January 2016

Sacramento Municipal Utility District SMUD's 2015 Smart Thermostat Usability Study

Karen Herter, Ph.D.
Herter Energy Research Solutions

Yevgeniya Okuneva, Statistician
Herter Energy Research Solutions

Powering forward. Together.

SMUD®

SMUD's 2015 Smart Thermostat Usability Study

A simultaneous, multi-user, paired comparison test of communicating thermostats for task efficiency, preference, and perceived usefulness of advanced features

January 2016

Herter Energy
RESEARCH SOLUTIONS

Prepared by: Herter Energy Research Solutions, Inc.
2201 Francisco Drive, Suite 140-120
El Dorado Hills, California
www.HerterEnergy.com

Authors: Karen Herter, Ph.D.
Yevgeniya Okuneva, Statistician

Prepared for: Sacramento Municipal Utility District
Sacramento, California

Senior Program Manager: Lupe Jimenez

Project Manager: Vikki Wood

SMUD Contract No: 4500071792

Suggested Citation:
Herter, Karen, and Yevgeniya Okuneva. *2015 Communicating Thermostat Usability Study*. Prepared by Herter Energy Research Solutions for the Sacramento Municipal Utility District. December 2015.

TABLE OF CONTENTS

List of Figures

List of Tables

EXECUTIVE SUMMARY

The main objective of this study was to better understand the features that contribute to ease of use and preference for communicating thermostats available for retail purchase in 2015 to enable procurement of thermostats for SMUD's energy efficiency and demand response programs. A secondary objective was test whether the standard System Usability Scale could be used in lieu of extensive lab-based usability testing as a time and cost-saving measure.

Time-on-task and survey data were collected during a 4-day lab study, during which 12 thermostats were tested by 155 participants. Each participant tested two different thermostats for 20 minutes each. Efforts were made to ensure a roughly even distribution of participants across thermostats by age and home ownership, the two demographic variables found to affect performance in SMUD's 2013 Smart Thermostat Usability Study. To prevent order bias, the sequence of thermostat pairs was reversed after each set of paired tests.

Based on an extensive literature review, and in accordance with the Common Industry Format (CIF) for usability[1], the following metrics were chosen to represent the usability of each thermostat tested in this study:

> ***Efficiency***, scored on a scale from 0 to 100, was calculated using the time required to complete a set of standard tasks for each thermostat. Thermostat features found to significantly affect Efficiency scores included a larger screen size and a smaller number of steps needed to change modes.
>
> ***Preference*** was calculated for each thermostat as the percentage of participants that would choose that thermostat over the other one tested for installation in their homes (assuming equal cost between the two thermostats tested). The only thermostat feature found to significantly affect Preference scores was a larger menu text size.
>
> ***Satisfaction*** ratings for Ease of Use, Sound and Feel, and Appearance were derived from participant responses to surveys questions.

Efficiency, Preference, and Satisfaction scores were averaged to create a grading system that ranked the Honeywell 9320, Venstar T7900 and Allure Eversense as the three most usable and likeable Wi-Fi thermostats tested in 2015 (Figure 1).

[1]ISO/IEC Standard 25062, 2006

FIGURE 1. TOP SCORING COMMUNICATING THERMOSTATS FOR 2015

Honeywell 9320

Venstar T7900

Allure Eversense

Each survey also included the 10-question System Usability Scale (SUS) survey, which SMUD wished to consider as a potential substitute for future usability testing, informing thermostat procurement at SMUD. The 12 thermostat SUS scores showed a 91% correlation with the final Usability Test scores calculated from the Efficiency, Preference, and Satisfaction scores, suggesting that the SUS could be a reasonable alternative to a full usability lab study.

TABLE 1. SMART THERMOSTAT USABILITY RESULTS SUMMARY

Rank	Thermostat	Number of Tests	Average Thermostat SUS	Usability Test Score	Grade
1	Honeywell 9320	26	73	74%	A
2	Venstar T7900	26	81	71%	A
3	Allure Eversense	27	62	64%	B+
4	Ecobee 3	26	51	60%	B
5	Emerson Sensi	26	61	59%	B
6	Schneider Wiser Air	26	65	59%	B
7	Carrier Cor	28	58	59%	B
8	Nest 3rd generation	26	50	57%	B-
9	Trane XL-824	25	48	51%	C+
10	Lux GEO	25	52	51%	C+
11	Honeywell Lyric	24	45	50%	C
12	Radiostat CT-80	25	43	49%	C

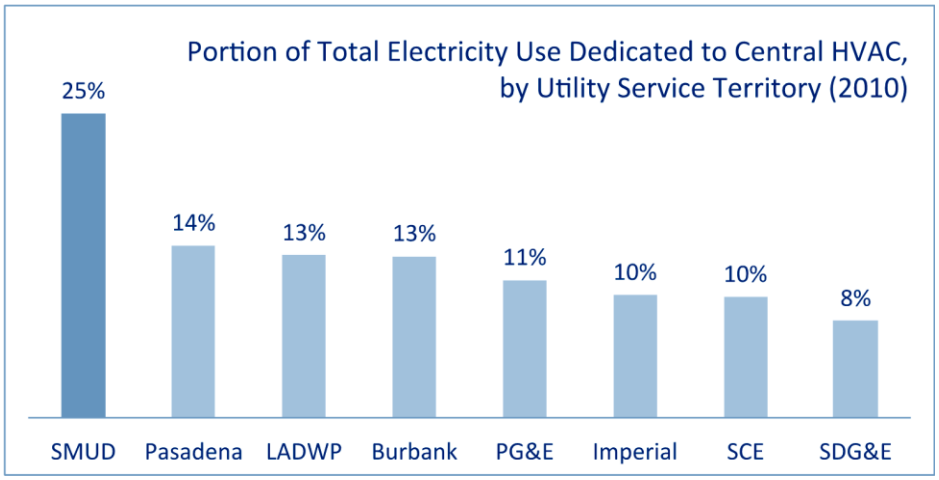
Based on these and other findings of this study, the research team recommends that SMUD establish as a prerequisite for future thermostat purchases a minimum average SUS score of 50, which corresponds to at least a B grade. The average SUS score should be the average of at least 22 SUS surveys based on a power analysis using the data collected in this study.

1 BACKGROUND

Communicating thermostats, or "smart thermostats" are Internet-connected devices responsible for controlling heating, ventilation and air conditioning (HVAC) systems in homes and small commercial buildings. In addition to the standard functionality of programmable thermostats, smart thermostats allow the user to control their HVAC system remotely, using smartphones and other Internet-connected devices. Many smart thermostats have additional advanced features that take advantage of information available through the Internet – things like weather, user location, and home energy use.

Smart thermostats have recently become the focus of utility efficiency and demand response efforts because of their ability to automate, control and optimize HVAC systems – the largest residential end-use in SMUD's service territory. Overall, thermostats control one-quarter of all electricity consumption in the SMUD service territory (Figure 2), and on the hottest summer days, residential air-conditioning is responsible for about one-third of SMUD's total 3,000-megawatt peak demand.

FIGURE 2. RESIDENTIAL HVAC ELECTRICITY USE IN CALIFORNIA



Source: California Energy Commission, 2009.

In 1995, the U.S. Energy Star created standards for programmable thermostats, only to rescind them about a decade later, after several studies showed that Energy Star compliant thermostats failed to bring about energy savings (Gunshinan 2007). Many of these studies pointed to evidence that energy saving schedules were not being used properly or at all.

Since then, some newer thermostats have been designed to minimize user input. Field studies using these new self-programming thermostats show modest savings; however, one of several recent SMUD studies showed that self-programming thermostats might not be effective at saving energy in low-income homes (Herter and Okuneva 2014a, 2014d).

1.1 Smart Thermostat Field Research at SMUD

In 2012 and 2013, SMUD ran several pilots designed to investigate the potential for communicating thermostats to reduce energy and demand in the residential sector. Many of these pilots also included one or more complementary savings measures, including time-of-use rates with critical peak price events (TOU-CPP), utility control of thermostats, enhanced information, audits, and weatherization measures. An overview of the pilots and their measures are provided in Table 2.

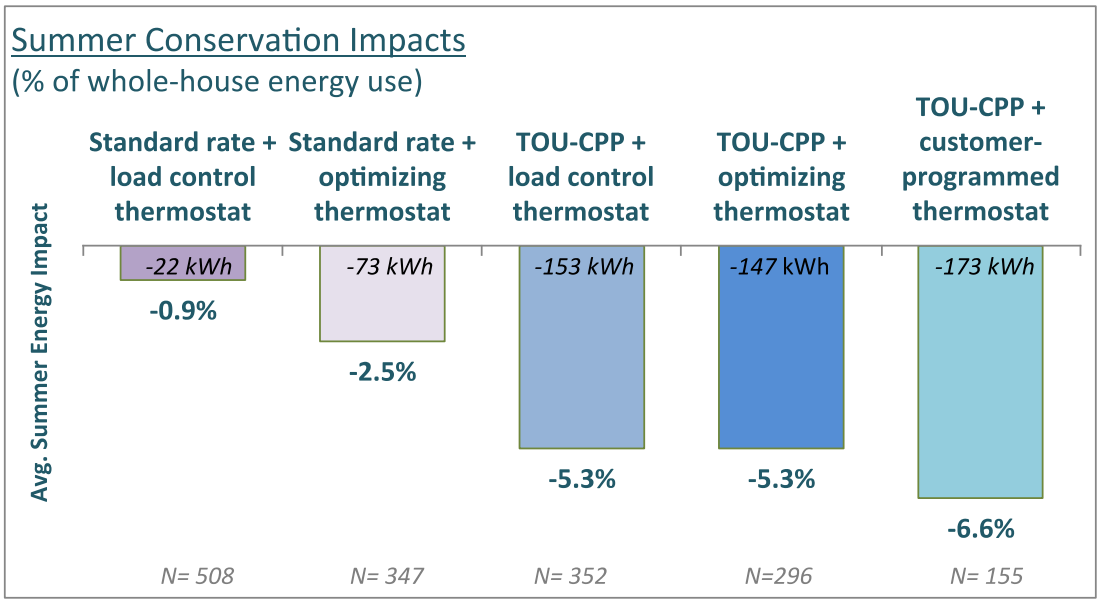
Table 2. Summary of recent smart thermostat field research at SMUD

			Thermostat Functions					
Pilot	Year	Participants	Utility Controlled Events	Customer Programmed Events	Self-Scheduling / Optimizing	TOU-CPP Rate	Real-time Energy Information	EE Audits / Measures
Single-Family Summer Solutions	2011-2012	310	✓	✓		✓	✓	✓
PowerStat Precooling	2012	180	✓					
Multi-Family Summer Solutions	2013	267		✓		✓	✓	
Smart Thermostats	2013	693			✓	✓		
Low-income Energy Management	2013	525			✓			✓

Overall, the smart thermostat pilots completed at SMUD indicated significant savings potential. Figure 3 shows that whole-house summer energy savings ranged from 1.0% to 7.9% under different rate and thermostat control scenarios.

Load control participants, who remained on the standard 2-tier rate and allowed SMUD to control their thermostats during demand response events, showed just 0.9% summer energy savings. Participants given self-programming or "optimizing" thermostats saved 2.5% of their summer energy use when they stayed on the standard 2-tier rate, while those who switched to a time-varying TOU-CPP rate more than doubled this savings to 5.3%. Across all of the SMUD pilots, the greatest savings (6.6%) came from customers who scheduled their own smart thermostats and programmed them to respond automatically to the TOU-CPP events. (Response to TOU-CPP events shown in Figure 4.)

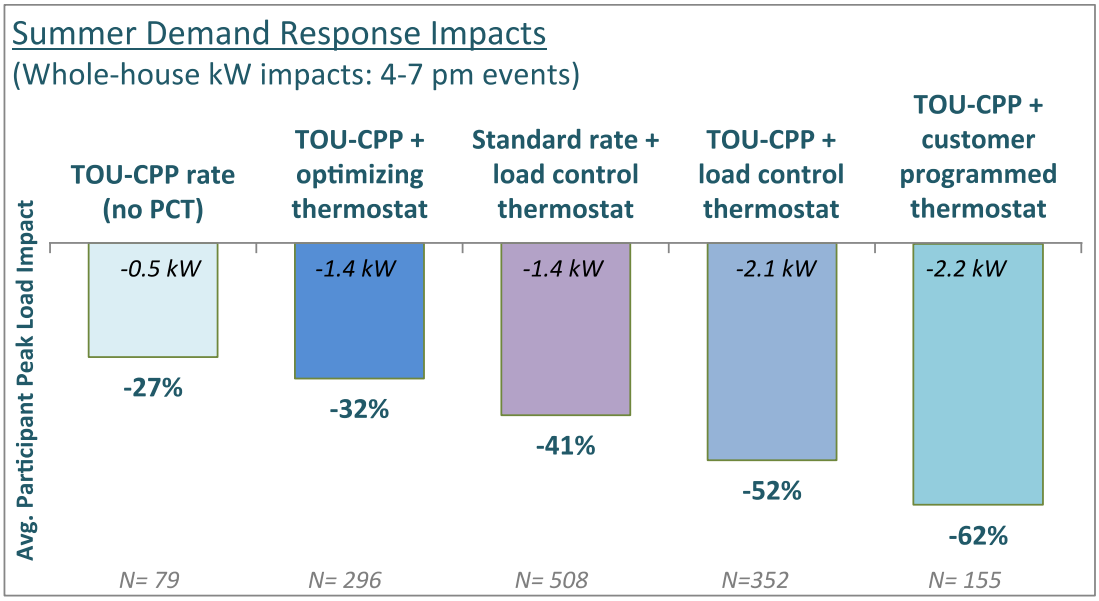
FIGURE 3. SUMMER CONSERVATION IMPACTS FOR SMART THERMOSTATS



Source: Herter and Okuneva 2014a

Figure 4 shows the load shed values for demand response events, which ranged from 27% to 62% (0.5 kW to 2.2 kW) under different rate and thermostat control scenarios. Demand response was lower for participants without smart thermostats, for those with optimizing thermostats that did not responded automatically to the TOU-CPP event signal, and for those on the standard rate with load control. Once again, savings were greatest for the group of participants who automated their own response to the TOU-CPP events.

