asked what feature would be helpful that is currently not available today, a number of respondents indicated “help with maintenance”, while others responded that carbon dioxide sensors would be helpful. A few interviewees wanted better access to data. By far the most useful notification is when the HVAC zone is operating outside of business hours, and if the zone temperature does not reach the target temperature (Figure 6).

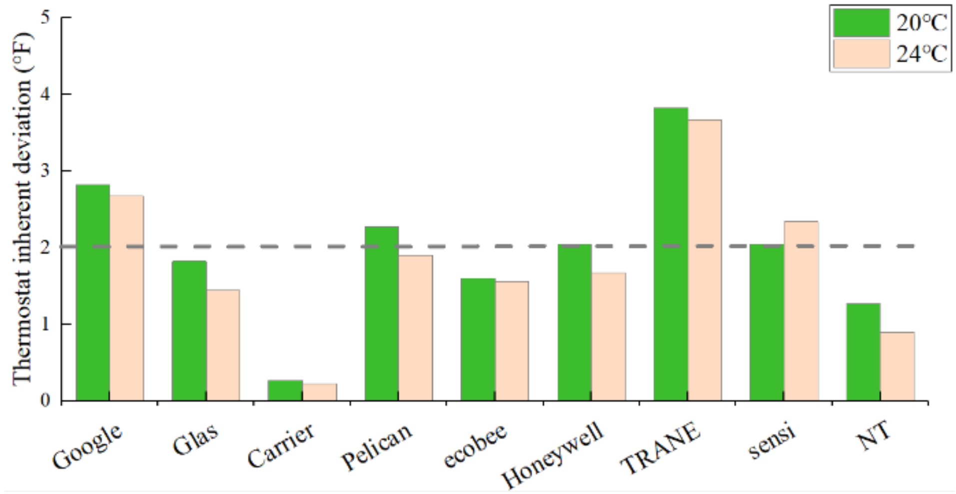
Figure 6: Response to what notification do you find useful

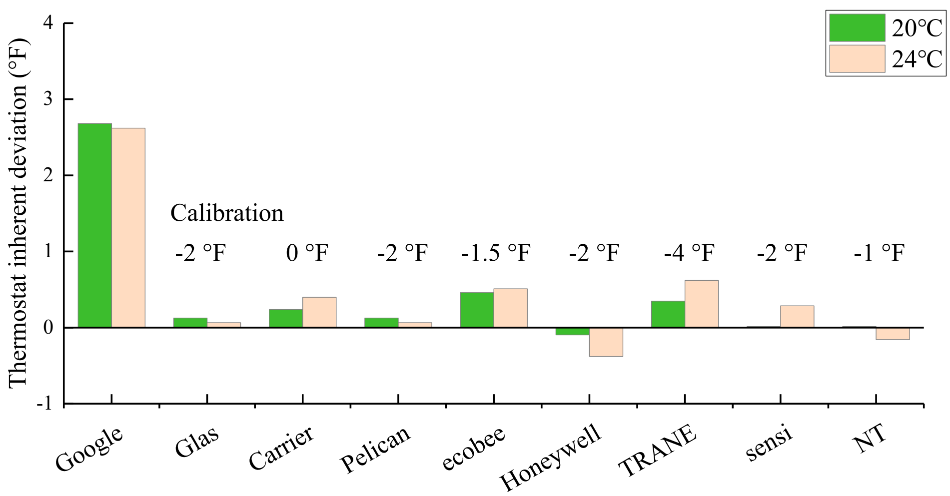


Sensor testing

The thermal symmetry testing, before calibration, evaluated the temperature sensor of each thermostat compared to a highly accurate control sensor. The Carrier thermostat tested as the most accurate one with the deviation less than 0.5°F, while other thermostats’ deviations varied from 1 to 4°F. The thermostats that had the capability of calibrating the sensor were then adjusted. After calibration, all the thermostats had deviations less than 0.5°F, except for the Google Nest which had no calibration function.

