Facilitating Energy Savings with Programmable Thermostats: Evaluation and Guidelines for the Thermostat User Interface

Journal:	*Ergonomics*
Manuscript ID:	TERG-2011-0341.R2
Manuscript Type:	SI - Ergonomics and Sustainability
Date Submitted by the Author:	n/a
Complete List of Authors:	Peffer, Therese; University of California, Berkeley, California Institute for Energy and Environment Perry, Daniel; University of Washington, Department of Human Centered Design & Engineering Pritoni, Marco; University of California, Davis, Mechanical & Aeronautical Engineering Aragon, Cecilia; University of Washington, Department of Human Centered Design & Engineering Meier, Alan; Lawrence Berkeley National Laboratory,
Keywords:	consumer ergonomics < Application Domains, product design < Product, Equipment and Built Environment Design, user testing < Product, Equipment and Built Environment Design, controls and input devices < Human-Machine Systems

Facilitating Energy Savings with Programmable Thermostats: Evaluation and Guidelines for the Thermostat User Interface

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

[a]*California Institute for Energy and Environment, 2087 Addison Street, 2nd Floor, University of California Berkeley, Berkeley, California, U.S.,* [b]*Department of Human Centered Design & Engineering, University of Washington, 407A Sieg Hall, Box 352315, Seattle, WA 98195,* [c] *Mechanical & Aeronautical Engineering, University of California Davis, Davis, CA 95616,* [d]*Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720*

corresponding author's information: 00-1-510-289-4278, therese.peffer@uc-ciee.org

Facilitating Energy Savings with Programmable Thermostats: Evaluation and Guidelines for the Thermostat User Interface

Thermostats control heating and cooling in homes—representing a major part of domestic energy use—yet, poor ergonomics of these devices has thwarted efforts to reduce energy consumption. Theoretically, programmable thermostats can reduce energy by 5-15%, but in practice little to no savings compared to manual thermostats are found. Several studies have found that programmable thermostats are not installed properly, are generally misunderstood, and have poor usability.

After conducting a usability study of programmable thermostats, we reviewed several guidelines from ergonomics, general device usability, computer-human interfaces, and building control sources. We analyzed the characteristics of thermostats that enabled or hindered successfully completing tasks and in a timely manner. Subjects had higher success rates with thermostat displays with positive examples of guidelines, such as visibility of possible actions, consistency and standards, and feedback. We suggested other guidelines that seemed missing, such as navigation cues, clear hierarchy, and simple decision paths.

Keywords: thermostat, user interface, energy, usability, heuristic evaluation, residential

Practitioner Summary: Our evaluation of a usability test of five residential programmable thermostats led to the development of a comprehensive set of specific guidelines for thermostat design including visibility of possible actions, consistency, standards, simple decision paths, and clear hierarchy. Improving the usability of thermostats may facilitate energy savings.

1 Introduction

Many energy efficient items of equipment or energy reduction measures save energy from the time they are installed, such as efficient refrigerators or building insulation; however, others require the active participation of an informed human. Programmable thermostats (that can automatically relax temperatures at night or during unoccupied periods) have been promoted all over the world to save energy used to heat and/or cool people's homes. In the U.S., where nearly two-thirds (64%) of the residential heating systems use central air, energy for residential heating and cooling amounts to 9% of the total primary energy use (Energy Information Administration (EIA) 2010, U.S. Department of Energy (DOE) 2011a). A widely used rule of thumb is a savings of 1.8% per degree C (1% per degree F) for an eight hour adjustment (Nelson and MacArthur 1978). Yet programmable thermostats have largely failed to save energy due to poor ergonomics. A recent literature review (Peffer *et al.* 2011) uncovered many reasons: improper installation (e.g., mounted sideways, in dimly lit corridors, too high/too low), poor interface (e.g., buttons too small, icons/terms not understood), and misunderstanding of how thermostats work in particular and how heating/cooling systems work in general.

The ergonomics of a thermostat involves understanding its context to analyze the demands placed on the user's capabilities during its use. A typical residential thermostat controls the heating and/or cooling equipment, provides a user interface for the occupant to read current status and adjust the control, and contains at minimum a temperature sensor to provide control feedback. The thermostat is typically mounted on a wall and wired to the heating, cooling, and/or ventilation system; ideally, this wall is an interior wall insulated from outdoor conditions and centrally located in the house. The de facto standard in the U.S. for placement of the thermostat on the wall is 1.524 meters (60 inches) from the floor. The physical location alone not only affects the functionality of the device (due to the temperature sensor) but several user capabilities, such as reach and vision. In our literature review, we found examples of programmable thermostats located in inaccessible places (e.g., behind furniture), turned sideways (i.e., to accommodate the aspect dimensions of the previous thermostat so the users didn't have to repaint the wall), too high or too low to read/access easily, and in dimly lit corridors, precluding visibility.

The human factors of the thermostat also include the user interface, which impacts human vision, communication, cognition, and dexterity. The basic thermostat interface needs to: allow the user to provide a comfortable temperature (whether heating or cooling), set a schedule for convenience (e.g., heat before getting out of bed in the morning) and for energy savings (e.g., turn off heat/cool systems when no one is at home or when one is going out for a few hours or several days; adjust temperatures at night to reduce heating and cooling).

Modern programmable thermostats are following two other consumer electronics trends: from using an analogue display to digital (and often from graphic to numeric display) and from mechanical operation (e.g., knobs, sliders) to electronic (e.g., push buttons, touchscreen). In addition, programmable thermostats are increasingly complex, beyond the base features of setting start and end times for desired temperatures for the day, and days of the week. Today's programmable thermostats have more control features and display parameters (e.g., status of filter and battery, amount of energy consumed).

In general, programmable thermostats do not have much market penetration in the U.S. and have not conclusively demonstrated energy savings. Although programmable thermostats have been available for more than 30 years, only 30% of U.S. households have installed them (Energy Information Administration (EIA) 2005c). Several studies indicate that many people do not use programmable thermostats as designed. Only 55-60% are used to adjust temperatures at night for cooling and heating seasons respectively (Energy Information

Administration (EIA) 2005b) (Energy Information Administration (EIA) 2005a). Approximately half are in “hold” mode, effectively disabling the programming features (Archacki 2003). Several studies have indicated no significant savings with programmable thermostats (Cross and Judd 1997, Nevius and Pigg 2000, Haiad *et al.* 2004).

The U. S. Environmental Protection Agency reviewed many field studies and concluded that consumers were not using programmable thermostats effectively due to programming difficulties and lack of understanding of terms such as set point (Harris 2008). As a result, the EPA discontinued the EnergyStar programmable thermostat program in December 2009. Indeed, several recent usability tests with thermostats indicate continued problems (Karjalainen 2009, Sauer *et al.* 2009, Combe *et al.* 2011, Perry *et al.* 2011), suggesting that the industry in general has not responded to improve these interfaces nor outlined means of testing them with users. [We note that Consumer Reports tests thermostats in well lighted rooms with users sitting down.]

While the study of ergonomics specific to thermostats is not new, thermostat manufacturers in general do not seem to have applied this in their design of thermostats. Thirty years ago Moore and Dartnall (1982) described issues such setting the time on programmable thermostats and Dale and Crawshaw (1983) observed the effect of font size and controls (Moore and Dartnall 1982, Dale and Crawshaw 1983). A quick review of modern programmable thermostats indicates that not only have the findings in these early studies been ignored, but also technology has changed dramatically and the functions have increased in complexity. One hypothesis is that in general industries have grown from product-driven and consumer focused to financially-driven (Foroohar 2011); we can only guess that manufacturers typically balk at the cost of design of such a mundane device as a thermostat.