FIGURE 4. SUMMER PEAK IMPACTS FOR SMART THERMOSTATS – EVENT DAYS



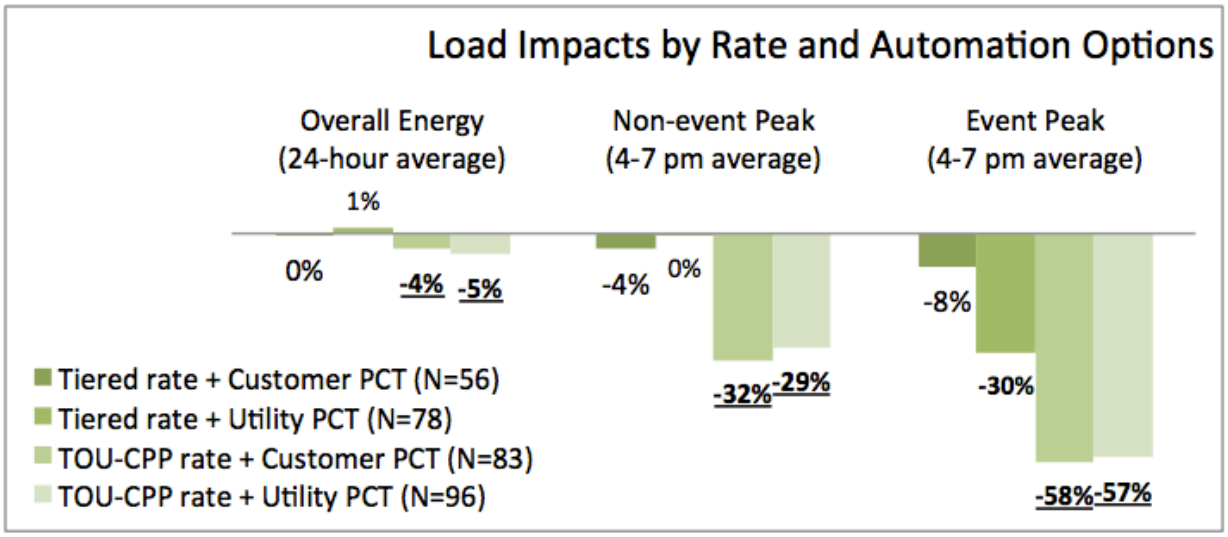
Source: Herter and Okuneva 2014a

2011-2012 Residential Summer Solutions Study

SMUD's Residential Summer Solutions Study was an integrated energy efficiency and demand response research pilot implemented in the summers of 2011 and 2012 in Folsom and downtown Sacramento – the two SMUD areas first outfitted with meters capable of measuring hourly energy use.[2] The study investigated the effects of real-time energy data and TOU-CPP rates on the energy use of customers that volunteered for smart thermostats that were either controlled by SMUD or programmed by customers to respond automatically to 12 summer CPP events (Herter & Okuneva 2014b).

Of the four rate and automation options offered, the TOU-CPP rate with customer-programmed automation provided the greatest savings, with 4% energy savings, daily weekday peak savings of more than 30%, and an average event peak load shed of nearly 60%. On average, customers who signed up for the dynamic TOU-CPP rate saved an average of $145 per summer, while those who chose to remain on the standard 2-tier rate saved $40 per summer

Figure 5. Rate and thermostat automation impacts are significant



*Impacts significantly different from the "Tiered rate + Customer PCT" group are marked in **bold**.*
Impacts significantly different from the "Tiered rate + Utility PCT" group are underlined.

[2] Folsom and downtown Sacramento are not representative of the SMUD population.

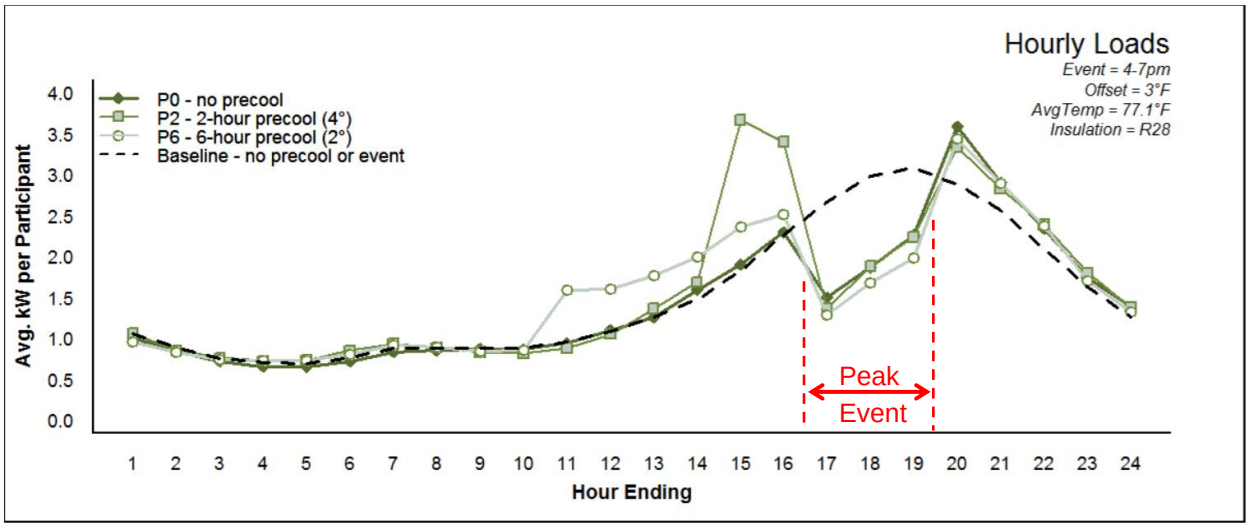
2012 Residential PowerStat Precooling Pilot

The Residential PowerStat Precooling Pilot solicited 180 residential participants with the primary objective of testing the effects of event-based precooling and peak temperature offsets on energy use, peak demand, and occupant comfort (Herter and Okuneva, 2013).

Three different precooling strategies or "treatments" were tested 8 times during the summer of 2012. Treatment 1, the base case, had no precooling before the peak period. Treatment 2 was a 2-hour, 4-degree precool before the peak period. Treatment 3 was a 6-hour, 2-degree precool before the peak period.

Participant survey responses and interval meter data were collected to enable comparison of the impacts of the three treatments to determine that the different precooling strategies had different effects on hourly load shapes, daily energy use, and occupant comfort. Overall, the 6-hour 2-degree precooling had the greatest peak period energy impacts of the three precooling strategies (Figure 6).

Figure 6. 6-hour 2-degree precool has the greatest peak impacts



** The average 3-hour peak impact for the P6 precool was -1.26 kW. The differences between the P6 impact and the impacts for P0 (-1.03 kW) and P2 (-1.08 kW) are statistically significant (α = 0.05).*

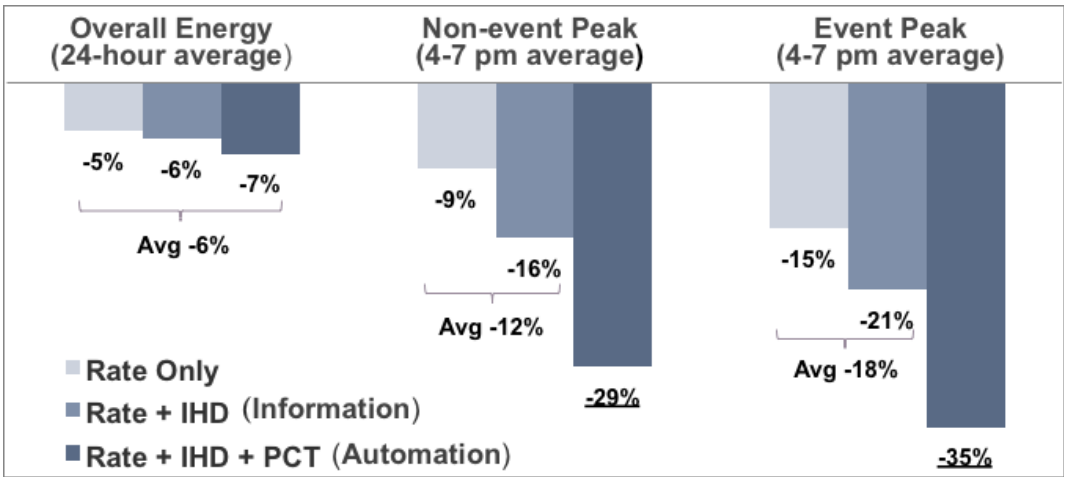
On average, the 6-hour, 2-degree precooling strategy provided statistically significant summer energy savings for participants with ceiling insulation of at least R38, suggesting that the adding insulation to participant attics could not only improve the effectiveness of peak load reduction programs, but also lower summer energy use and customer bills.

2013 Multifamily Summer Solutions

This study considered the effects on multifamily homes of three energy and demand reduction measures: an experimental TOU-CPP rate, an Energy Aware Power Tab energy display (IHD) that provided real-time whole-house kWh, and an Energate Pioneer Z100 programmable communicating thermostat (PCT) that was capable of automatically adjusting target temperatures in response to price signals from the utility (Sutter et al., 2014).

All three measures resulted in statistically significant energy and demand response savings compared to the standard two-tiered flat rate. Energy savings ranged from 5% to 7%. The PCT automation group had peak load savings of 29% and 35% for non-event and event days, respectively — more than double the savings of the rate and information groups (Figure 7).

Figure 7. Automation doubles peak savings in multifamily dwellings

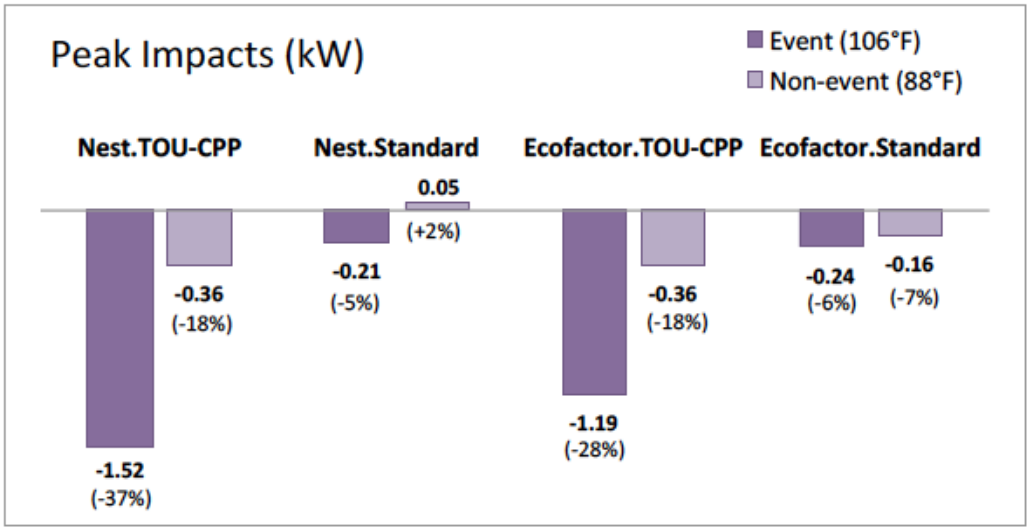


(Source: SMUD 2015) All savings are significantly different from standard rate at p≤ 0.05 Underlined values are significantly different from other groups.

2013 Smart Thermostat Pilot

SMUD's Smart Thermostat Pilot investigated the energy and demand impacts of the Nest thermostat and Ecofactor system under two rates: a standard 2-tier rate and a TOU-CPP rate. Results showed that both thermostats provided statistically significant annual whole-house energy savings of between 1.6% and 3.3%. In addition, those on the TOU-CPP rate showed substantial and statistically significant savings during the 4 to 7 pm peak period on event and non-event days, as shown in Figure 8 (Herter & Okuneva 2014c).

Figure 8. Nest and Ecofactor peak impacts under tiered and TOU-CPP rates



Note: All load impact values are statistically significant. All differences between load impact values on like days are statistically significant with the following exceptions: (1) Nest.Standard vs. Ecofactor.Standard on event days, and (2) Nest.TOU-CPP vs. Ecofactor.TOU-CPP on non-event days.

2013 Low-income Weatherization and Energy Management Pilot

SMUD's Low-income Weatherization and Energy Management Pilot measured the energy and demand impacts of the Nest thermostat in the homes of low-income weatherization participants. In contrast to the savings found in the Smart Thermostat Pilot, which excluded participation by low-income customers, results showed that low-income participants with the Nest used 7.1% more energy annually, and had a 5.4% higher peak summer demand than a control group made up of similar low-income weatherization participants without the Nest (Herter & Okuneva 2014d).

1.2 THERMOSTAT USABILITY RESEARCH AT SMUD

The need for a method to identify user-friendly thermostats for procurement purposes has motivated SMUD to conduct thermostat usability studies. The first usability study was conducted in 2013 (Herter and Okuneva 2014e). The second is described in this report.

The National Institute of Standards and Technology (NIST), the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) have developed the Common Industry Format (CIF)[3] for reporting the results of summative usability tests such as this one "to facilitate incorporation of usability as part of the procurement decision-making process for interactive software products." The standard does not specify how to conduct usability tests; rather it specifies the types of data to collect in usability testing and how the test results should be reported to ensure they mean the same thing from product to product. The intent was for vendors to commission independent agents to conduct studies so prospective buyers could compare the usability of information technologies and products on the market.

Thermostat vendors have not adopted the practice of conducting publicly available usability tests, possibly because their primary customers are retail and not institutional. SMUD's thermostat usability studies adhere to the CIF to allow test results to be used in SMUD's thermostat sourcing process, and in the hope that thermostat manufacturers will be able to apply the results to improve their products' usability.

In SMUD's first thermostat usability study, conducted in 2013, ten communicating thermostats and two non-communicating thermostats were tested by a total of 163 participants, who were evenly distributed by age and education. A focus group facility was used for testing, to record individual thermostat interactions, collect participant surveys, and conduct focus groups.

The study found that Preference scores were similar across participants of differing age, gender, education, income, home ownership, and technology IQ. Preference scores were significantly higher for thermostats with color displays and high "Overall Feel and Sound" ratings. Efficiency scores were significantly higher for thermostats with larger screens and higher ratings for "Overall Ease of Use" ratings. Efficiency was also influenced by the user characteristics, with younger users and homeowners having significantly higher Efficiency scores.

[3]After several years of development under NIST, the CIF became an American National Standards Institute (ANSI) standard in 2001, and was approved as ISO/IEC Standard 25062 in 2006. Subsequently, ISO and IES have expanded the CIF under ISO/IEC TR 25060 (2010), ISO/IEC 25064 (2013) and ISO/IEC 25063 (2014).

2 Literature Review

The National Institute of Standards and Technology (NIST) defines usability as "The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use." (NIST 2001)

Some of the basic elements of usability testing are as follows (Rubin and Chisnell 2008):

- Development of research questions rather than hypotheses
- Use of a representative sample of end users, which may or may not be randomly chosen
- Representation of the actual environment
- Observation of end users who either use or review a representation of the product
- Controlled interviewing and probing of the participants by the test facilitator
- Collection of quantitative and qualitative performance and preference measures
- Recommendation of improvements to the design of the product

This study made use of a specific implementation of usability test called a *comparison test*, where the purpose is to explicitly compare two or more products. The basic methodology involves side-by-side comparison of two or more different product designs. Performance and preference data are collected for each product and the results are compared. Comparison tests are typically used to establish which design is easier to use or learn, and to better understand the advantages and disadvantages of different designs (Rubin and Chisnell 2008).