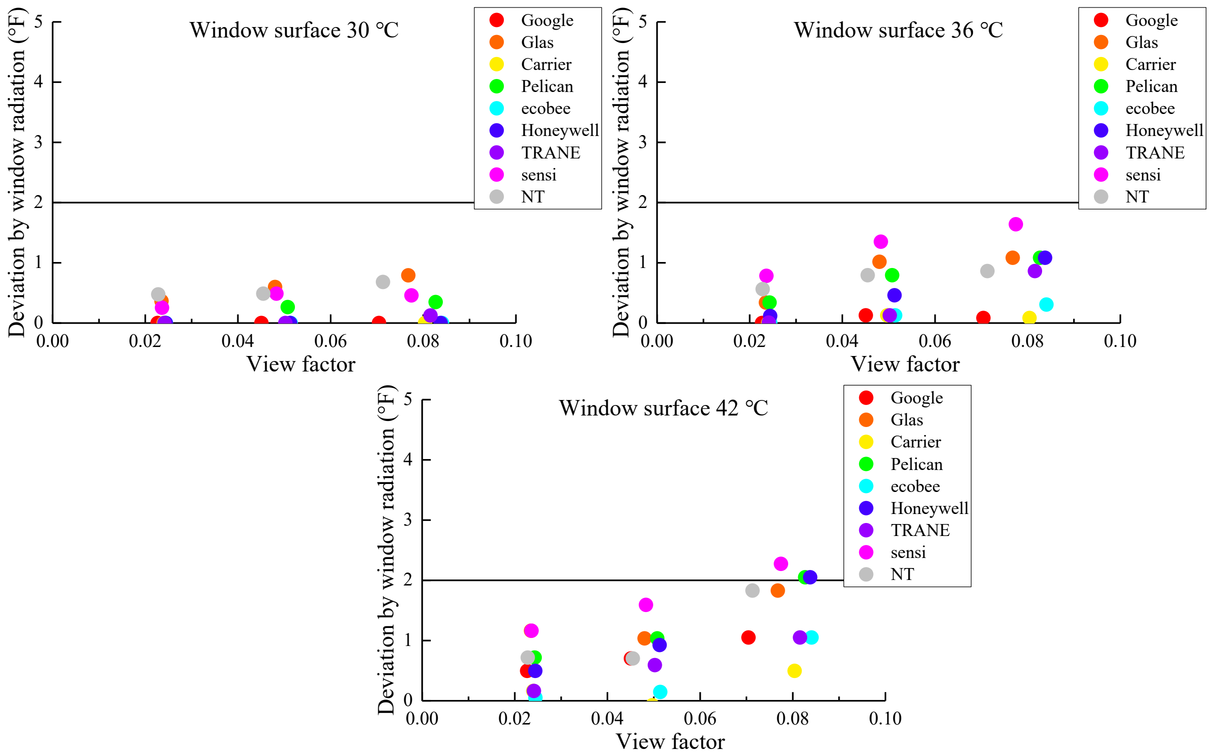
Figure 7. Thermostat deviations in the thermal symmetry testing at 68F (20C) and 75F (24C). Top: before calibration; Bottom: after calibration





In the thermal asymmetry testing, the measured temperature deviation increased with view factor and with the higher temperature surface (97F (36C), and 107.6F (42C)). Most of the deviations caused by the warm window surface were less than 2°F for all window temperatures tested, as shown in Figure 8, except for the highest temperature (107.6F (42C)). and greatest view factor. The Carrier thermostat was the least affected by the warm window; the deviation is always lower than 0.5°F. Other thermostats (Google Nest, Trane, and Ecobee) also performed well with the highest deviation of 1.0°F.

Figure 8. Thermostat deviations in the thermal asymmetry testing, for exposure to a heat panel simulating a window surface: at 30C, 36C and 42C.