We did not find any existing guidelines or heuristics specific to guiding better ergonomics design of residential programmable thermostats. However, many ergonomics textbooks (e.g., Sanders and McCormick 1993) provide the basics for visual display, cognition and basic controls (Sanders and McCormick 1993). More domain-specific are guidelines on human-computer interfaces. These are applicable to this study because current programmable thermostats may be described as embedded devices—having computer systems that are limited in scope and designed to do dedicated functions. The human-computer interface design realm provides many guidelines, such as on Internet or web interfaces, and even regarding smart phone interfaces and touchscreens (Nielsen and Molich 1990, Nielsen 1994b, Cooper *et al.* 2007, Shneiderman and Plaisant 2009). Two guidelines target commercial building controls (Wyon 1997, Bordass *et al.* 2007), and one recent guideline was specifically developed for thermostats for offices (Karjalainen 2008). Many of these latter principles are specific to office buildings and not residential settings.

This study examines several guidelines and uses them to evaluate user mistakes and success performing tasks on several programmable thermostats in a usability test. We use the guidelines in our analysis of the parameters of the thermostats that led to errors and confusion as well as those that led to successful completion. In evaluating what worked and what was missing, we then developed a new set of guidelines for residential thermostat interface design.

2 Methods of evaluating interfaces

We reviewed several guidelines for user interfaces to assess their potential for evaluating programmable thermostats. The first list (Table 1) suggests potentially relevant parameters from a basic ergonomics textbook. These principles are grouped into those applicable to location, vision, controls, and cognitive abilities. While these guidelines may be generally applied, they are fairly detailed and specific. Many of these guidelines were developed through measurements, testing and observations of humans.

Table 2 provides guidelines from general device interface usability (Polson and Lewis 1990, Norman 2002), the computer-human interface realm (Nielsen and Molich 1990, Nielsen 1994b, Shneiderman and Plaisant 2009), commercial building controls (Bordass *et al.* 2007), and commercial building thermostats (Karjalainen 2008). These are grouped horizontally as best possible to see overlap among the guidelines. These guidelines are fairly generally applied, and without much detail. The next few paragraphs describe the various guidelines.

The ergonomics guidelines have more specifics regarding the "how" of interface design—where located for best reach and vision, what kind of display is best for the type of task, and the size and type of font. The other guidelines are more heuristic in nature in providing general design guidance.

Norman, Polson and Lewis provide general device usability guidelines. Norman discusses natural mappings between the real world and how users think. Polson and Lewis describe attributes of "walk-up-and-use" applications, which we find particularly appropriate to residential thermostats (Polson and Lewis 1990).

The Shneiderman & Plaisant and Nielsen & Molich user interface design guidelines seem applicable to embedded devices (e.g., visible system status, use of conventions and standards, and minimizing errors). Limitations of embedded devices such as programmable thermostats as compared to general-purpose computers include limited screen size and dedicated functions (to reach a certain state). This limits the ability of the user to fully explore or gather data more freely. For example, many programmable thermostats would not be able to support the help function and wizards typical of most computer software programs.

Bordass, Leaman and Bunn describe end use requirements for more usable controls commonly found in commercial buildings (such as lighting, fans, windows, and thermostats). The Bordass et al. list provides criteria that are specific to controls (need for fine control, amount of use), but does not take into account the potential of an embedded device as a control.

Karjalainen reviewed six different usability guidelines in developing his own guideline for office thermostats (Karjalainen 2008). He noted in general that the guidelines available do not take into consideration thermal inertia (e.g., the time delay in reaching the desired temperature), psychological, behavioral and physiological components of human thermal comfort, the occupant's lack of knowledge of how the heating/cooling system works, the occupant's false idea of comfortable temperatures (e.g. in practice one's thermal comfort range is much wider than an occupant often thinks), and the characteristics of heating and cooling systems. Karjalainen provides several specific guidelines for temperature controls, detailing the type of feedback (both controls and room); he also suggests providing advice on comfortable room temperatures. Finally, he suggests usability testing as part of the design process.

There is much overlap in these guidelines: easy to use and understand, visibility, need for feedback, aesthetics, simplicity, and not relying on people's memory in using the device. The latter two guidelines were designed for commercial buildings, and thus do not consider context and motivation for using the controls from a residential perspective (e.g., comfort for guests, saving money).

We developed a number of research questions: What can we learn from these guidelines to help evaluate the usability of programmable thermostats—what is useful and what is missing? Specifically, what element(s) of the interface renders it easier for the subject to complete the task? What features create frustration and prevent task completion?

3 Research design

In 2010, we conducted a usability study on five commercially-available residential programmable thermostats (three touchscreen, one web, and one button-based—see Table 3), with 31 participants, involving five tasks to evaluate device usability and effectiveness. Each subject interacted with two thermostats (one at a time) installed at a height in the lab typically seen in most U.S. houses. The details of the test and usability metrics developed to evaluate the devices' usability and the users' effectiveness at performing common thermostat tasks are described in (Perry *et al.* 2011). A video recording of each session was used to input numerous categories of data including task completion, time on task, function path (buttons and function interactions), interaction motions (press, slide, hold, etc.), interaction errors, and experimenter observations regarding users' confusion during the task.

This study looked at each trial, defined as a subject attempting one of five tasks on a given thermostat, to determine what parameters led to good versus poor usability. We used the guidelines to categorize these parameters. We analyzed each subject's actions during successfully completed tasks compared with non-completed tasks (included tasks not completed or not successfully completed). In addition, we observed the path length to complete each task and the total time on task, either to complete the task or ending time when the task was not successfully completed.

We developed tasks reflecting the typical functions of programmable thermostats. The most important were: ability to turn on/off heat or cooling, temporarily turning up or down heating or cooling, checking the current and target temperatures, setting time (and changing time at the beginning and end of Daylight Savings Time), and adjusting systems to use less energy when one is away. These tasks were also chosen in consideration of their effect on residential energy consumption. The five tasks finally selected for this study are in Table 4.

4 Analysis of existing designs

This section describes the success rate and time on task for subjects for each thermostat per task. Success refers to accomplishing the stated goals of the task; time on task marked the ending time for each task, when the subject stated that he/she was finished (whether successful or not). We note here that there were cases where the subject felt he/she had completed the task correctly, but in fact were unsuccessful, cases where the subject was not sure whether he/she had completed the task correctly or not, and cases where the subject clearly gave up, not knowing what else to do to complete the task. The development of the time and success metric is outlined in (Perry *et al.* 2011); another complementary study looked at optimal path length—the minimum number of button pushes and other actions required to complete the task (Pritoni *et al.* 2011). We also show time for incomplete tasks; this includes subjects verbally stating he/she was finished with the task (whether or not he/she had completed it) and incorrectly completing the task. For each task we describe various elements of the interfaces that seem to contribute to success and timely task completion as well as those that created confusion and difficulties. Figure 1, Figure 2, Figure 3, Figure 4 and Figure 5 show annotated pictures of the thermostats.

4.1 Task 1: Set thermostat to HEAT mode.

This was considered to be the easiest task; the optimal path length—the shortest series of actions necessary to complete the task—ranged from 1 to 4 (see Figure 6). In general, the best performing thermostat interface was the web interface—100% of the subjects completed this task, and all under 30 seconds. However, the worst performing thermostat only had a 46% success rate.

The two thermostats with the highest success rate had the "switch" visible at the home screen/default level. With the worst performing interface (HYB), the switch was hidden by a small cover, with no affordance or design features that hint that it was in fact openable. The BTN interface also had a cover; perhaps because the cover was larger led to better recognition by subjects. For the SMT thermostat, one had to go into the Details menu. Besides not being able to find the switch, other problems included confusion over the correct switch (between fan (ON-AUTO) switch and heat switch (OFF-HEAT)[1], and confusion over what is the current setting (e.g. the thermostat is currently in heat mode or off) versus whether system is on (e.g., furnace is currently producing heat).