The unique contribution of user testing is that it exposes what people actually *do*, as opposed to what they *say* – or say they will do. Paired-comparison testing was chosen for this study to elicit more honest critical feedback than would single-unit testing. With paired comparison testing, users are forced to choose one product over the other, and are given the opportunity to explain in more detail precisely what they liked about one product in relation to what was *not as good* about the other (Enerson 2012).

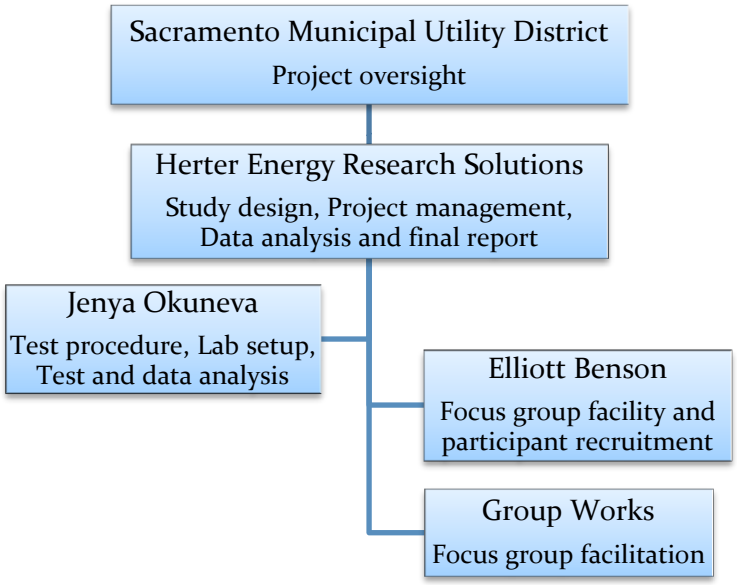
3 Approach

3.1 Research Team

Figure 9 illustrates the organization of the research team for this project. The client team is led by Vikki Wood, Principal Demand-side Specialist at the Sacramento Municipal Utility District, a large customer-owned electric utility located in California's central valley in and around Sacramento County.

The consultant team is led by Karen Herter, President of Herter Energy Research Solutions, who has worked in the energy industry for 20 years, investigating and reporting on a broad range of topics including utility-scale wind energy potential, energy efficiency, dynamic pricing, demand response, utility-customer communications, smart grid, interval meter data, consumer behavior, and home automation.

Figure 9. Research team organizational chart



Herter Energy Research Solutions was supported by three additional firms to carry out the 2015 smart thermostat usability study. Jenya Okuneva is an independent statistical consultant that has been conducting statistical evaluation of hourly energy data for electric utilities since 2012. Group Works is a market research firm that has been providing consulting services to a wide range of industries for over 30 years. Elliott Benson is a focus group facility located in downtown Sacramento.

3.2 Study Overview

This document describes the implementation and evaluation of a simultaneous multi-user, multi-device comparison test of a sample of thermostats available for purchase in 2015.

Goals and Objectives

The primary goal of this study was to assess the features and functions of a variety of communicating thermostats to determine which characteristics might be recommended or required in specifications for thermostats promoted by or implemented for future programs at SMUD. The objectives of the study were to:

- Calculate **Efficiency** scores from time-on-task measurements of how quickly users perform a given set of common tasks
- Calculate **Preference** scores as the frequency that each thermostat was preferred over the others as determined through surveys
- Determine **Satisfaction** for individual features based on participant ratings collected through surveys
- Determine **System Usability Scale** (SUS) scores and compare to metrics for Efficiency, Preference, and Satisfaction.
- Rate the **Interest** level for advanced features, based on participant surveys

Research questions

- How do Efficiency, Preference, and Satisfaction metrics compare between products?
- What features are most helpful to users in completing common tasks?
- What features are common in thermostats most frequently preferred by users?
- How do participants rate the advanced features they reviewed?

Logistics

- Test 12 advanced communicating thermostats under controlled lab conditions
    - Devise a list of common tasks for each participant to perform
    - Video record participant attempts to complete the task list
    - Have participants fill out paper surveys at test stations
    - Conduct discussion sessions to gain further qualitative insights
- Data input
    - Record survey responses
    - Review the video recording of tasks
        - Record done or not done for each task (success)
        - Record start and end times for each task (time-on-task)
- Data Analysis
    - Calculate thermostat Efficiency as a function of success and time-on-task
    - Calculate average Satisfaction scores from surveyed feature ratings
    - Calculate System Usability Scale ratings and compare to other metrics
    - Regress thermostat features, participant characteristics, and satisfaction ratings on dependent variables Efficiency and Preference

This methodology is considered to be the most appropriate for this application for the following reasons:

- The lab study allows for a relatively large sample size (good statistical power) completed quickly and at a reasonable cost relative to a field study.
- The mixed methods approach uses quantitative survey ratings and efficiency scores combined with quantitative open-ended answers and focus group discussions to provide a thorough and well-rounded study.
- The metrics can readily be transformed into procurement requirements: the regression of efficiency and preference scores on thermostat characteristics will allow for statistically valid determination of important features to look for in procurement.

SCHEDULE

Month	Task
October 2015	Literature review
	Procure equipment
	Draft scripts, surveys and other documentation
November 2015	Recruit test participants
	Usability test
	Input survey data
	Review video for time-on-task and success rates
	Data input and analysis
	Draft Report
December 2015	Final Report
	Final Presentation at SMUD

3.3 Roles During Testing

Following are the roles of the people involved during the three days of lab testing.

Principal Investigator

The principal investigator was present for the entirety of the testing to:

- Observe testing and take notes
- Direct facilitator and technical support staff as needed
- Address issues that could not be resolved by the technical support staff or facilitator
- Answer viewer questions

Facilitator

The facilitator was present for the entirety of each test session to:

- Provide directions to participants
- Indicate start and stop times for testing
- Conduct group discussion sessions
- Respond to non-technical requests for assistance

Technical Support Staff

The technical support person was present for the entirety of each test session to:

- Monitor recording equipment
- Resolve technical problems with thermostats or other equipment

Participants

For each of the two thermostats tested, the participant's role was to:

- Attempt to complete a set of representative task scenarios as efficiently as possible
- Fill out a post-test questionnaire
- Participate in a discussion session to provide honest opinions regarding the usability and likability of the thermostats and supporting applications

Facility Staff

Facility staff members were present for the entirety of the testing to:

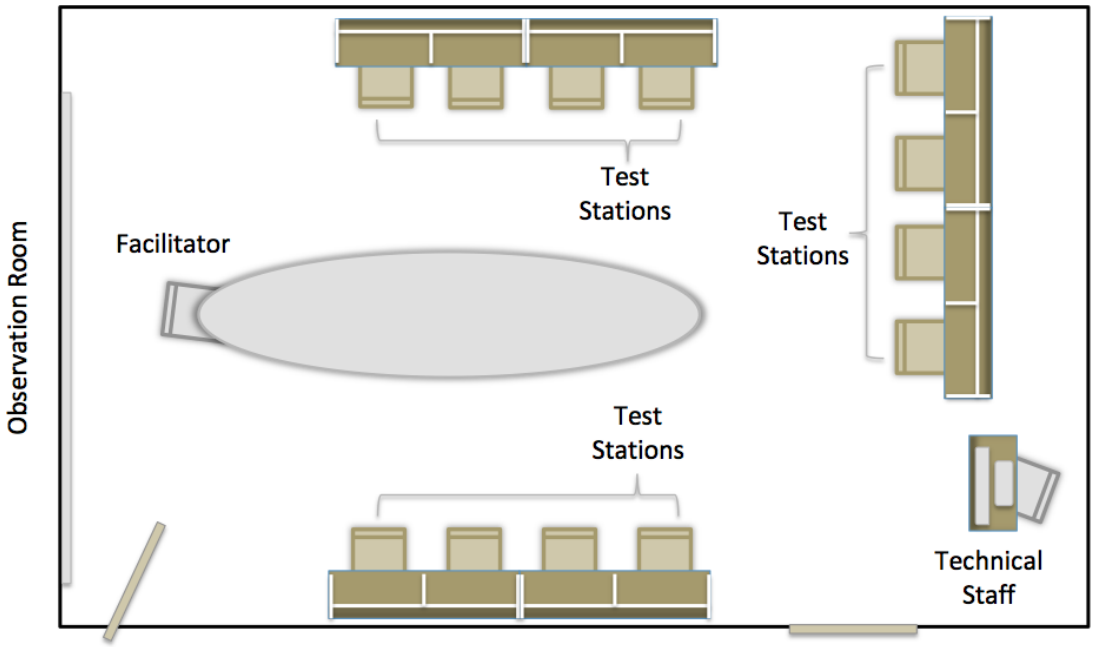
- Greet and direct participants as needed
- Provide assistance with internet connectivity issues
- Provide miscellaneous items that aided in testing

3.4 Test Lab and Equipment

The usability testing took place at a facility within the SMUD service territory. The test lab was equipped with multiple test stations affording each participant some privacy. A thermostat and its supporting smartphone application were mounted in each cubicle, in view of the video camera. Video of participant faces was not recorded.

During testing, the facilitator and one technical support staff were seated in the same room as the participants, while observers monitored the sessions in the observation room.

Figure 10. Test room layout



Video and Audio Recording

Small cameras were affixed to each cubicle to record user interactions with the thermostats and supporting applications. One video was recorded for each of the 310 tests for later review. The facility also provided audio recordings of all group discussions that took place at the large table in the center of the room, before, between, and after the testing of the thermostats at the test stations.

Web Access

All of the thermostats tested made use of WiFi Internet connectivity for remote control by the associated smartphone application.

THERMOSTATS

A total of 12 Wi-Fi thermostats were selected for testing as shown in Table 3.

TABLE 3. WIFI THERMOSTATS TESTED – ALL UNITS HAVE WIFI AND SMARTPHONE APP

ID	Device	Display	Input	Price	Image
1	Nest 3rd generation	Color, graphical	Push-dial	$249	
2	Ecobee 3	Color, graphical	Touchscreen	$249	
3	Honeywell Lyric	Color, graphical	Dial + Buttons	$175	
4	Honeywell RTH9320	Color, graphical	Touchscreen	$175	
5	Emerson Sensi	Mono, text	Buttons	$125	
6	Venstar T7900	Color, graphical	Touchscreen	$195	
7	Trane XL824	Color, graphical	Touchscreen	$395	
8	Lux GEO	Mono, text	Push-dial	$150	
9	Schneider Wiser Air	Color, graphical	Touchscreen	$239	
10	Carrier Cor	Color, graphical	Touchscreen	$249	
11	Allure Eversense	Color, graphical	Touchscreen	$249	
12	Radiostat CT-80	Mono, text	Touchscreen + Buttons	$200	

3.5 Participant Sample

A total of 180 residential SMUD customers were recruited for participation in the thermostat usability study based on a sample size power analysis showing the need for a minimum of 24 participant tests per thermostat (see Appendix C). Using the recruitment script provided in Appendix D, about 30 participants were recruited for each of 12 cells defined by 6 age categories. Of the 180 recruited participants, 155 attended the usability testing, as summarized in Table 4.

Table 4. Number of participants that attended the testing, by age and education

Year of Birth	Age	Total
1986 - 1995	20 - 29	**18**
1976 - 1985	30 - 39	**26**
1966 - 1975	40 - 49	**31**
1956 - 1965	50 - 59	**26**
1946 - 1955	60 - 69	**24**
pre - 1945	70 +	**30**
Total		**155**

SMUD's 2013 usability study indicated that gender, income, education and confidence using a thermostat were not significant indicators of efficiency or preference for thermostats. The same study showed that younger participants and homeowners scored significantly higher on Efficiency ratings than did older participants and renters. Building on this evidence, the 2015 study assigned participants to thermostat pairs in a way that ensured roughly even distribution of age (Figure 11) and homeownership (Figure 12).

Figure 11. Age distribution

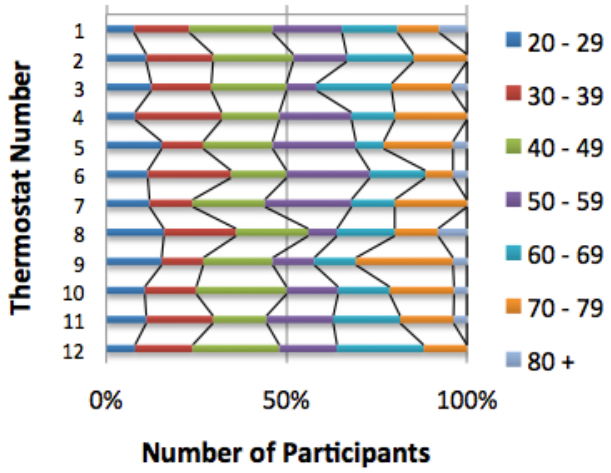
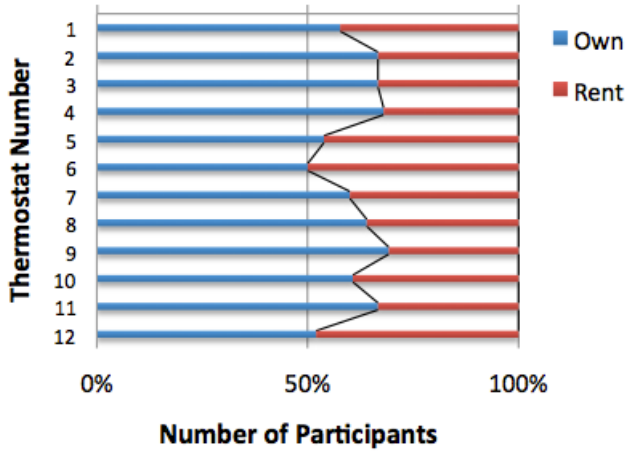


FIGURE 12. HOMEOWNERSHIP DISTRIBUTION



3.6 THERMOSTAT ASSIGNMENT

The 155 participants each evaluated 2 thermostats for a total of 310 individual tests. To avoid order bias, each thermostat was the *first* unit tested in roughly half of the tests, and the *second* unit tested in the remaining tests. All 66 potential thermostat pairs were tested at least once, and 119 of the 132 ordered pairs were tested at least once.

Table 5 shows the final count of participants that tested each ordered pair of thermostats. The first unit tested is listed across the top as column headers, and the second thermostat tested is listed along the left side as row headers. For example, 2 participants tested thermostat 1 followed by thermostat 2, and 1 participant tested thermostat 2 followed by thermostat 1.