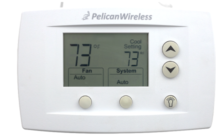
Heuristic evaluation

The first phase of the heuristic evaluation was using the wall interface of each thermostat to perform the tasks listed in Table 3. Table 4 lists the results of each heuristic for each thermostat, with green marking support of the heuristic and orange suggesting lack of support. Not any one thermostat was perfect; some were certainly easier to use than others. Many had good visibility of the available options on the home screen. Only half had a "wide and shallow decision tree[3]." Most had decent navigation cues so one could move through the choices in the display and know where one was and how to return. Nearly all of the thermostats displayed a clear hierarchy of display, which is useful for providing the most common information quickly. Most of them had standard text and icons (e.g., flame for heat, snowflake for cooling). Not many thermostats supported natural mappings: a few had very clear menus, but a few had touch/scroll features that were difficult to master. On one, you couldn't tell whether text displayed was a touch button to toggle or merely displaying the state. Some thermostats showed better error prevention, using "Done" or "Cancel" prompts. Feedback from controls also showed mixed results with some thermostats having audible chirps to confirm selection or knobs to turn and push.

Some thermostats had advanced features: forming a group of thermostats and providing a single schedule (Honeywell, Trane, Pelican) or holiday schedule/mode (Trane, Johnson Controls, Carrier, Ecobee), and a few with air quality features.

[3] A good analogy of a wide and shallow decision tree is nested folders on a computer: having multiple choices at the top denotes a "wide" tree, and the number of levels (folders inside of folders inside of folders) refers to the depth.

Table 4: Results of heuristic evaluation

Heuristic	Honeywell 	Trane 	Emerson Sensi 	Johnson Controls 	Carrier 	Google Nest 	Pelican 	Ecobee 	NT 
Visibility of available options (home)	Only current temp + heat mode + menu	Typical + hold	Words + icons	Can't turn off from Home screen	Only temp/fan icons, home icon	Can't easily turn off	Typical words	Icons	Text (small), Icon
A wide and shallow decision tree	Not easy to find "set time/day"	Many choices	Two deep	Too deep	Have to scroll thru 15 icons (6 at a time)	Need to scroll through many choices	*Only Tasks 1 and 3 accomplished with wall interface*	Menu screen	Poor for setting time/day
Navigation cues	Done, Back, but lots of scrolling required	Done, Back, Apply	Back arrow visible	Pull down menu, Scroll dots, right/left arrow swipe	Home icon with two dots left and right—very clear	Linear scroll through choices, easy to get lost		Menu and arrows	Check marks, back arrow, up/down
Clear hierarchy of display	Large display of current temp	Large display of current temp	Large display of current temp	Large display of current temp	Large display of current temp	Large display of setpoint but not current temp	Large display of current temp	Large display of current temp	Large display of current temp
Consistency and standards	Terms, icons, orange/blue colors	Terms, icons, orange/blue colors	Terms, icons, orange/blue colors	Ok, Terms, some icons	Icons, colors, means of setting day/time	Terms, icons, orange/blue colors	Up/down arrows	Terms, icons, colors; color of icon very subtle	Icons not standard but intuitive; requires toggling
Natural mappings	Menu icon, "heating to" text; hold until text	Touch and hold is not easy to use	Not easy to change temperature in auto mode	Two settings have same icon	Touch scroll not easy to use	Target temp shown only at beginning of period		Clean menu navigation	Unclear when icon is a touch button or reflecting state
Error prevention and recovery	No Save or Confirm; information icon	Text explanation of choices; small arrow buttons;	Ok. Cancel, Save.	Done, Cancel	Several ways to access some features	Done		Text explanation	Flashing in set mode
Feedback from controls	Touch scroll difficult to control	Ok	Ok	Not easy to swipe and scroll	Audible chirp when making selection; sometimes push buttons nonresponsive	Turn knob to scroll and push knob to select		Touchscreen, menu bar, arrow selection	No color feedback

Dark Green: thermostat supports heuristic very well; Green: support heuristic; Grey: barely supports heuristic; Orange: does not support heuristic

Conclusion

This paper described the evaluation and testing of several advanced thermostats for the purpose of guiding procurement officers and facility managers in making decisions about thermostat selection. The research team reviewed relevant literature, interviewed 20 energy managers, conducted sensory accuracy tests, and performed a heuristic evaluation of several advanced thermostats currently on the market.