4.2 Task 2: Set time and day

For task 2, the WEB interface was exempt, since its time and day stamp came from a networked computer. The Touchscreen (TCH) interface performed the best, with 100% of subjects completing the task, however a few subjects took more than two minutes to complete the task. See Figure 7.

One major problem was not finding the place (whether button or menu) to change the day/time. The TCH model had a Clock button at the home/default level. For the HYB, one could either press on the current time and day display to change or go through the Menu to Set Time/Day; both used up or down arrows on the touchscreen to change, but the Menu editor required saving by pressing “Yes”. However, for the SMT interface, one had to press the More button, then Settings, then Preferences to get to the clock function, then eight more button presses to finish the task. The BTN interface had the day/time button under a cover. In addition, subjects were confused whether they were setting universal time or time/day for a

[1] Forced air systems in the U.S. commonly have a fan switch (ON-AUTO), which controls the blower fan inside the house; ON recirculates air when no heating or cooling is taking place and AUTO allows the system to turn on blower fan as necessary. The system switch can be as complicated as OFF-HEAT-COOL-AUTO to control which system is running (furnace or air conditioning) or AUTO whether the system should decide depending on the interior conditions.

programmed temperature schedule. In general, subjects found it tedious to set time because each button push only incremented the time by one minute; if one held the button down, the time would scroll faster, but this was not obvious.

4.3 Task 3: Identify current target or "set to" temperature

This task showed a wide disparity among the thermostats with subjects performing well on three thermostats and poorly on the other two (Figure 8). Three devices showed the current temperature and the target temperature setpoint in the main screen. The information was presented clearly, with labels indicating "set to" or "set temperature" for the temperature setpoint compared to "current", "inside", "room", or "current temperature," but some of the labels used were quite small.

A common problem was in navigation: subjects were confused how to get to the temperature setpoint. For the Button (BTN) thermostat interface, one either immediately grasped the procedure or not at all; the target temperature was accessed from pressing either one of the up/down arrow push buttons that one would use to change the setpoint—not an obvious means of control. The Hybrid thermostat allowed one to press the screen where the current temperature was displayed to access the target temperature—again, not an obvious or natural connection.

One common error was the subject providing the current temperature instead of the set or target temperature, especially for the devices where only one temperature was shown (BTN and HYB). Another common error was the subject providing the set or target temperature for the wrong time of day.

4.4 Task 4: Identify future target or "set to" temperature

This task required the subject to identify a target temperature for a time in the future, and thus subjects had to look at the programmed schedule. The best performance came from using the web interface with a 71% completion rate. However subjects had difficulties completing the task with the other thermostat interfaces (Figure 9).

The main problem was that subjects did not know how to access the schedule of time and temperatures. Three of the five interfaces tested had a means of viewing the schedule; for the others, it was necessary to enter the “Edit Program” mode. The WEB interface had the schedule on the home screen, so this was the most accessible; the SMT had a button labeled “program” on the home screen and the TCH had a button labeled “Sched” for Schedule. The thermostat on which subject performed the worst (BTN) had a panel covering the program button and had serial access to this data, meaning that one had to scroll through all days of week and all time modes; one couldn’t skip to the desired time. In addition, only the start time for each period was indicated, and if one passed the setting for the desired time, one could not go backwards. Access using the HYB thermostat required several steps pressing areas on the touchscreen: first pressing Menu, and then pressing the Scroll button to find Set/review heat progrms and then selecting Yes, and then pressing the Next button to go through each day’s four time modes (MORN, DAY, EVE, NITE) until one reached the desired day/time. The best performing thermostats (WEB and SMT) had a two-dimensional graphic depicting day/time and temperature in a tabular form; the web interface provided information on single schedules just by hovering with the cursor over the calendar.

4.5 Task 5: Set for Away/Vacation

The final task asked subjects to set the thermostat for the condition that the house would be unoccupied for several days. Subjects using the WEB interface performed reasonably well, with 75% completion. Subjects using the Touchscreen (TCH) and Smart (SMT) interfaces were less successful. Less than half of those using the Button (BTN) interface were able to complete the task, and only one subject using the Hybrid (HYB) interface completed the task—taking five minutes to do so (Figure 10).

For each thermostat there were a few different ways of completing this task. Three thermostats (WEB, SMT, and BTN) had a "vacation" or "away" or "energy savings" mode; all five had Hold modes, in which one could change the temperature setting and "hold" it. Ironically, the thermostat with one of the shortest path lengths (HYB) posed the most challenge; one had to press on the word Hold on the screen and then the current temperature display to change the setpoint, but there is no affordance to suggest that Hold is in fact a button and is touch sensitive. Many subjects tediously changed the setpoints for all days and all time modes in the day. A common error was not saving properly, so the changes were lost. In general, subjects were confused regarding the terms/functions temporary override, timed hold, permanent hold, permanent override, away, and vacation.

4.6 Application of the guidelines

Some guidelines are meant to be applied in the field, and did not apply to our lab test of the programmable thermostats, such as location and feedback from the effect of the equipment. Other guidelines we did not explicitly test, such as glare, and font and icon size. As was shown in Table 2, many guidelines were similar. We attempted to group and summarize the most pertinent and applicable guidelines. The following table (Table 5) show this reduced set of guidelines and shows how each thermostat fared accordingly. We evaluated the parameters of each thermostat that seemed to correspond with successful and timely completion or incomplete and longer time on task. We categorized them according to each guideline. Positive examples of each guideline are labelled with a plus sign (+) and shaded in grey; examples that violate the guideline are labelled with a minus sign (-). Blank cells merely indicate that the thermostat did not clearly provide a positive example nor a violation of the guideline.

5 Discussion

Several of the interfaces were complicated and difficult for users to understand, leading to frustrations and major barriers for completing the tasks. The data from the usability test (successful completion, time to complete and ending time for incomplete tests) provided a means of evaluating the devices as well as probing the causes of problems and keys to success. Another combined measure of efficiency, ease of use, and few errors was actual path length compared to the ideal path length. In general, we assumed that a successful completion of a task meant the subject found the interface usable from a strictly functional point of view—in the case of this lab test, perhaps easy to learn was the key usability attribute.

Subjects performed better with some displays than others, and this varied by task. For example, subjects using the WEB display in general performed well. However, the WEB display is a special case in that it is not an embedded device, and has more capability (e.g., larger screen, familiar menu driven display). While more and more thermostats include control by smart phone and tablets, we anticipate that some sectors of the population will always require a stand-alone, simple, non-networked device. Some of the elements of the WEB display can be developed in embedded devices, such as the tabular display of temperature, time and day. We found all thermostats provided attributes that supported guidelines; even the least usable device had positive qualities, such as a responsive touchscreen that provided feedback and hierarchy of display. The variability in performance in general across thermostats and tasks indicates a variety of solutions for usability; we do not develop these guidelines towards a "one size fits all" solution.

Many of the guidelines were quite useful in understanding the ease or difficulty with which subject completed each task. One of the most important guidelines appeared to be the *visibility of available options*, associated with walk-up-and-use applications. For example, the cover or door on two thermostats seemed to reduce performance in hiding a few of the available options. Users got lost when the action choice was not available on the home screen, and the terminology of the choices was not clear (e.g., setting the time on the Smart thermostat). Arguably, a cover may make the thermostat more aesthetically pleasing with a more simplified appearance; however, in this case, the manufacturers could certainly provide better affordances to indicate that a panel or door is openable and how to open it.