TABLE 5. PARTICIPANT-THERMOSTAT ASSIGNMENTS: FIRST AND SECOND UNITS TESTED

First Unit → / Second Unit ↓	1	2	3	4	5	6	7	8	9	10	11	12	Total
1		1	1	1	1	1	1	1	2	1	2	2	**14**
2	2		1	1	1	1	1	1	1	2	1	2	**14**
3	2	2		0	1	1	0	1	0	1	2	2	**12**
4	1	2	1		1	0	1	1	1	1	1	2	**12**
5	0	2	2	2		1	1	1	0	1	1	1	**12**
6	1	1	1	2	2		1	1	1	1	1	1	**13**
7	1	1	1	2	2	2		0	1	0	1	0	**11**
8	1	1	1	2	1	1	2		1	1	1	1	**13**
9	1	1	1	0	2	2	2	1		2	1	0	**13**
10	1	1	1	1	1	2	2	2	2		0	1	**14**
11	1	1	1	1	1	1	2	2	2	2		1	**15**
12	1	0	1	1	1	1	1	1	2	2	1		**12**
Total	**12**	**13**	**12**	**13**	**14**	**13**	**14**	**12**	**13**	**14**	**12**	**13**	**155**

3.7 PROCEDURE

A total of 15 sessions were conducted over the span of four days. Each session accommodated up to twelve participants and took 90 minutes, roughly following the schedule shown in Table 6.

TABLE 6. AGENDA FOR EACH SESSION

	Segment	Minutes
1	Introduction	15
2	Thermostat test #1	20
3	Group discussion #1	15
4	Thermostat test #2	20
5	Group discussion #2	15
6	Wrap up and goodbye	5
	Total	90

INTRODUCTION

The facilitator briefed participants on the usability test procedure, including:

- the purpose of the study
- the importance of their involvement
- the facilitator's role
- the room configuration, recording systems, observers, etc.
- the testing protocol and agenda

THERMOSTAT TESTING

The facilitator briefed the participants on the testing process, stressing that the thermostats – not the participants – were being evaluated. The facilitator explained that the amount of time taken to complete each task was measured, and that exploratory behavior outside the task flow should not occur until after completion of all tasks. Participants were given 20 minutes to complete the entire task list and fill out the survey. (See Facilitator's Guide, Appendix. E)

The first task of each test was conducted to enable comparison of the short-term or "walk-up" usability of the thermostats, meaning that participants were not provided with user manuals or coached in any way prior to their interaction with the thermostats. At the end of this 30-second pre-test, participants who could not finish the task were shown how to complete it.

Given the relatively short test time of 20 minutes per thermostat, the tasks were designed to be the most common of available functions (see Table 7). These common tasks were identical for all thermostats, with minor variations every other test to limit the need for lab staff to reset thermostats to default settings after each test. To accomplish this alternating test pattern, every thermostat was assigned one Booklet A and one Booklet B. After each test, during discussion sessions, lab staff swapped from A to B or from B to A, and thermostat settings were adjusted for tasks that were not successfully completed.

The test commenced when users were told by the facilitator to begin. Participants were then directed to flip to the first card and begin task 1. The task ended when the participant marked that they did or did not complete the task on their task checklist. The next task began when they flipped over the next task card, and so on. Participants were directed to contact technical support staff immediately should any of the equipment fail to operate during testing.

FIGURE 13. PHOTOGRAPH OF TESTING IN PROGRESS

THERMOSTAT SURVEYS

Participants were given one survey for each of the two thermostats they tested. Surveys used 5-point Likert scales to rate multiple factors contributing to three categories of features for both the thermostat and the app: Ease of Use, Feel and Sound, and Appearance.

The standard 10-question System Usability Scale was also included in each thermostat survey. The SUS questions on the surveys distributed in the first 8 (of 15) test sessions referred to the system as "the thermostat." After the first 8 sessions, the survey was

changed to refer to the system as "the thermostat/app system" to limit confusion about whether participants should be rating the thermostat only, or the entire system.

After the second thermostat survey, an additional set of questions was provided to collect (1) information on which of the two thermostats was preferred, and (2) 5-point ratings for the desirability of a list of advanced thermostat features.

Group Discussions

Short focus group discussion took place after each thermostat test. Conducting a group discussion between the two thermostat evaluations ran the risk of providing some information to participants about the thermostat tested in round two. Despite this possibility, a discussion was conducted between the two tests for the following reasons:

- Each thermostat was tested as the first and second unit an equal number of times. Any bias inherent in the second test resulting from discussion between the two tests, or increased familiarity with the process and tasks in the second test, was evenly distributed across thermostats.
- Richer, more relevant feedback was made possible by having two discussions; i.e. the experience with the first unit may have been lost if the two tests were contiguous.
- Lab staff needed that time for logistic purposes – checking to see that the thermostats were ready for the second test.

4 Data Analysis and Results

For this comparison study, qualitative data, satisfaction ratings, task efficiency metrics, and preference metrics were collected for use in comparing between products and features.

4.1 Timed Tasks

Two metrics were determined from review of the thermostat test video recordings. The first task was devised to uncover issues with unusual input mechanisms. The remaining tasks were combined to determine an overall Efficiency score as described below.

Input Mechanism Discovery

An important feature of this study is the fact that participants were not provided with instructions on how to use the thermostat or smartphone application prior to testing. The reason for this is that thermostats are generally acquired with a newly purchased home rather than purchased outright. As a result, new occupants are often required to interact with their inherited thermostat without instructions.

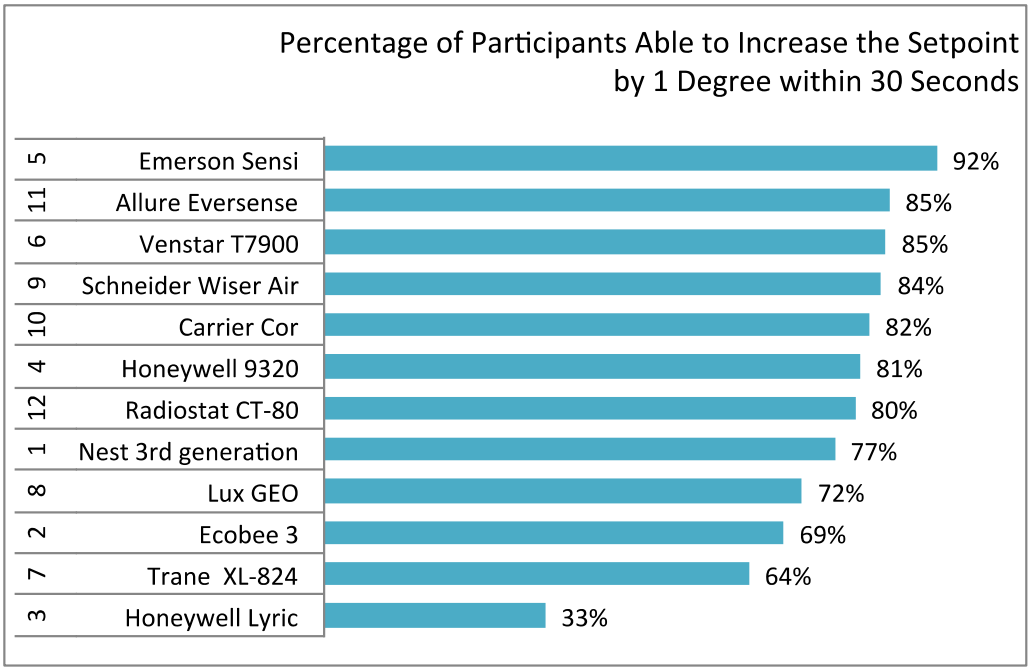
In SMUD's first thermostat usability study, it took an unusually long time for many participants to figure out how to interact with the Nest thermostat, which at the time was relatively new on the market and had a very different design and input mechanism from all other thermostats at the time. In the first study, about half of participants that tested the Nest were unable to figure out the input mechanism during the 20-minute test period. Participants swiped and tapped the glass screen, but did not discover that the entire face of the Nest had to be pushed and turned for input. As a result, a majority of the Nest tasks received an Efficiency score of zero, and frustrated Nest testers scored the unit poorly in the surveys (Herter Energy Research Solutions 2014).

While acknowledging that input mechanism discovery is an important part of walk-up usability, the research team decided to separate the input discovery from the task completion because users discover the input mechanism just once – the very first time the unit is used. In contrast, tasks like modification of target temperatures and setpoint schedules are expected to occur hundreds or thousands of times. To avoid unfairly weighting an action that in the real world constitutes an inconsequential portion of interaction time, this study tested the ability for participants to discover the input mechanism on their own, but did not include success rates for this task in the overall thermostat Efficiency scores.

Based on this experience, this second thermostat usability study incorporated a simple pre-test to measure participants' ability to discover the input mechanism: participants were given 30 seconds to increase the current target temperature by 1 degree. After the pre-test,

participants who were unable to complete the task were shown the correct method by lab staff, thus ensuring that further testing would not be inhibited by a lack of understanding about the input mechanism. Input mechanism discovery success rates are provided separately in Figure 14.

FIGURE 14. INPUT MECHANISM DISCOVERY



EFFICIENCY: COMPLETION TIME FOR COMMON THERMOSTAT INTERACTIONS

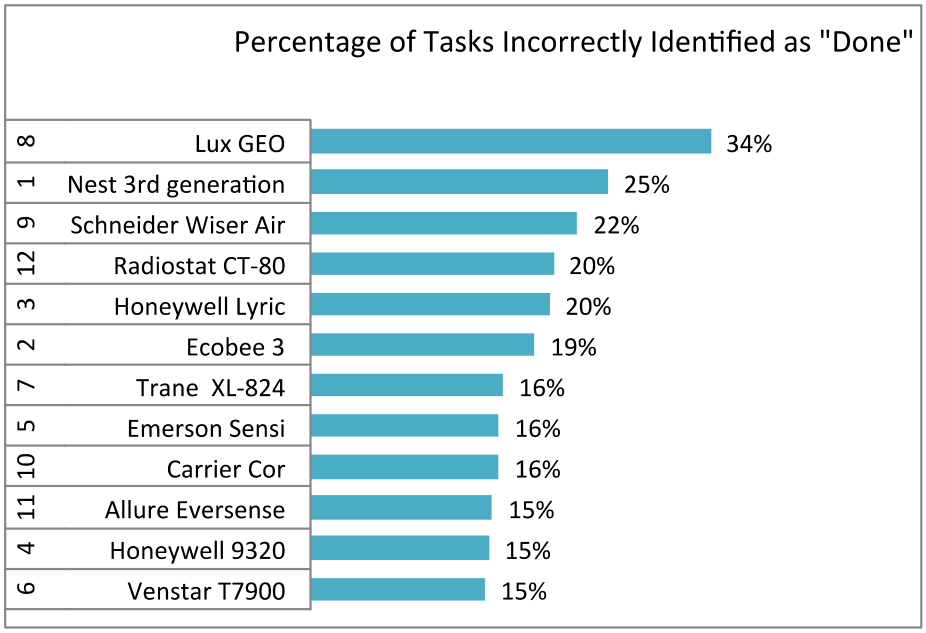
All 310 video recordings of the usability tests were reviewed to capture Time-on-task and Success measures for each thermostat, participant, and completed task. These measures were then used to calculate the individual task and average thermostat Efficiency metrics for each thermostat.

For all tasks marked "Done" on the task checklist, the task start time was recorded at the moment the numbered task card became visible by the camera, and the end time was recorded as the participant marked their checklist. In some cases, participants did not complete the task on the first try, marked their checklist Not Done, then returned to and completed the same task later in the session. In these cases, the two times were added together to obtain the total Time-on-task metric.

Successful completion of the task was determined using the video recording. Tasks successfully completed received a Success score of 1, while those that were not successfully completed received a Success score of 0, even if the participant marked that task on their

checklist "Done." Figure 15 shows the percentage of tasks that were marked done on the checklist, but were never completed. It is not clear whether participants marked done because they felt they had completed the task, or whether they were indicating that they were done trying to complete the task. Note that these values are highly correlated with the Efficiency scores described below.

FIGURE 15. INCOMPLETE TASKS THAT WERE MARKED "DONE" ON THE CHECKLIST



EFFICIENCY METRIC

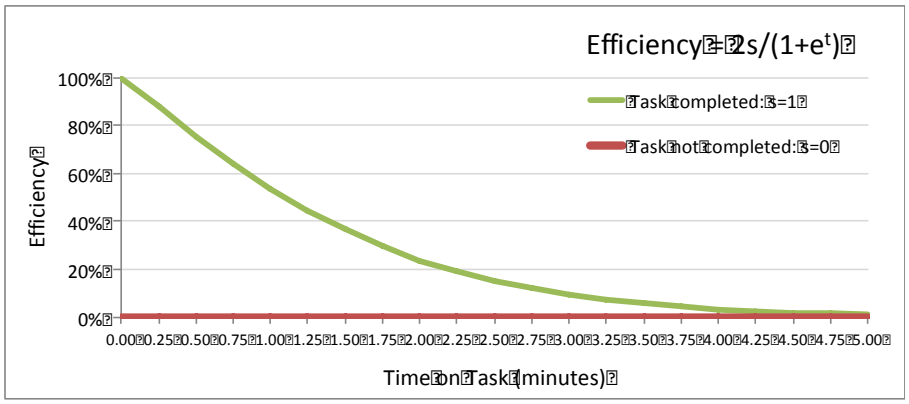
Together, the Time-on-task and Success metrics were used to calculate the *Efficiency* metric, defined on a scale from 0% to 100%, such that 0% indicates that the task could not be completed at all, and 100% indicates successful completion in no time (Eq. 1). A similar metric is described in Perry et al., 2011.

$$\text{Efficiency} = 2s / (1+e^{t}) \quad (1)$$

Where

- s = Success = {0 for failed tasks; 1 for completed tasks)
- t = Time-on-Task = time to complete the task, in minutes.

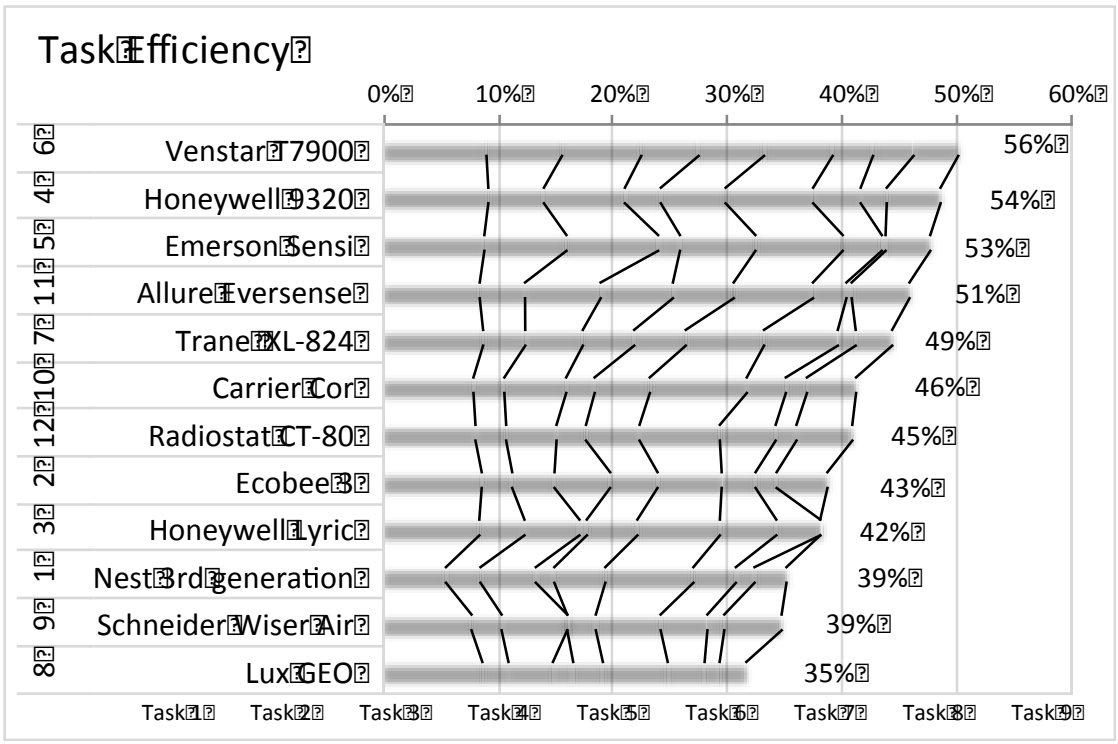
FIGURE 16. EFFICIENCY METRIC



EFFICIENCY SCORES

Figure 17 ranks the twelve thermostats tested in this study by final Efficiency scores calculated using Equation 1. Proportional contributions from each task are shown as different colored sections. The top ranked Venstar T7900, Honeywell 9320, and Emerson Sensi thermostats scored significantly higher than the Lux Geo. The differences between Efficiency scores for all other pairs of thermostats were not statistically significant.