From this research, we learned that facilities managers and service providers primarily use the following features: remote access (100% of those interviewed), grouping thermostats and changing grouped setpoints (95%), setting up schedules (90%), and disabling manual override 70%). A vast majority (95%) of respondents said that they use the energy reports. The most useful thermostat notification is when the HVAC zone is operating outside of business hours (80%), and if the zone temperature does not reach the target temperature (60%). Respondents enjoyed creation of unique schedules and temperature band lock-out. Surprisingly, only three (15%) used the Open API of the thermostat for access and only two (10%) used utility demand response functions.

We discovered that thermostats differed in sensor accuracy: lighted colored screens may cause self-heating, which can throw off accuracy by as much as 4F; even a degree or two Fahrenheit can make a difference in saving energy. Thus, we recommend that installations of thermostats include a temperature calibration check. We studied thermal asymmetry: the sensitivity of the thermostat temperature sensor to a source of radiant heat, such as a large window, sunlight on a wall or any warm surface in the room, such as large equipment. The measured temperature deviation increased with view factor and with the higher temperature surface. Four of the thermostats showed a deviation of 1F or less even with the highest temperature and greatest view factor.

The heuristic evaluation was conducted of the wall interface by a single researcher and showed that while most of the thermostats performed well, some of these advanced thermostats may be difficult to use. Many had good visibility of the available options on the home screen. Only half had a "wide and shallow decision tree;" this can correlate with a user getting lost among the choices, especially if the basic navigation cues are not well established. and usability tests. Not many thermostats supported natural mappings: a few had very clear menus, but a few had touch/scroll features that were difficult to master.

Finally, this study was hampered by COVID-19 in delays in testing and modifying plans regarding the usability testing. In addition, we were not able to obtain one thermostat, and discovered that two thermostats (Glas and NT) have been discontinued at the time of this writing. The next steps include conducting a heuristic evaluation of the web interface of each thermostat and a survey to understand why more energy managers are not using utility demand response functions.

Acknowledgements

We are grateful to the Technical Advisory Committee members for their service, for the facilities managers and energy service providers we interviewed and finally to the California Energy Commission for funding this work under grant number EPC-17-034.

Disclaimer: This paper does not necessarily represent the views of the Energy Commission, its employees or the State of California. The Energy Commission, the State of

California, its employees, contractors and subcontractors make no warranty, express or implied, and assume no legal liability for the information in this paper; nor does any party represent that the uses of this information will not infringe upon privately owned rights. This paper has not been approved or disapproved by the California Energy Commission nor has the California Energy Commission passed upon the accuracy or adequacy of the information in this paper.

References

Bordass, B, A Leaman, and R Bunn. 2007. "Controls for End Users a Guide for Good Design and Implementation." Reading, Berkshire, UK.

California Energy Commission (CEC). 2019. "Occupant Controlled Smart Thermostats." 2019. https://www.energy.ca.gov/rules-and-regulations/building-energy-efficiency/manufacturer-certification-building-equipment-7.

Consortium for Energy Efficiency (CEE). 2021. "CEE Residential Heating and Cooling Systems Initiative: Connected Thermostats." Middleton, MA. https://library.cee1.org/system/files/library/14493/CEE_ResHVAC_ThermostatSpecs_15Jan2021.pdf.

He, Yingdong, Edward Arens, Nianping Li, Zhe Wang, Hui Zhang, A Yongga, and Chenzhang Yuan. 2021. "Modeling Solar Radiation on a Human Body Indoors by a Novel Mathematical Model." *Building and Environment* 187 (107421). https://doi.org/10.1016/j.buildenv.2020.107421.

Hoyt, Tyler, Edward Arens, and Hui Zhang. 2015. "Extending Air Temperature Setpoints: Simulated Energy Savings and Design Considerations for New and Retrofit Buildings." *Building and Environment* 88 (June): 89–96. https://doi.org/10.1016/J.BUILDENV.2014.09.010.

Inncom. 2010. "Energy Savings and Operations Report: April 2010, Queens College, the City University of New York."

Johnson, Karl, Therese Peffer, and Jonathan Woolley. 2012. "Technical Report: Advanced Thermostat Control." Berkeley.