Another pertinent guideline was *consistency and standards*. Both the Touchscreen and Hybrid thermostats had touchscreens; however, with the Hybrid thermostat, the "buttons" did not look like buttons. For example, only one subject was able to figure out the Hold function with the Hybrid; this may also reflect the lack of familiarity with this term. This also seemed to affect setting the day and time using the Hybrid (73% completion vs. 100% for the Touchscreen). In general, the terminology was neither standardized across thermostats nor natural or familiar; terms such as settings or setpoint, current, and hold seemed to create confusion. The Hybrid also was inconsistent in operation. For example, one could press on the time or day to set the time or day, but pressing on the current temperature provided the target temperature. It was difficult to know which words were touch-sensitive or not.

Lack of *feedback* was another common issue. Many users made errors when they failed to save changes. There was no confirmation prompt (e.g., do you want to save?) as is so common in computer interfaces. Previous studies have shown feedback on user behavior is vital to performance as well as satisfaction (Sauer *et al.* 2007).

Several issues were not addressed by the guidelines. More important than the path length (i.e., number of total actions), was a *broad and shallow decision tree (Shneiderman 1988),* that is, having many options available at first glance, and not very many "layers" or levels of choices (e.g., a menu (level 1) with x choices, each of which (level 2) have y choices, each of which (level 3) have z choices and so on). Each decision point represented a chance to get lost. Of course, the "width" of the decision tree at the first level is not infinite;

further studies may suggest a practical limit of the number of choices to keep the interfaces simple. For example, many web designers state that it takes longer to make a decision when presented the option within a large set of options versus a smaller set (attributed erroneously as Hick's Law) (Seow 2005, Johnson 2010).

One issue not fully addressed by the guidelines was in *navigation*, or knowing where one was. A common example was when the subject was confused between whether he was in edit mode and actually making changes to the scheduled program, or just viewing the scheduled program.

With respect to improving visual aesthetics and reducing cognitive load, another issue is the development of a *clear hierarchy*, so that the most often used functions are the most prominently displayed, such as having the largest font; less often used or perhaps functions for expert use might be buried but accessible down a level or two.

Both navigation cues and a clear hierarchy would complement the wide and shallow decision tree. These attributes would help prevent the user from getting lost, and thus develop more confidence about the interface.

The unique challenge of embedded devices such as programmable thermostats is that they are not full-fledged computers (with all the attention given human-computer interfaces), but they have more functions than "dumb" controls. As programmable thermostats evolve, we see larger screens and touchscreens, but this study shows that these do not ensure usability. Compounding the issue is that the number of functions seems to be increasing (an example of feature creep), which aggravates usability issues.

However, a positive development one author notes is the return of the moving pointer-fixed analogue display as a more intuitive mode of understanding the difference between the setpoint and current temperature (Peffer 2009). Thermal comfort is subjective and relative (de Dear and Brager 1998, American Society for Heating Refrigerating and Air-

Conditioning Engineers (ASHRAE) 2004) so an analogue display output suits this purpose; a temperature setpoint is a discrete number, which suggests a digital input.

The categorization of thermostat attributes according to each guideline was qualitative and arguably subjective; a related analysis showed that the combined quantitative measure of time to complete and success rate was a good proxy for usability of the thermostat per task (Perry *et al.* 2011). While a usability test and single metric can provide feedback to manufacturers on existing and prototyped thermostat designs, a new set of guidelines specifically for residential thermostat interfaces can inform these designs to begin with, as a proactive approach rather than reactive.

6 Recommendations

Our general recommendations include visibility of available options on the home screen, a wide and shallow decision tree, navigation cues, clear hierarchy of display, consistency and standards, natural mappings, error prevention and recovery, and feedback. We recommend that usability testing be part of the design process. More specific recommendations include the following:

- Include all important and often used actions at the home level; consider no covers or clearly provide affordances.
- Use a graphic tabular form to view the temperature setpoints for the time of day and day of week.
- When possible, include confirmation prompts (e.g., do you want to save?), or some other means of confirming when something is edited or changed.
- Use plain English wherever possible (no abbreviations) and standard icons.
- Use clear affordances
    - For touchscreens, buttons should look like and act like buttons

- if required, covers should be clearly marked so they look "openable".

Specific recommendations by Karjalainen include providing a clear way to adjust room temperature and detailing the type of feedback: there should be an adequate and fast effect on room temperature, and clear and sufficient feedback to the user after the temperature adjustment. He suggests acceptable default settings and providing advice on comfortable room temperatures. Finally, he suggests usability testing as part of the design process, especially including females.

7 Conclusion

While programmable thermostats are theoretically capable of saving 5-15% of energy to supply heating and cooling for residences, in practice, they save little to no energy due primarily to human factors issues. We reviewed several design guidelines intended to improve ergonomics and usability from various fields. Many of the guidelines converged, promoting visibility, feedback, consistency and standards, natural mappings, and recovery from errors.

After conducting a usability study with 31 subjects with five programmable thermostats in a lab setting, we used the converged guidelines to analyze the usability attributes of the thermostats. We answered several research questions. Some guidelines seemed more important to task completion than others. Both positive examples and violations of guidelines were critical to examine with respect to thermostat display usability. Finally, we developed guidelines to describe some issues not covered by the set of guidelines we used. Several guidelines proved vital to usability such as visibility of available actions, feedback, and consistency and standards; subjects had the highest success rate of task completion with the thermostats that embraced these guidelines. Missing guidelines included navigation cues, clear hierarchy, and wide and shallow decision trees; these are useful in completing desired actions with confidence and preventing users from getting "lost."

Improving the usability of thermostats must draw upon traditional proven ergonomics principles, with emphasis on quantitative measurements, detailed specifics on types of suitable controls, and analysis in developing a more holistic and quantifiable approach to design heuristics. As such, we embrace both qualitative proactive approaches, such as developing specific heuristic guidelines for residential programmable thermostats, as well as quantitative approaches, such as our previous work in developing usability metrics to provide a single quantitative measure of usability. At the same time, new techniques are needed to both assess and solve the unique problems related to embedded controls. We recognize that improving the usability of programmable thermostats may only represent one step towards facilitating energy savings, but it is a vital one. Finally, we anticipate these heuristic guidelines may be applied to other devices to facilitate energy savings, such as dishwashers, audio-visual equipment, and water heaters.

Acknowledgements: We wish to thank Jessica Granderson, Dhawal Mujumdar, Becky Hurwitz and Margarita Kloss for their contributions.

8 References

American Society for Heating Refrigerating and Air-Conditioning Engineers (ASHRAE), 2004. *Standard 55-2004: Thermal environmental conditions for human occupancy.* Atlanta.

Archacki, R., 2003. Carrier thermostat mode summary: Summer 2003. (Personal correspondence to Gaymond Yee).

Bordass, B., Leaman, A. & Bunn, R., 2007. *Controls for end users a guide for good design and implementation.* Reading, Berkshire, U.K.: Building Controls Industry Association.

Combe, N., Harrison, D., Craig, S. & Young, M.S., 2011. An investigation into usability and exclusivity issues of digital programmable thermostats. *Journal of Engineering Design*.

Cooper, A., Reimann, R. & Cronin, D., 2007. *About face 3: The essentials of interaction design* San Francisco: Wiley.