FIGURE 17. EFFICIENCY SCORES FOR ALL THERMOSTATS TESTED



Statistical significance bounds: ±18% (α=0.1)

TABLE 7. DESCRIPTIONS OF 9 COMMON TASKS USED FOR EFFICIENCY SCORING

Task	Task Description
1	Identify indoor temperature
2	Set to cool. Identify target cooling temperature.
3	Change target cooling temperature
4	Use app to identify a scheduled cooling temperature
5	Set to heat. Identify target heating temperature.
6	Change the current target heating temperature.
7	Use app to increase the target heating temperature
8	Use the app to postpone heating until you get home
9	Access the Wi-Fi network screen on the thermostat

4.2 SURVEYS

Participants filled out surveys for each thermostat they tested. This section summarizes the most pertinent findings from the survey. The full set of questions and response frequencies are provided in Appendix A.

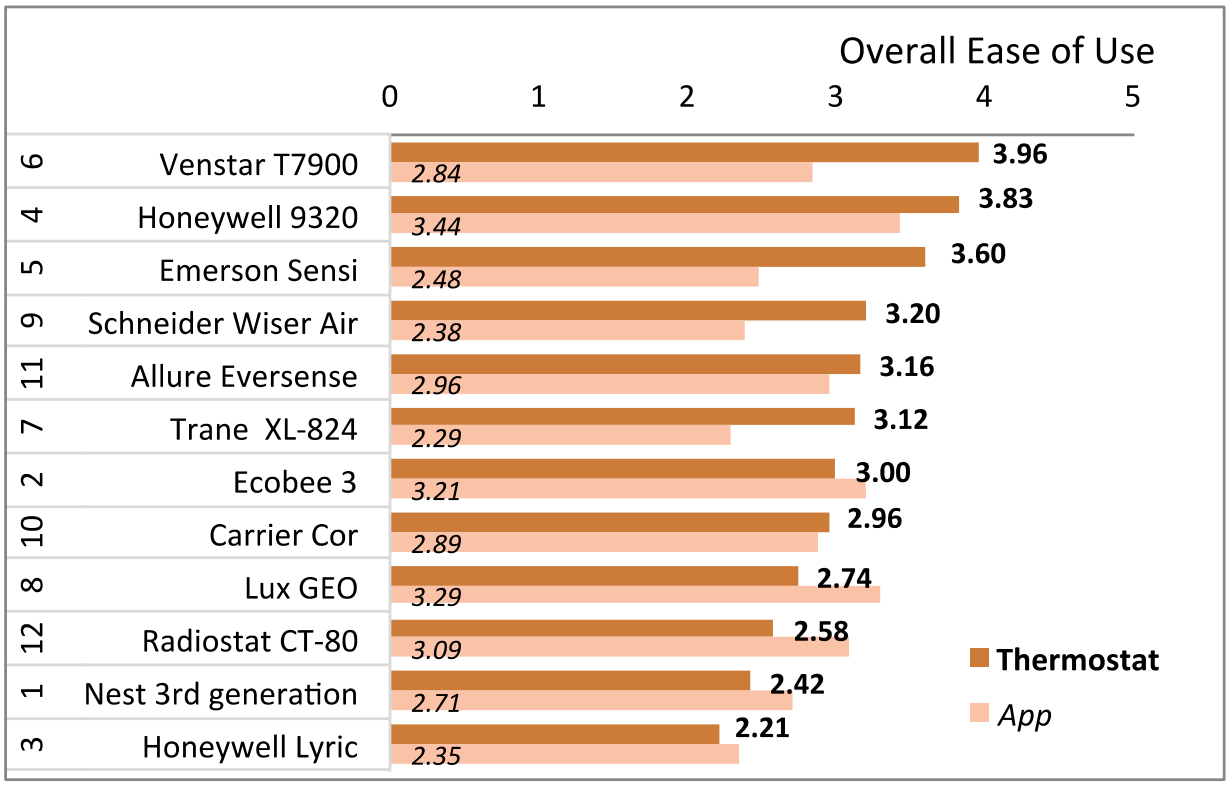
SATISFACTION RATINGS: THERMOSTAT AND APP FEATURES

This section provides the average thermostat and app ratings for the overall Ease of Use, Feel and Sound, and Appearance as rated by participants in these surveys.

EASE OF USE

Figure 18 summarizes participant ratings for "Overall Ease of Use" – question 3e in the survey (see Appendix A). The Venstar T7900, Honeywell 9320, and Emerson Sensi ranked in the top three spots, with average ratings significantly higher than the Honeywell Lyric. The smartphone app for the Honeywell 9320 earned the highest ratings, scoring significantly higher than did the smartphone apps for the Honeywell Lyric and the Trane XL-824. A notable trend that can be seen in Figure 18 is that the thermostats with the highest ratings were often paired with low-rating apps and vice-versa.

FIGURE 18. AVERAGE EASE OF USE RATINGS FOR THERMOSTATS AND APPS



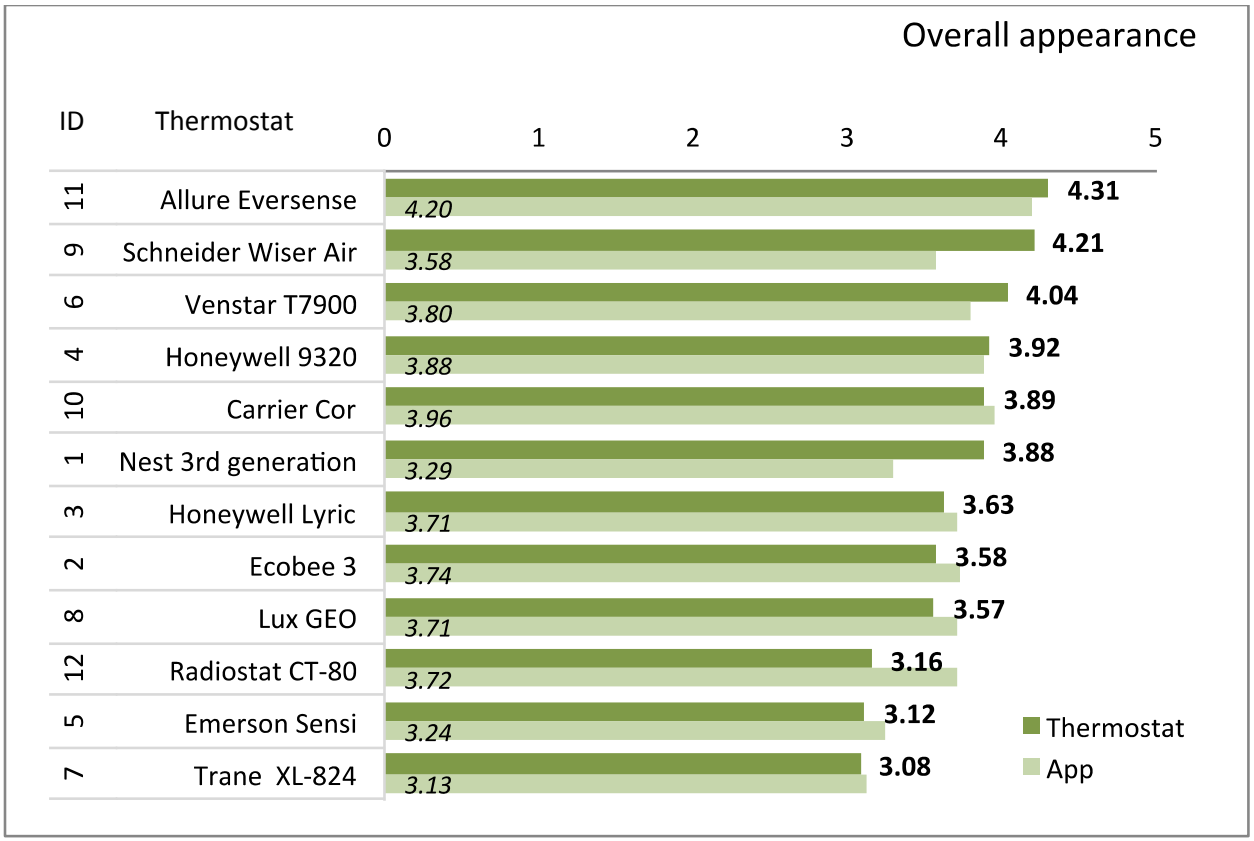
Statistical significance bounds for Thermostat: ±1.09; App: ±1.09 (α=0.1)

APPEARANCE

Figure 19 ranks the thermostats by participant scores for "Overall Appearance" – question 5e in the survey (see Appendix A). In this category, the Allure Eversense and Schneider Wiser Air had significantly higher average ratings than the Radiostat CT-80, Emerson Sensi, and Trane XL-824. The app for the Allure Eversense also earned the highest ratings, with a significantly higher average rating than the app for the Trane XL-824.

Appearance ratings for thermostats-app pairs were very similar, with a few exceptions. The Schneider Wiser Air and Nest had substantially higher ratings for the thermostat than for the app. Appearance ratings for the app exceeded thermostat Appearance ratings for the six lowest rating thermostats.

FIGURE 19. AVERAGE APPEARANCE RATINGS FOR THERMOSTATS AND APPS



Statistical significance bounds: Thermostat ±0.92; App ±1.07 (α=0.1)

FEEL AND SOUND

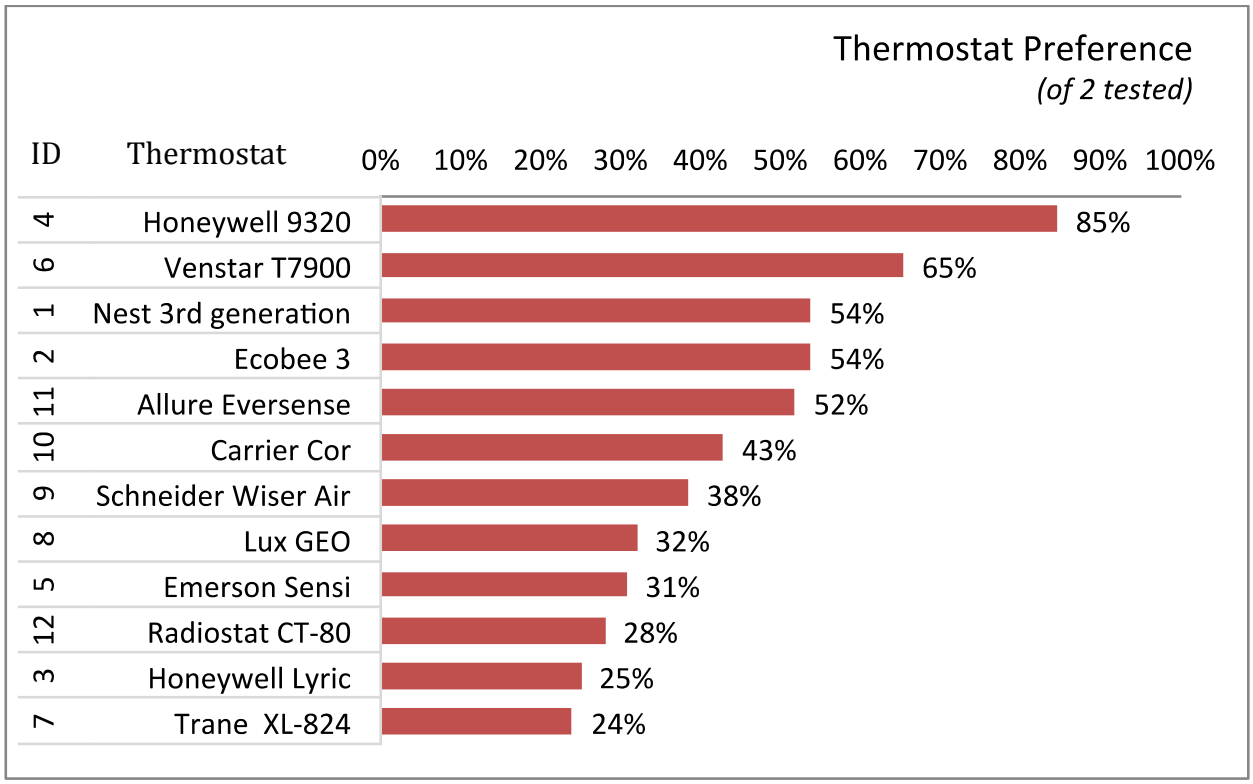
In contrast to the findings reported in SMUD's 2013 thermostat usability study, the 2015 thermostat lineup showed no statistically significant differences between participant ratings for "Overall Feel and Sound" (α=0.1). Average scores for Overall Feel and Sound can be found in Appendix A.

THERMOSTAT PREFERENCE: CHOICE OF 2 UNITS TESTED

In the survey filled out after the second thermostat test, participants were asked to choose the thermostat they would purchase given the choice of the two units they had tested. Preference scores were calculated for each thermostat as the percentage of participants that preferred that thermostat to the other thermostat tested (Figure 20).

Eighty-five percent of participants that tested the Honeywell 9320 chose it as their preferred thermostat, with participants indicating that it was easy to use, responsive to the touch, and had large easy-to-read text. More than half of participants picked the Venstar T7900, Nest, Ecobee 3, and Allure Eversense over their competitors.