Kaplan, Katherine. 2009. "ENERGY STAR Programmable Thermostat Suspension Memo." Washington, DC. https://www.energystar.gov/ia/partners/prod_development/revisions/downloads/thermostats/Spec_Suspension_Memo_May2009.pdf.

Karjalainen, S. 2008. "The Characteristic of Usable Room Temperature Control." Helsinki University of Technology. http://lib.tkk.fi/Diss/2008/isbn9789513870607/.

Nielsen, J. 1994. "Heuristic Evaluation." In *Usability Inspection Methods*, edited by J. Nielsen and R.L. Mack. John Wiley & Sons. http://www.useit.com/papers/heuristic/heuristic_list.html.

Nielsen, Jakob. 1993. *Usability Engineering*. San Francisco: Morgan Kaufmann.

Norman, Donald A. 2002. *The Design of Everyday Things*. 1st Basic. New York: Basic Books. https://www.goodreads.com/work/best_book/18518-the-psychology-of-everyday-things.

O'Leary, Kat. 2012. "How Much Is 1° Worth?" Energy Hub. 2012. https://info.energyhub.com/blog/how-much-is-one-degree-worth.

Outcault, Sarah, Claudia Barriga, Kristin Heinemeier, John Markley, and Dan Berman. 2014. "Thermostats Can't Fix This: Case Studies on Advanced Thermostat Field Tests." In *American Council for an Energy Efficient Economy 2014 Summer Study for Energy Efficiency in Buildings*. Asilomar: ACEEE.

Peffer, T., D. Perry, M. Pritoni, C. Aragon, and A. Meier. 2013. “Facilitating Energy Savings with Programmable Thermostats: Evaluation and Guidelines for the Thermostat User Interface.” *Ergonomics* 56 (3). https://doi.org/10.1080/00140139.2012.718370.

Peffer, T., M. Pritoni, A. Meier, C. Aragon, and D. Perry. 2011. “How People Use Thermostats in Homes: A Review.” *Building and Environment* 46 (12). https://doi.org/10.1016/j.buildenv.2011.06.002.

Peffer, T, Gabe Fierro, Marco Pritoni, Moustafa AbdelBaky, Daniel Lengyel, Greg Thomson, Anand Prakash, et al. 2019. “XBOS-DR: Customer-Controlled, Price Mediated, Automated Demand Response for Commercial Buildings Final Report California Energy Commission.” Sacramento. http://uc-ciee.org/.

Perry, D., C. Aragon, A. Meier, T. Peffer, and M Pritoni. 2011. “Making Energy Savings Easier: Usability Metrics for Thermostats.” *Journal of Usability Studies* 6 (4): 226–44.

Polson, Peter G, and Clayton H Lewis. 1990. “Theory-Based Design for Easily Learned Interfaces.” *Human-Computer Interaction* 5 (2): 191–220. https://www.researchgate.net/publication/238670665_Theory-Based_Design_for_Easily_Learned_Interfaces.

Shneiderman, B. 1988. “We Can Design Better User Interfaces: A Review of Human–Computer Interaction Styles.” *Ergonomics* 31 (5): 699–710.

Shneiderman, B, and C Plaisant. 2009. *Designing the User Interface: Strategies for Effective Human–Computer Interaction*. 5th ed. Reading, MA: Addison Wesley.

Telkonet, New York University &. 2011. “Webinar on Occupancy Sensing Thermostats.” http://www.telkonet.com/resources.php.

US Department of Energy (DOE). 2021. “Programmable Thermostats.” Energy Saver. 2021. https://www.energy.gov/energysaver/programmable-thermostats.

———. 2022. “Smart Thermostats, Wi-Fi Enabled, Digital & Programmable.” EnergyStar. 2022. https://www.energystar.gov/products/smart_thermostats.

Woolley, Jonathan, M. Pritoni, M. Modera, and T. Peffer. 2014. “Why Occupancy-Responsive Adaptive Thermostats Do Not Always Save - and the Limits for When They Should.” In *Proceedings of the ACEEE Summer Study on Energy Efficiency in Buildings*. Washington, DC: American Council for an Energy Efficient Economy.