Cross, D. & Judd, D., Year. Automatic setback thermostats: Measure persistence and customer behaviored. *International Energy Program Evaluation Conference (IEPEC)*, Chicago
Dale, H.C.A. & Crawshaw, C.M., 1983. Ergonomic aspects of heating controls. *Building Services Engineering Research and Technology,* 4 (1), 22-25.
De Dear, R.J. & Brager, G.S., 1998. Developing an adaptive model of thermal comfort and preference. *ASHRAE Transactions,* 104 (1a), 145-167.
Energy Information Administration (EIA), 2005a. Residential energy consumption survey, housing characteristics tables, table hc14.5: Space heating usage.
Energy Information Administration (EIA), 2005b. Residential energy consumption survey, housing characteristics tables, table hc15.7: Air-conditioning usage.
Energy Information Administration (EIA), 2005c. Residential energy consumption survey: Preliminary housing characteristics tables. Energy Information Administration.
Energy Information Administration (EIA), 2010. Table 1.1 primary energy overview, 1949-2009.
Foroohar, R., 2011. Driven off the road by M.B.A.s: The rise of business schools coincided with the fall of American industry. *Time,* July 18, 2011, 22.
Haiad, C., Peterson, J., Reeves, P. & Hirsch, J., 2004. *Programmable thermostats installed into residential buildings: Predicting energy savings using occupant behavior & simulation*.
Harris, A.D., 2008. *Energystar is examining thermostat approach*: Air Conditioning, Heating and Refrigeration NEWS.
Johnson, R., 2010. *Ten laws to design by* [online]. Design+Strategy Available from: http://3.7designs.co/blog/2010/07/ten-laws-to-design-by/ [Accessed May 29, 2012
Karjalainen, S., 2008. *The characteristic of usable room temperature control*. Helsinki University of Technology.
Karjalainen, S., 2009. Thermal comfort and use of thermostats in finnish homes and offices. *Building and Environment,* 44 (6), 1237-1245.
Moore, T.G. & Dartnall, A., 1982. Human factors of a microelectronic product: The central heating timer/programmer. *Applied Ergonomics,* 13 (1), 15-23 Available from: http://www.sciencedirect.com/science/article/B6V1W-47YN42S-1T1/2/2f3f837ec75990d73af1ab39b02cca2a.
Nelson, L.W. & Macarthur, J.W., 1978. Energy savings through thermostat setbacks. *ASHRAE Transactions,* 83 (AL-78-1 (1)), 319-333.
Nevius, M. & Pigg, S., 2000. Programmable thermostats that go berserk: Taking a social perspective on space heating in wisconsin. *Proceedings of the 2000 ACEEE Summer Study on Energy Efficiency in Buildings*, 8.233-8.244 Available from: http://72.36.212.11/prod/berserk.pdf.
Nielsen, J., 1994b. Heuristic evaluation. *In* Nielsen, J. & Mack, R.L. eds. *Usability inspection methods.* New York: John Wiley & Sons.
Nielsen, J. & Molich, R., 1990. Heuristic evaluation of user interfaces. *CHI'90 Conference.* Seattle, WA: ACM, 249-256.
Norman, D., 2002. *The design of everyday things* New York: Basic Books.
Peffer, T., Pritoni, M., Meier, A.K., Aragon, C. & Perry, D., 2011. How people use thermostats: A review. *Building and Environment,* 46 (12), 2529-2541.
Peffer, T.E., 2009. *California dreaming: The design of residential demand responsive technology with people in mind*. University of California Berkeley.
Perry, D., Aragon, C., Meier, A., Peffer, T. & Pritoni, M., 2011. Making energy savings easier: Usability metrics for thermostats. *Journal of Usability Studies,* 6 (4), 226-244.

Polson, P.G. & Lewis, C.H., 1990. Theory-based design for easily learned interfaces *Human-Computer Interaction,* 5, 191-220.

Pritoni, M., Meier, A. & Perry, D., 2011. *Human factors in climate controls for small commercial buildings*: Institute for Building Efficiency: Johnson Controls, Technology Brief.

Sanders, M.S. & Mccormick, E.J., 1993. *Human factors in engineering and design,* Seventh ed. San Francisco: McGraw-Hill.

Sauer, J., Schmeink, C. & Wastell, D.G., 2007. Feedback quality and environmentally friendly use of domestic central heating systems. *Ergonomics,* 50 (6), 795-813.

Sauer, J., Wastell, D.G. & Schmeink, C., 2009. Designing for the home: A comparative study of support aids for central heating systems. *Applied Ergonomics,* 40 (2), 165-174.

Seow, S.C., 2005. Information theoretic models of hci: Comparison of the hick-hyman law and fitt's law. *Human-Computer Interaction,* 20 (3), 315-352.

Shneiderman, B., 1988. We can design better user interfaces: a review of human-computer interaction styles. *Ergonomics,* 31 (5), 699-710.

Shneiderman, B. & Plaisant, C., 2009. *Designing the user interface: Strategies for effective human-computer-interaction,* 5th ed.: Addison Wesley.

U.S. Department of Energy (DOE), 2011a. *2008 residential energy end-use splits by fuel type (quadrillion btu)* [online]. http://buildingsdatabook.eren.doe.gov/ChapterIntro2.aspx?2#1 [Accessed April 6 2011 Table 2.1.5].

U.S. Department of Energy (DOE), 2011b. *Thermostats and control systems* [online]. http://www.energysavers.gov/your_home/space_heating_cooling/index.cfm/mytopic=12720 [Accessed Aug 2, 2011

Wyon, D.P., 1997. Individual control at each workplace for health, comfort, and productivity. *Creating the Productive Workplace Conference.* London.

9 List of Tables

10 List of Figures

Table 1: Relevant parameters from an ergonomics textbook (Sanders and McCormick 1993).

Location:	Within reach for control and within sight on display, sufficiently lighted.
Visual display:	Illumination on screen, contrast between content and background, glare, font size and type (segmented font vs. dot matrix (fewer errors)). Button and switch size and position. Symbol size.
Control interface:	Fixed scale with moving pointer provides rapid clue of approximate quantity (and relative rate of change) and set-in quantity (natural relationship between control and display motions). Compatible/consistent with human expectations (faster learning and reaction time, fewer errors) Spatial compatibility: physical similarity of displays and controls. Movement compatibility: (e.g. up arrow or move to right or clockwise indicates increase) Hierarchy of control (rate/first order vs. higher order control) Push button for discrete information vs. sliding lever or turning knob for transmitting continuous information Coding controls: shape, texture, size, location, operational method, colour, labels Push vs. hold down: feel of resistance
Cognitive:	Alpha-numeric display good for identification and small space, also superior to analog when a precise numeric value is required and values are not continually changing. Symbolic signs preferable if the code symbol has an already fairly universally established association (no recoding from symbol to words to concept). Easiest to read straight line scale (vs. curved) with moving pointer and control moves the pointer (Heglin 1973).

Table 3: Description of thermostats tested.

Device	Type	Description
BTN	Buttons/ Switches	Button-based programming; full cover over device; user instructions on cover; 7-day programming.
HYB	Buttons with touchscreen	Hybrid of touchscreen (primary programming), switches under a cover (heating and cooling controls), and button for lighting; 7-day programming; ability to view past energy usage.
TCH	Touchscreen	Touchscreen with black/white display; 7-day programming.
SMT	Smart with touchscreen	Smart WiFi enabled device; full-colour LCD touchscreen; 7-day programming; quick save function.
WEB	Web portal	Web platform; 7-day programming; synched with wall device.

Table 4: Description of tasks.

Tasks	*Description*
Set Heat	Set the thermostat to HEAT mode. (Setting was OFF at the start of the task).
Set Time & Day*#	Set the thermostat to the current day and time. (The time settings were programmed to Monday at 12:00 AM for the start of the task.)
Current setting	Identify and read aloud the temperature that the thermostat was set to reach at that current time.
Future setting	Identify and read aloud the temperature that the thermostat was set to reach at a future period (Thursday at 9 pm). (No need to change any settings).
Vacation/Away/Hold	Set the thermostat to maintain the same temperature during a five-day period when one is away.