FIGURE 20. THERMOSTAT PREFERENCE SCORES

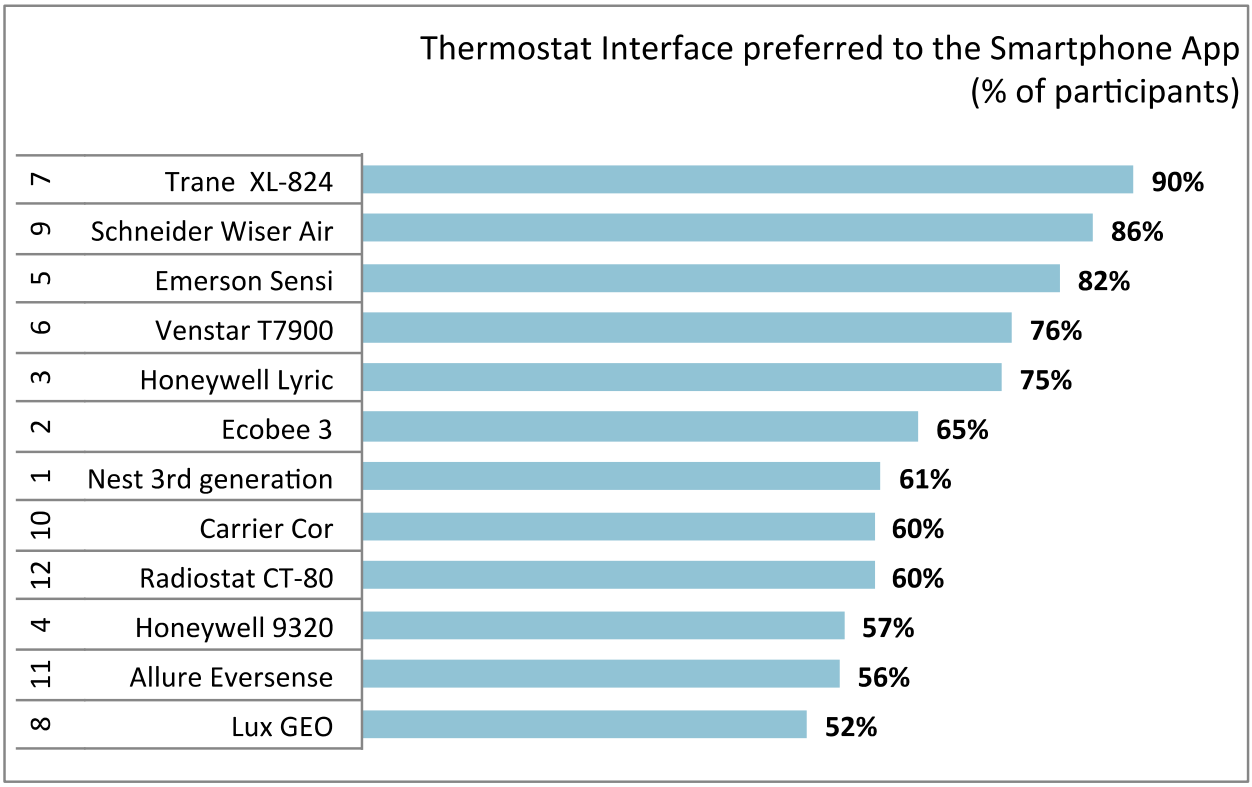


Statistical significance bounds: ±53% (α=0.1)

Interface Preference: Thermostat vs. Smartphone App

A statistically significant proportion of participants who tested the Trane XL-824, Schneider Wiser Air, Emerson Sensi, Venstar T7900, Honeywell Lyric, and Ecobee 3 preferred the thermostat interface to the smartphone app. Since these scores reflect the preference of the smartphone app over the thermostat interface (not relative preferences among apps) this could either indicate that the thermostat interface was particularly easy to use, or that the smartphone app was relatively difficult to use.

Figure 21. Preference of thermostat interface over the smartphone app

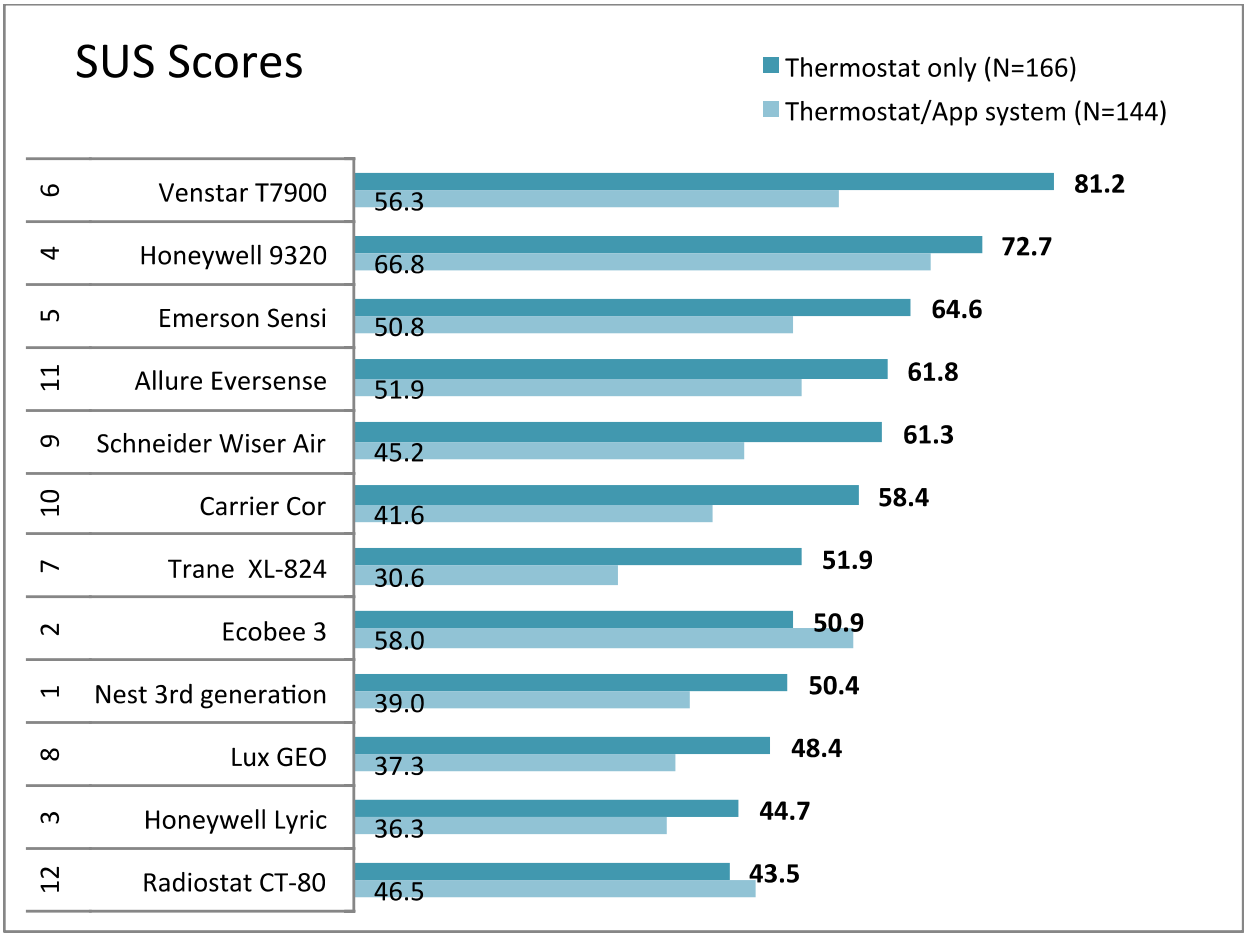


System Usability Scale

The most important change to the survey from the 2013 thermostat usability study was the addition of the System Usability Scale, a standard 10-point survey designed to assess the usability of any system. The SUS was selected among four commonly used usability questionnaires because it has been used often enough to have a known distribution, and has shown itself to be reliable (produces consistent results), valid (measures what it intends to measure) and discriminating (sensitive to differences between systems) (Tullis and Stetson, 2004). Also, it is the shortest of the standard usability questionnaires.

The SUS questions on the surveys distributed in the first 8 test sessions referred to the system as "the thermostat." For the last 7 sessions, SMUD requested that the survey refer to the system as "the thermostat/app system." For the remainder of this report, the SUS score collected in the first 8 sessions will be called the Thermostat SUS, and the SUS score collected in the last 7 sessions will be called the thermostat/app system SUS or just System SUS. Average SUS scores for both are provided in Figure 22.

Figure 22. Thermostat and system SUS scores



Statistical significance bounds (α=0.1): Thermostat ±29; System ±36

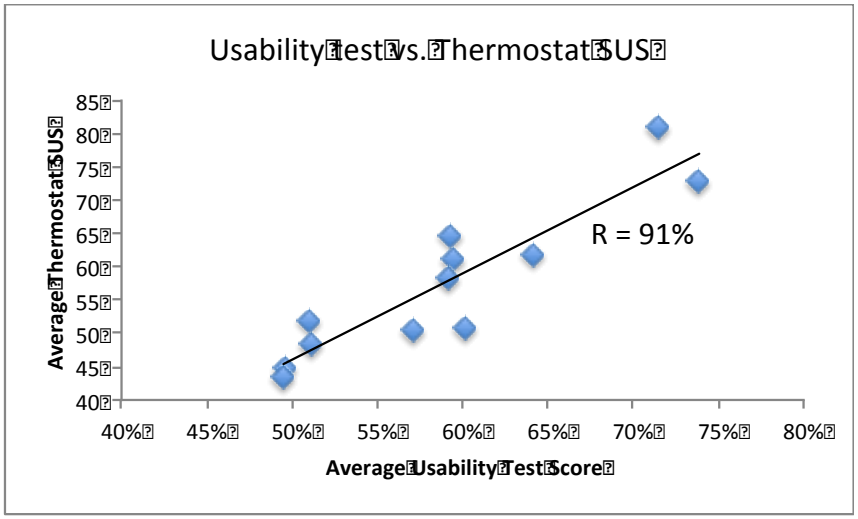
The average Thermostat SUS score for the Venstar T7900 was significantly higher than scores for the RadioStat CT-80, Honeywell Lyric, Lux GEO, Nest, Ecobee 3, and Trane. The average Thermostat SUS score for the Honeywell 9320 was significantly higher than the RadioStat CT-80. Differences between Thermostat SUS scores for all the other pairs of thermostats were statistically insignificant.

The average System SUS score for the Honeywell 9320 was significantly higher than the System SUS score for the Trane. System SUS scores for all the other pairs of thermostats were statistically insignificant.

To allow comparison of the usability test metrics with the SUS metric, average usability study test scores were calculated as the average of scores for Efficiency, Preference, Ease of Use, Feel and Sound, and Appearance. Pearson product-moment correlation coefficients were then calculated to determine the degree of correlation between the 12 average SUS scores and the 12 average usability test scores.

Correlation of average usability test scores with the Thermostat SUS was 91%, suggesting that the average score derived from a Thermostat SUS survey might be a reasonable proxy for the full usability testing completed for this study (Figure 23). Correlation of the average usability test scores with the System SUS scores was slightly lower at 83%.

FIGURE 23. 91% CORRELATION BETWEEN THERMOSTAT SUS AND USABILITY TEST SCORES



Power analysis using the standard deviation of the Thermostat SUS scores collected in this study indicates that at least 22 SUS surveys should be collected to achieve 90% confidence that the average SUS score is within 10 points of the average of the 22 recorded scores.

Advanced Feature Ratings

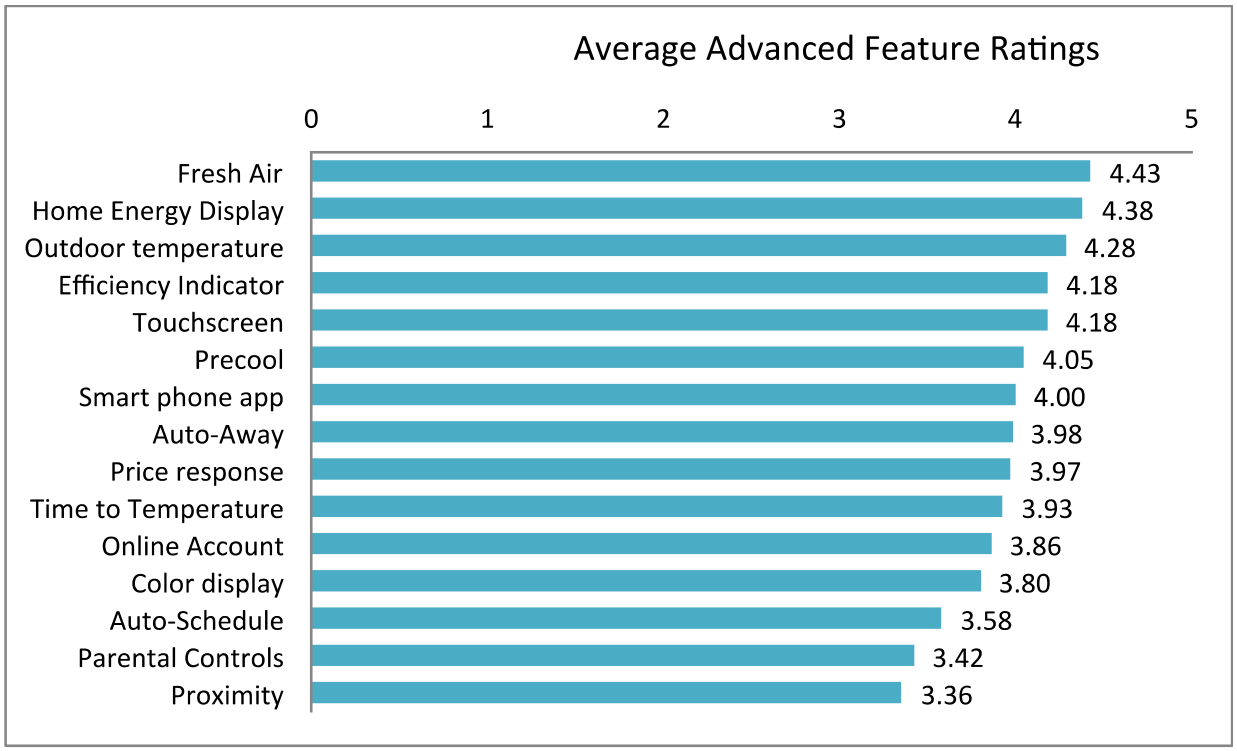
The second survey, completed after the completion of both thermostat tests, presented participants with the question: "Do you think you would find the following features useful on a thermostat in your home?" Possible responses ranged from 1, Not at all, to 5, Definitely. See Appendix A for the full list of survey questions.

Table 8. Survey questions to rate perceived usefulness of advanced features

7	Do you think you would find the following features useful on a thermostat in your home?
A	**Auto-Away**: The thermostat automatically adjusts the temperature when it senses your home is unoccupied.
B	**Auto-Schedule**: The thermostat programs your temperature preferences for you, based on your adjustments in the first week or two.
C	**Color display**: The main display has more than 2 colors.
D	**Efficiency Indicator**: The thermostat indicates when you adjust it to an energy efficient temperature setting.
E	**Fresh Air:** The thermostat alerts you when opening windows would be more economical than running the system.
F	**Home Energy Display**: The thermostat displays the amount of energy used by your home.
G	**Online Account**: You can use a computer to adjust your thermostat settings remotely.
h	**Outdoor temperature**: The thermostat can display the outdoor temperature.
i	**Parental Controls**: You can set your thermostat to allow changes to settings only after a password is provided.
J	**Precool**: You can set the thermostat to precool your home before a peak period in summer.
K	**Price response**: You can program the thermostat to adjust settings when the price of electricity changes.
l	**Proximity**: Your thermostat knows your location and automatically switches between home and away settings.
m	**Smart phone app**: You can use a smart phone to adjust your thermostat settings remotely.
n	**Time to Temperature**: The thermostat displays how long it will take to reach the target temperature.
o	**Touchscreen**: The main screen is also an input device.

Figure 24 shows the average ratings for the 15 advanced features described in Table 8. The top-rated advanced feature included a "Fresh Air" feature[4], where the thermostat notified occupants of an opportunity to save energy by opening windows, a real-time home energy display built into the thermostat[5], and a place for the current outdoor temperature. These three features scored significantly higher than color display, auto-schedule, parental controls and proximity functionality.

FIGURE 24. PERCEIVED USEFULNESS OF 15 ADVANCED FEATURES



Statistical significance bounds: ±0.44 (α=0.1)

Beyond these advanced features, some users used the discussion sessions to suggest a few additional thermostat features they would like. These included:

- Ability to enlarge the font or zoom in
- Ability to integrate with other devices, e.g. whole house fans
- Display bill balance on thermostat
- Simple climate control dials, with red for hot and blue for cold
- Voice control, through the thermostat and/or through the smartphone app
- Wireless unit that can be placed anywhere in the home

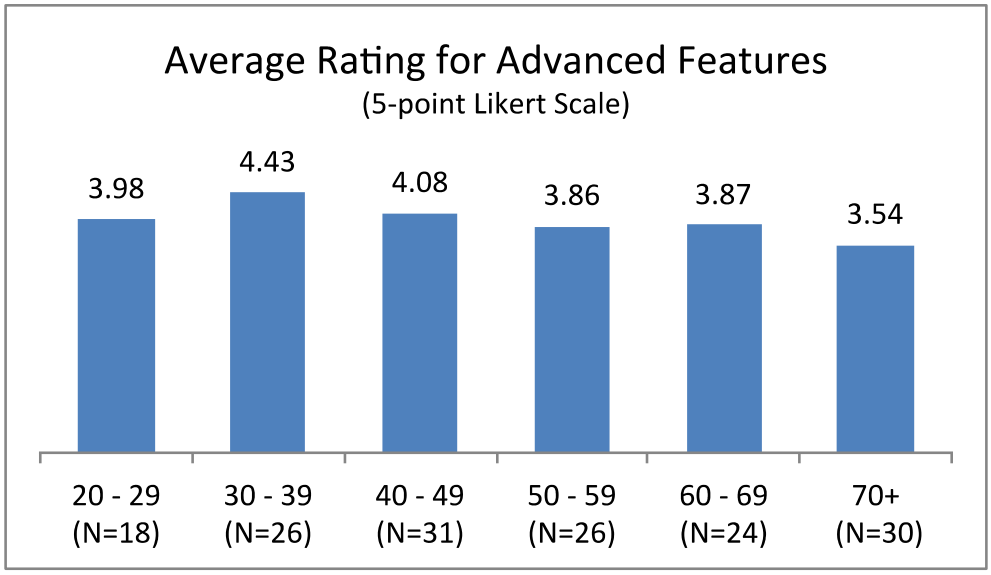
[4] Thanks to Kristin Heinemeier, Principal Engineer at UC Davis Energy Efficiency Center for suggesting we add this advanced feature to the survey.

[5] The author has a thermostat with a built-in real-time home energy display made by RCS Technologies.

Advanced Feature Preferences by Age

In general, younger users were more likely to consider the group of advanced features useful, as indicated by the average ratings provided in Figure 25. A notable exception to this is the 20-29 year olds, whose average rating for the 15 advanced features was significantly lower than that of the 30-39 year olds. In fact, the average rating for the latter group was significantly higher than any of the other age groups. The average rating for the 70+ group was significantly lower than any other age group.

Figure 25. Average ratings across all 15 advanced features, by age



Statistical significance bounds: ±0.21 (α=0.1)

Figure 26 and Figure 27 consider ratings by age group of each of the 15 individual advanced features. Most features follow a distribution of scores by age group that is similar to the distribution shown in Figure 25, with the highest ratings coming from the 30-39 year olds and the lowest ratings coming from the 70+ year olds.

FIGURE 26. AVERAGE RATING FOR EACH ADVANCED FEATURE, BY AGE GROUP (PART 1)

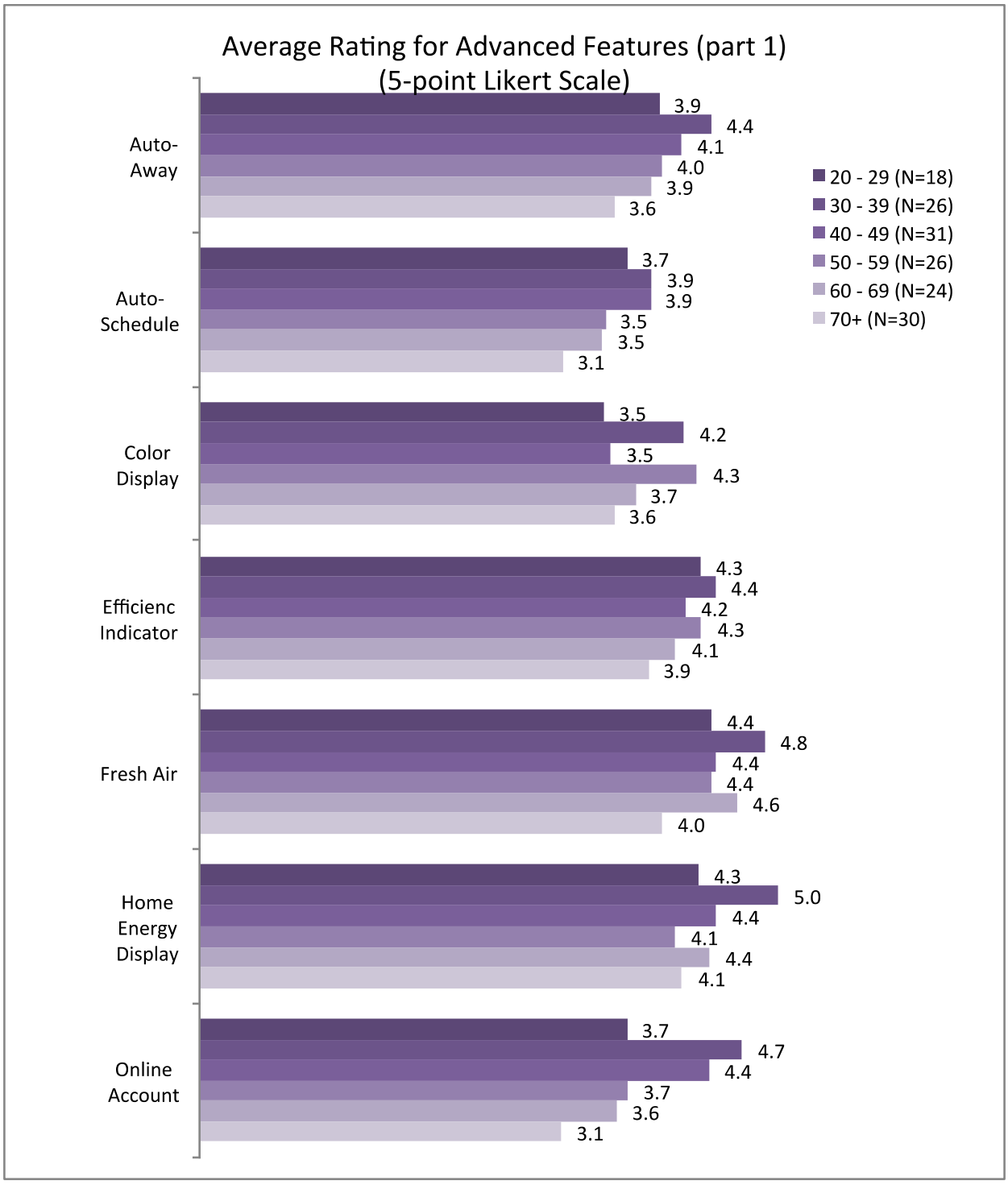
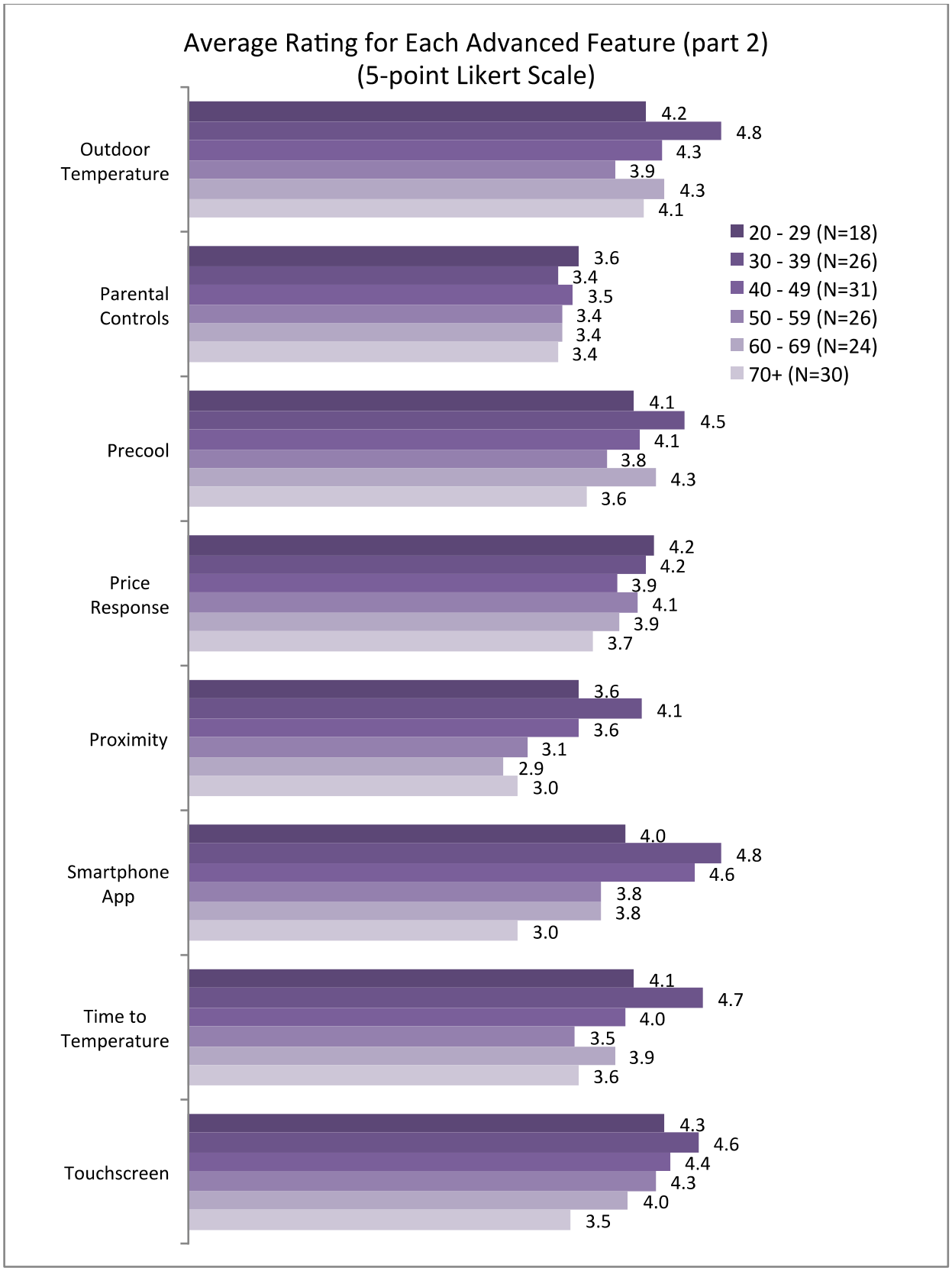


FIGURE 27. AVERAGE RATING FOR EACH ADVANCED FEATURE, BY AGE GROUP (PART 2)



4.3 Regression Models

To investigate the effects of thermostat features and participant characteristics on participant thermostat Preference and Efficiency, two ordinary linear regression models were constructed. A total of 26 potential independent variables were considered and divided into three categories: thermostat characteristics, participant characteristics, and satisfaction ratings collected through surveys, as shown in Table 9. Those found to have significant correlations with the dependent variables were then analyzed for multicollinearity.

Table 9. Regression model variables

Variable	Description	Variable Name
Dependent Variables	Thermostat **Preference** of 2 units tested	Preference
	Task **Efficiency** (see Equation 1)	Efficiency
Thermostat Characteristics	Size of the thermostat (cm^2)	Tstat_Size
	Screen size (cm^2)	Tstat_Screen_Size
	Color display screen (more than 2 colors)	Tstat_Screen_Color
	Thermostat has differentiated buttons for input	Tstat_Input_Buttons
	Thermostat has a touchscreen for input	Tstat_Input_Touchscreen
	Thermostat has a dial for input	Tstat_Input_Dial
	Size of the thermostat menu button/indicator (cm^2)	Tstat_Button.Size_Menu
	Height of the main menu text (mm)	Tstat_Text.size_Menu
	Height of the smallest text (mm)	Tstat_Text.size_Smallest
	Help button is available	Tstat_Buttons_Help
	Home button is available	Tstat_Buttons_Home
	Back button is available	Tstat_Buttons_Back
	Button to confirm chosen setting	Tstat_Buttons_Done
	Symbols are labeled with text	Tstat_Symbols_Labels
	Thermostat has input sounds and volume control	Tstat_Volume
	Number of steps needed to set thermostat mode	Tstat_Steps_Mode
	App looks and functions just like the thermostat	App_Consistent
Participant Characteristics	Participant age (years)	Participant_Age
	Participant rents their home	Participant_Renter
	Participant self-rated confidence using a computer	Participant_ComputerIQ
	Participant self-rated confidence using a smartphone	Participant_SmartphoneIQ
Satisfaction Ratings	Overall ease of use of the thermostat (Q3e)	Tstat_Rated_Ease.of.Use
	Overall feel and sound of the thermostat (Q4e)	Tstat_Rated_Feel.Sound
	Overall appearance of the thermostat (Q5e)	Tstat_Rated_Appearance
	Overall ease of use of the smartphone app (Q3e)	App_Rated_Ease.of.Use
	Overall appearance of the smartphone app (Q5e)	App_Rated_Appearance

Preference Model

The first model incorporated nine independent variables found to be correlated with Preference scores and selected to avoid multicollinearity. Table 10 lists the independent variables along with a summary of the model output.