*not performed on the WEB because time settings could not be modified

#setting the day not performed for the TCH because this required a code from the manual

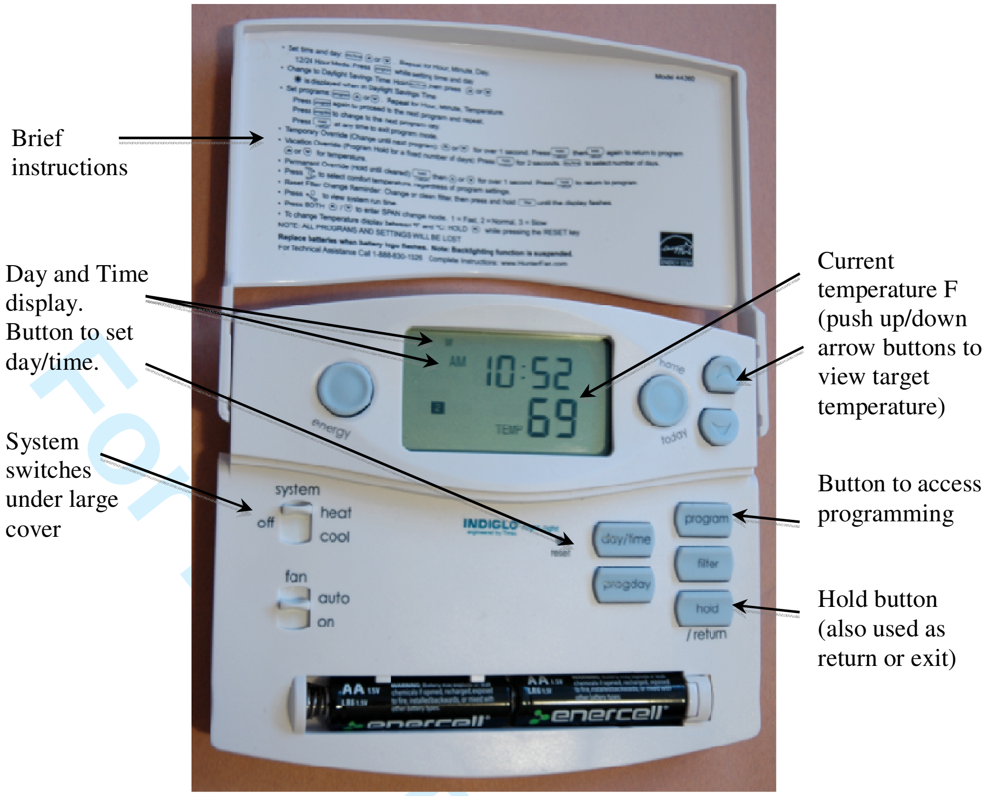

Figure 1: Button thermostat (BTN).

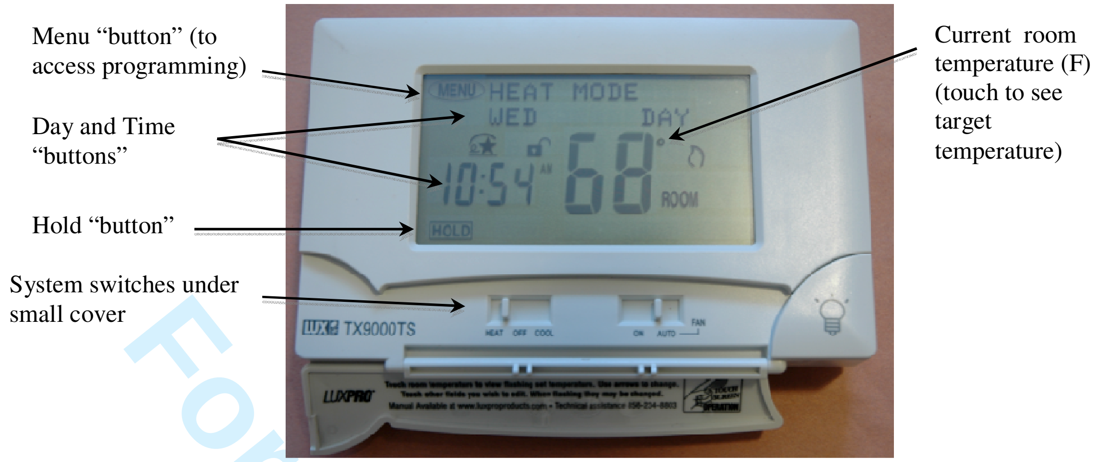

Figure 2: Hybrid thermostat (both touch screen and buttons/switches) (HYB).

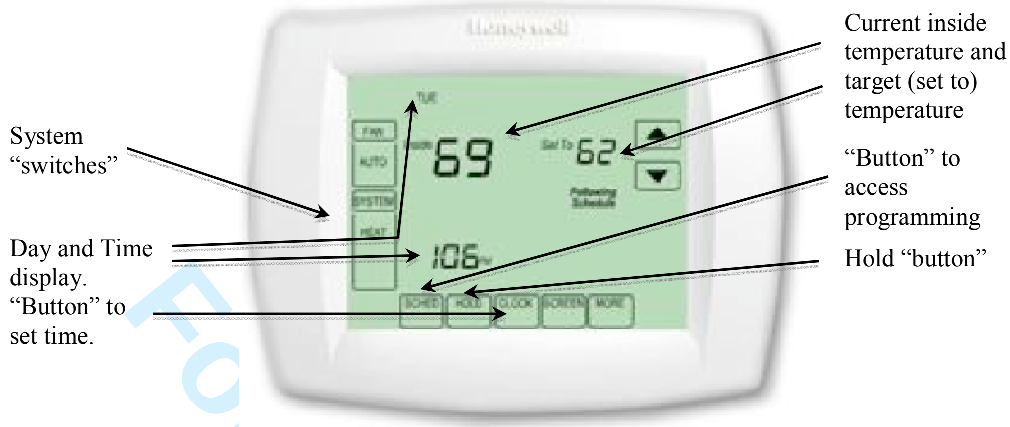

Figure 3: Touchscreen thermostat (TCH).

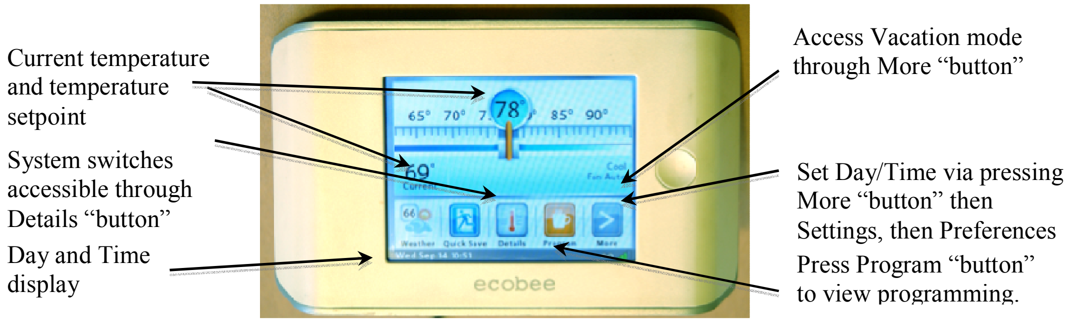

Figure 4: Smart thermostat (touchscreen with Home button) (SMT).

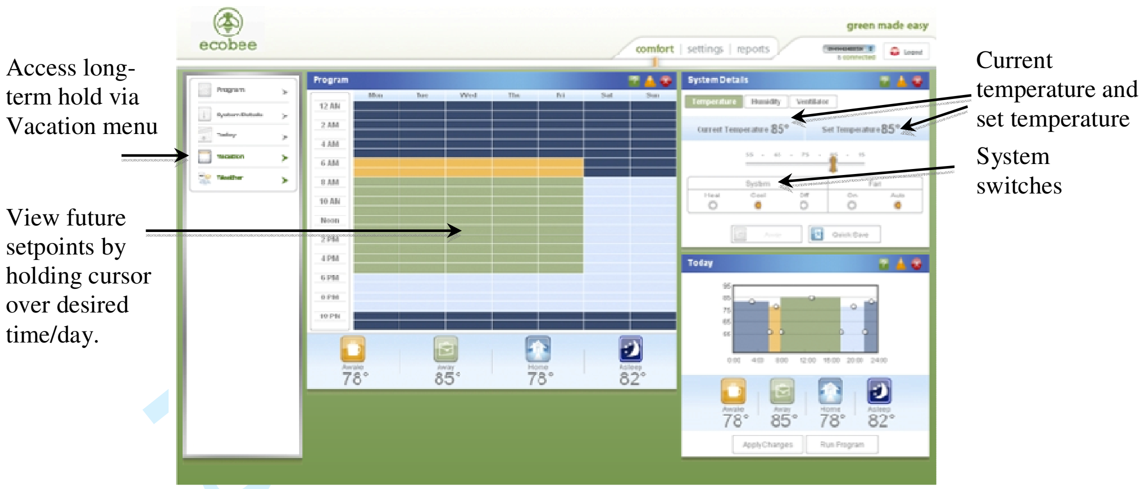

Figure 5: Web portal as a thermostat interface (WEB).

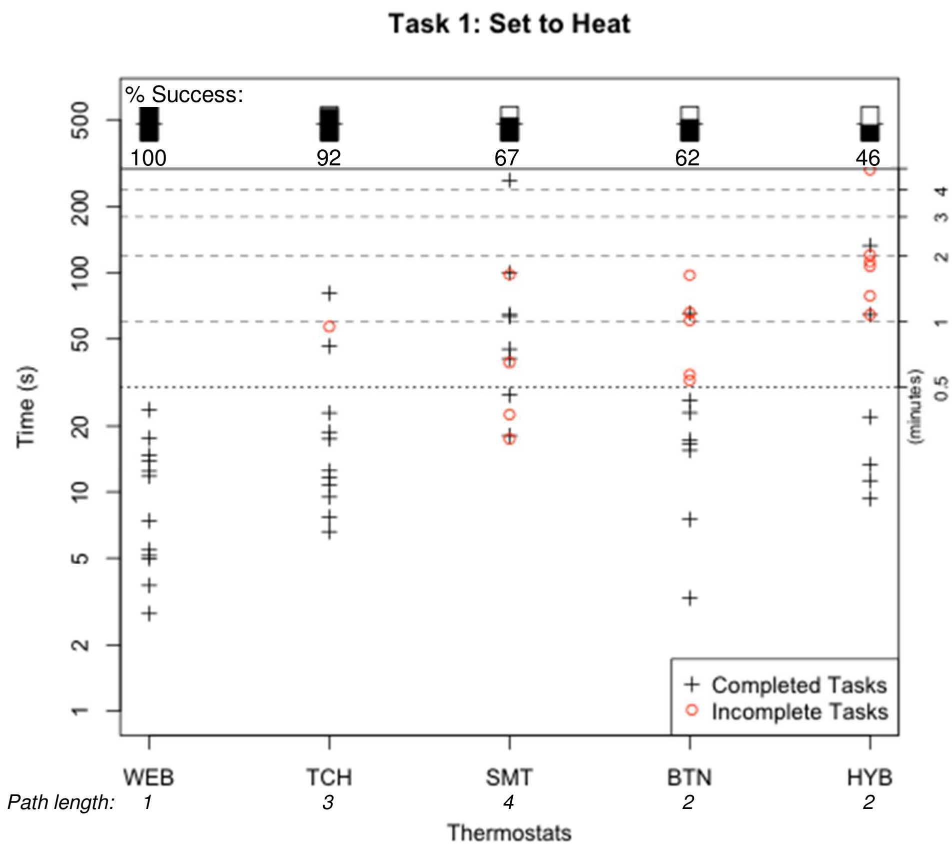

Figure 6: Time on task, success rate, and ideal path length for Task 1: Set to Heat.

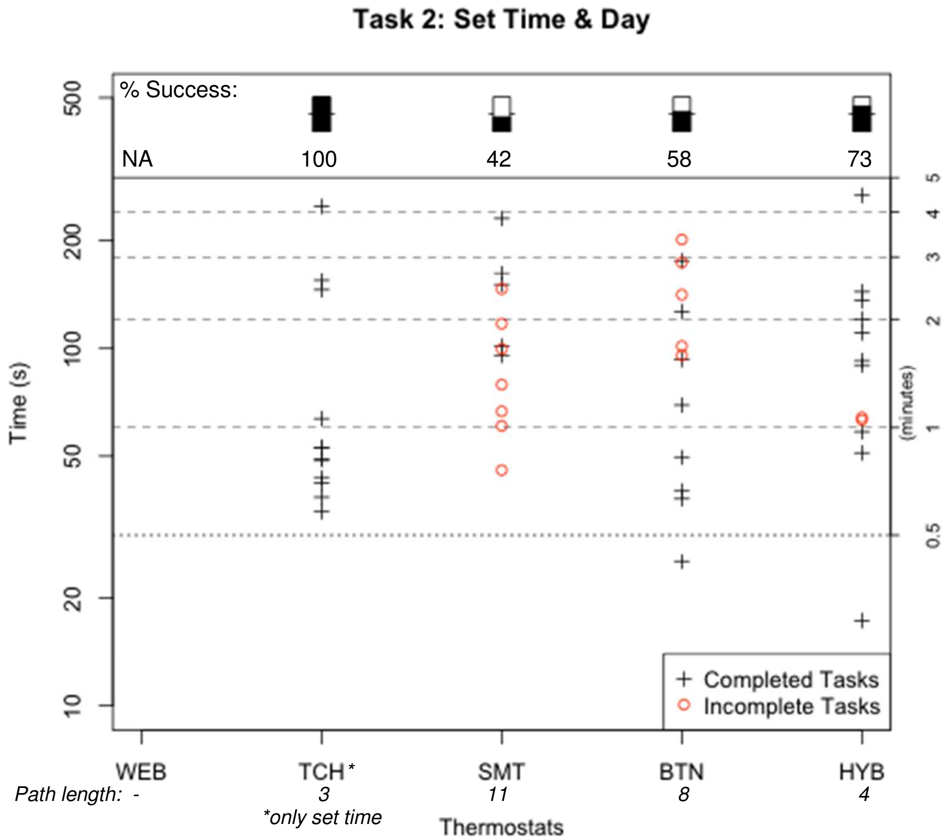

Figure 7: Time on task, success rate, and ideal path length for Task 2: Set Time & Day.

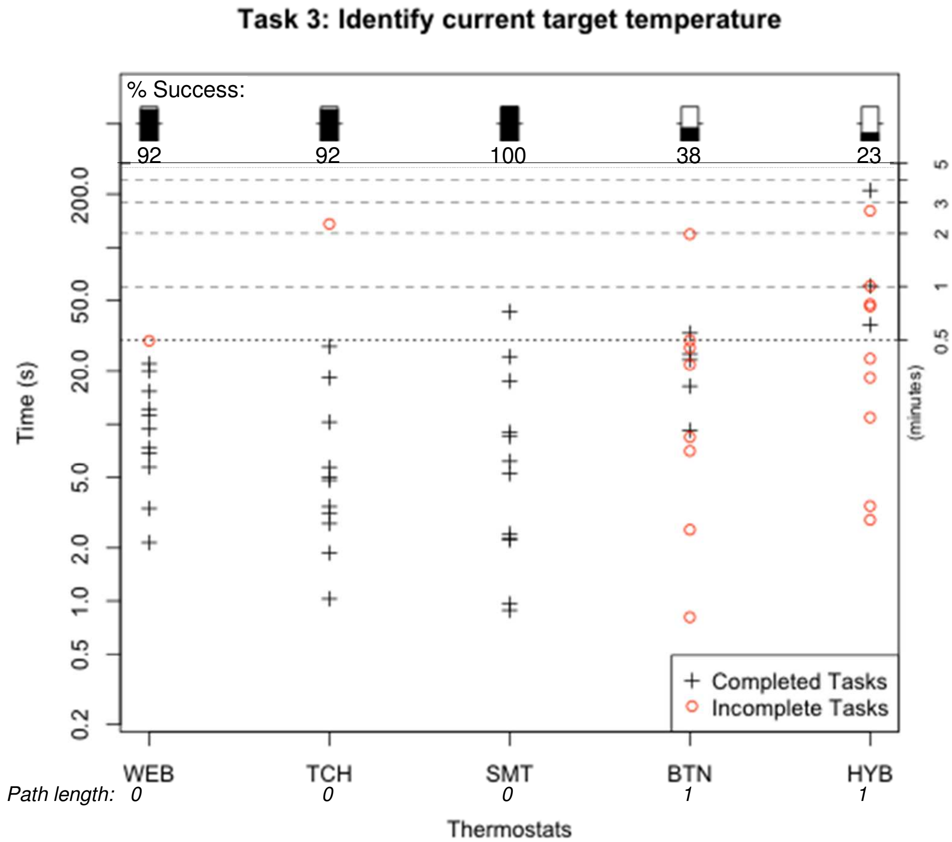

Figure 8: Time on task, success rate, and ideal path length for Task 3: Identify current target temperature.

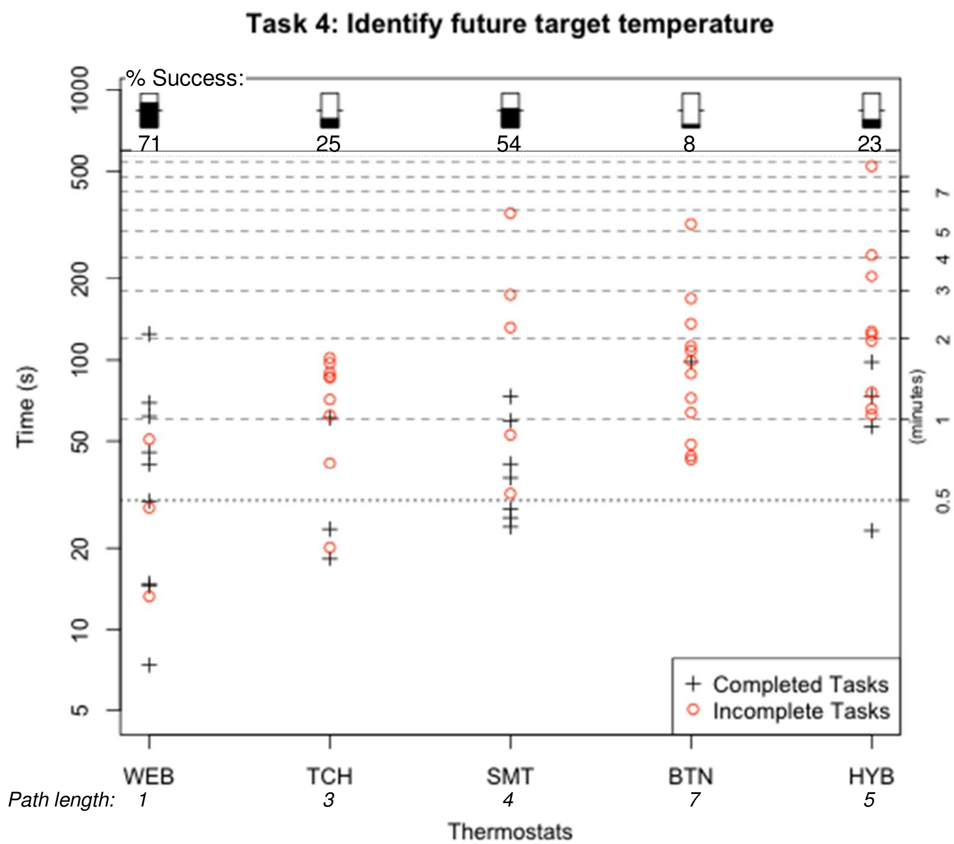

Figure 9: Time on task, success rate, and ideal path length for Task 4: Identify future target temperature.

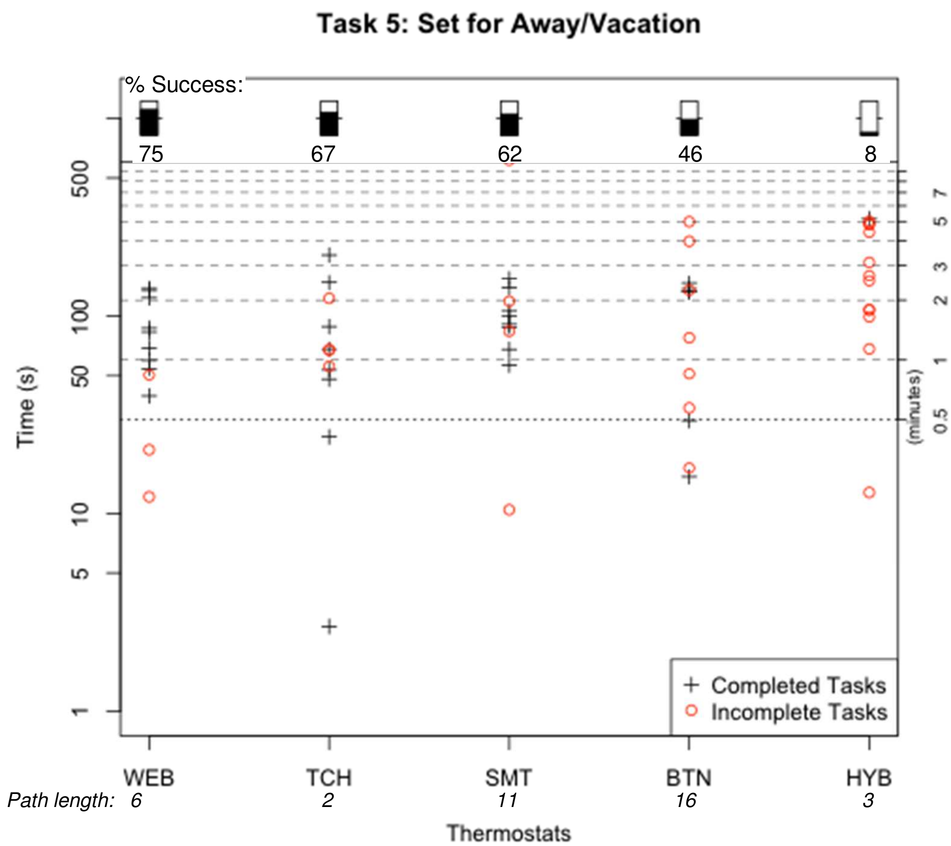

Figure 10: Time on task, success rate, and ideal path length for Task 5: Set for Away/Vacation.