Table 10. Preference model summary results

	Preference	Coefficient	Pr(>|z|)	[1]
Thermostat Characteristics	Tstat_Screen_Color	0.5324	0.2083	
	Tstat_Screen_Size	0.0078	0.5734	
	Tstat_Buttons_Help	0.3689	0.3376	
	Tstat_Text.size_Menu	**0.7012**	**0.0111**	*
Satisfaction Ratings	**Tstat_Rated_Ease.of.Use**	**0.4032**	**0.0197**	*
	Tstat_Rated_Feel.and.Sound	0.1734	0.3465	
	Tstat_Rated_Appearance	**0.5766**	**0.0017**	**
	App_Rated_Ease.of.Use	0.2017	0.2222	
	App_Rated_Appearance	-0.2561	0.1520	

[1] Statistical Significance codes: ***: α=0.001; **: α=0.01; *: α=0.05

Based on the Preference model output, one of the thermostat features and two of the satisfaction ratings included in the model significantly increased the likelihood that a participant would choose one thermostat over another. None of the participant characteristics had a significant impact on thermostat preference.

1. Thermostat: Larger menu text size
2. Satisfaction: Higher rating for "Overall Ease of Use" of the thermostat
3. Satisfaction: Higher rating for "Overall Appearance" of the Thermostat

For the full Preference model output and related analysis details, see Appendix B.

Efficiency Model

The second model investigated the individual effects of variables that were correlated with Efficiency. The list of variables and a summary of model results are provided in Table 11.

Table 11. Efficiency model summary results

	Efficiency	Coefficient	P	[1]
Thermostat Characteristics	Tstat_Screen_Color	0.0343	0.1696	
	Tstat_Screen_Size	**0.0038**	**0.0015**	**
	Tstat_Buttons_Help	-0.0269	0.2845	
	Tstat_Buttons_Done	-0.0288	0.3721	
	Tstat_Symbols_Labels	0.0042	0.8877	
	Tstat_Steps_Mode	**-0.0267**	**0.0002**	***
Participant Characteristics	**Participant_Age**	-0.0042	0.0000	***
	Participant_Renter	-0.0166	0.4385	
	Participant_ComputerIQ	**0.0441**	**0.0016**	**
	Participant_SmartphoneIQ	**0.0268**	**0.0263**	*
Satisfaction Ratings	Tstat_Rated_Ease.of.Use	0.0123	0.2548	
	Tstat_Rated_Feel.and.Sound	0.0190	0.1005	
	Stat_Rated_Appearance	-0.0132	0.2383	
	App_Rated_Ease.of.Use	**0.0352**	**0.0008**	***
	App_Rated_Appearance	-0.0064	0.5743	

[1] Statistical Significance codes: ***: α=0.001; **: α=0.01; *: α=0.05

Based on the Efficiency model output, two thermostat characteristics, three participant characteristics, and one satisfaction rating were associated with significantly higher thermostat Efficiency scores as follows.

1. Thermostat: A larger screen size
2. Thermostat: Smaller number of steps needed to change to cooling mode
3. Participant: Younger
4. Participant: Higher computer confidence
5. Participant: Higher smartphone confidence
6. Satisfaction: Higher rating for smartphone app "Overall Ease of Use"

For the full Efficiency model output and related analysis details, see Appendix B.

Significant Predictors of Efficiency and Preference

The following variables were significant predictors of Efficiency, Preference, or both.

Screen Size

The research team measured the dimensions for each screen and calculated the screen size in square centimeters. Screen sizes of the thermostats included in this study ranged from 12.5 cm^2 for the lowest ranked Honeywell Lyric to 56.4 cm^2 for the Trane XL-824. A larger screen size was associated with significantly improved Efficiency scores – but not with improved Preference scores.

In contrast to the 2013 usability study, a color display was not associated with significantly higher scores for Efficiency or Preference. Whereas color display ranked 7th of the 15 advanced features in the 2013 study, this time color display moved down to the 12th place position.

Menu Text Size

The research team measured and documented the height of the main menu item text in millimeters (mm). Main menu text sizes for this study ranged from 1.0 mm to 3.0 mm. Larger menu text sizes were significant predictors of Preference scores, but had no significant effect on Efficiency scores.

Rating for Ease of Use

The thermostats Ease-of-Use rating was a significant predictor of Preference, while the smartphone app Ease-of-Use rating was a significant predictor of Efficiency. This makes sense, since Preference ratings referred to the "thermostat you would choose to have installed," thus focusing on the thermostat unit rather than the app. In contrast, Efficiency scores were based on task completion, which was frequently achieved through the smart phone app.

Rating for Appearance

In contrast to the 2013 finding, the rating for the "Overall Feel and Sound" of the 2015 thermostats was not a significant predictor of Preference, while "Overall Appearance" was a significant predictor this time.

Steps to Change Mode

For each thermostat, the research team counted the number of steps required to put the thermostat at into heating or cooling mode. For the thermostats in this study, the number of steps ranged from 1 to 6. A smaller number of steps needed change mode was associated with higher Efficiency scores. No significant effect was indicated for Preference scores.

Participant Age

Participants in this study were recruited to fill six age groups designated by year of birth, beginning at 1995 and moving backward in 10-year increments. A decision to coordinate with the 2013 usability study using year of birth (rather than number of years since birth) was made based on the idea that generational impacts were likely stronger than experiential impacts. Regression analysis indicated that higher ages were associated with lower Efficiency scores, but had no influence on thermostat Preference. Age did, however, influence preferences for specific advanced features, with older participants generally being much less interested in advanced features than their younger counterparts.

Participant Confidence with Computers and Smartphones

Higher self-reported confidence ratings with computers and smartphones were significant predictors of thermostat Efficiency scores, but were not significantly associated with Preference scores. Participants' self-reported confidence with thermostats had no significant impact on Efficiency or Preference.

4.4 Results Summary by Thermostat

Table 12 lists the Satisfaction, Efficiency, and Preference rankings for all 12 thermostats and provides the Overall Grade based on an average of scores. Following this table are more detailed summaries of results and participant comments for each thermostat.

Table 12. Summary results of participant pros and cons of tested thermostats

ID	Device	Image	Satisfaction Rank *	Preference Rank	Efficiency Rank	Overall Grade
4	Honeywell 9320		2	1	2	A
6	Venstar T7900		1	2	1	A
11	Allure Eversense		4	5	4	B+
2	Ecobee 3		7	3	8	B
9	Schneider Wiser Air		3	7	11	B
10	Carrier Cor		6	6	6	B
5	Emerson Sensi		5	9	3	B
1	Nest 3rd generation		8	4	10	B-
8	Lux GEO		9	8	12	C+
7	Trane XL-824		10	12	5	C+
12	Radiostat CT-80		12	10	7	C
3	Honeywell Lyric		11	11	9	C

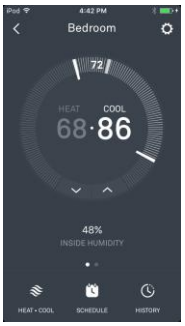
* Based on the average of the Overall Ease of Use, Feel and Sound, and Appearance ratings.

1. Nest 3rd Generation

Ease of Use

Despite the pre-test task designed to familiarize participants with thermostat input mechanisms, video recordings showed that nearly one-quarter (23%) of the participants who tested the Nest had trouble with the push-turn mechanism. Of these, most discovered the push and turn functionality, but had trouble putting the two together to complete tasks.

The majority of participants had no trouble figuring out the Nests input mechanism. If participants were unable to use the thermostat they completed tasks using the smartphone app.



Survey comments indicated a roughly even split between those who thought the Nest was easy to use and those who thought that the Nest was complicated, hard to navigate, or not intuitive. Some mentioned they liked the dial and said Nest was a smart system while others found the icons and colors confusing. A few participants said it was difficult to find and change settings on the thermostat. One participant locked himself out of the thermostat by accidently setting up a password. One wished for an instruction manual.

Feel and Sound

One participant said Nest would be easier to use if it was a touchscreen rather than push-turn thermostat. A few didn't like how the thermostat felt when pressed, with one mentioning it was hard to press. One participant said the Nest felt cheap.

REPORT CARD

SUS	45
Ease of Use	48%
Feel and Sound	67%
Appearance	78%
Efficiency	39%
Preference	54%
Overall Grade	B-

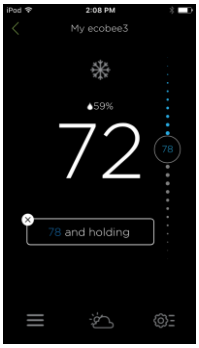
Appearance

Most participants liked the style of this thermostat and the fact it had a dial as a part of the input mechanism. One participant said he liked the "bold temperature number" right in the middle of the screen, but some participants mistook it for the current indoor temperature when in fact it represents the current target temperature. While most were fine with the size of Nest, one said it was too small.

2. Ecobee 3

Ease of Use

While most participants were able to figure out the touchscreen input mechanism of the Ecobee 3, a few had trouble figuring out the slider to change the target temperatures. Some participants thought the thermostat was difficult to program and navigate. A couple of the participants mentioned they would do a lot better on the tasks if they had an instruction manual. Some participants got stuck in Comfort Setting Home screen and adjusted comfort setting temperature rather than the target temperature.



While the smartphone app received a slightly higher Ease of Use rating, there was a significantly higher proportion of people who stated they would be more likely to use the thermostat (65%) regularly if they had to choose between the thermostat and the smartphone app.

Feel and Sound

Participants liked the responsiveness of the Ecobee 3 touchscreen. Video review revealed that many participants had trouble moving the slider, and had to try several times before setting the desired target temperature.

Appearance

Participants liked the modern sleek look of the Ecobee 3, but a few complained about the color theme. They thought white on black was hard to see and would prefer it the other way around – black on white, and perhaps using more colors. A few mentioned they liked the consistency between the smartphone app and the thermostat. Some found symbols to be somewhat confusing. One person said "…symbols did not open up to what I logically thought they stood for…" Several mentioned the thermostat and the screen were too small.

REPORT CARD	
SUS	54
Ease of Use	60%
Feel and Sound	72%
Appearance	72%
Task Efficiency	43%
Preference	54%
Overall Grade	B

3. Honeywell Lyric

Ease of Use

Most of the participants who tested Honeywell Lyric found it "not user friendly", "hard to figure out", and "not easy to use." Many felt there was need for an instruction manual. A couple commented on a poor main menu and no main menu button. Even so, a few participants said it was basic, easy to learn, and easy to adjust.

Most participants had trouble identifying scheduled cooling temperatures, which resulted in the low efficiency score for Task 4. None of the participants were able to find the Wi-Fi settings screen, which required holding one of the two buttons for five seconds. This resulted in a task Efficiency score of 0 for Task 9.

Many participants liked the idea of having a smartphone app to control settings on the thermostat but found it difficult to use. Three-quarters of participants said they would be more likely to use the thermostat on a regular basis if given a choice between the smartphone app and the thermostat.

Feel and Sound

The Honeywell Lyric received a relatively low overall feel and sound rating – lower than seven of the other tested thermostats.

Appearance

Some participants mentioned they liked the modern look and sleek design of Honeywell Lyric. One mentioned he liked the screen size.

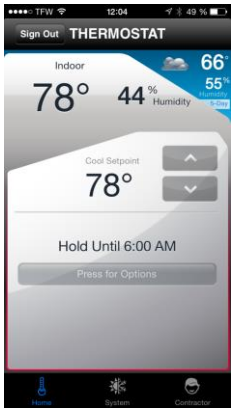
REPORT CARD

SUS	40
Ease of Use	44%
Feel and Sound	64%
Appearance	73%
Efficiency	42%
Preference	25%
Overall Grade	C+

4. Honeywell 9320

Ease of Use

Most participants found Honeywell 9320 easy to use and very intuitive. A few mentioned they wished there was an instruction manual or better on-screen instructions. The video review revealed that none of these participants clicked the help button available in most menu screens. Switching between cooling and heating modes required accessing the menu labeled "System" – a label that participants appeared to find unintuitive. Many participants were able to complete mode-switching tasks only after exploring the thermostat.



Some of the favorite features of the Honeywell 9320 were the ability to show the indoor and outdoor temperature and humidity level. Participants also liked the schedules and appreciated how easy it was to create them. Some participants mentioned they liked the clear arrows to set temperatures and they also liked current temperature and "set to" temperature right next to each other with menu on top of the screen.

Feel and Sound

Most participants liked the feel of the Honeywell 9320 but one thought the touchscreen needed too much pressure to activate. Some participants discovered the setting to turn the button sound on or off, so some participants experienced the Honeywell 9320 with sound and some experienced it without. Feedback on this was mixed, with some who had the sound wished it was completely silent or at least had volume control, while those who had no sound complained that there was no auditory feedback. Some wished there was a confirmation sound.

Appearance

While most participants liked the appearance of the Honeywell 9320, there were a few who said the thermostat was too thick, sticking out from the wall too far. The amount of information on the screen brought up different responses; some liked it while others said the screen was too busy. A few mentioned they liked different colors, bright display, and the ability to change color themes. A couple participants didn't like that the thermostat and smartphone app were not consistent.

REPORT CARD

SUS	70
Ease of Use	77%
Feel and Sound	76%
Appearance	78%
Efficiency	54%
Preference	85%
Overall Grade	A

5. Emerson Sensi

Ease of Use

Most participants found Emerson Sensi very easy to use. Some stated they needed an instruction manual but they mostly wanted instructions for the smartphone app. When asked which interface they would be more likely to regularly use, 82% of participants of 22 who responded chose the thermostat interface.

The reasons for the low efficiency score for task 9 are uncertain. Many participants quickly discovered the Wi-Fi screen but did not mark the task Done on their checklist and continued to search. It is likely that they were confused by the appearance of the Wi-Fi screen and so didn't realize that they had found it.

During testing there were connectivity problems with the thermostat about every hour – an issue that was not apparent for other thermostats. The only way to reset the Wi-Fi connection is to power the unit off.

Feel and Sound

Most participants were fine with buttons but one said the buttons required too much pressure.

Appearance

Some participants liked the classic look of the Emerson Sensi saying it looked "familiar," but there were a few who said it had an outdated appearance. A few participants complained that the text was hard to read and that the screen kept going dark, not staying lit for long enough. While several participants thought the screen was too small, they liked large print of the current and target temperatures. Some complained that the thermostat interface was not consistent with the smartphone app.

REPORT CARD

SUS	58
Ease of Use	72%
Feel and Sound	78%
Appearance	62%
Task Efficiency	53%
Preference	31%
Overall Grade	B

6. VENSTAR T7900

EASE OF USE

Most participants mentioned in the comments that the Venstar T7900 was easy to navigate. This is also supported by fact that the Venstar received the highest Efficiency score and Ease of Use rating. One participant said he needed an instruction manual. Some thought the home screen was simple and others thought there were too many choices, making it too complicated.

Many participants had trouble completing tasks using the smartphone app because they didn't push the "Send" button that appeared at the top of the screen every time they attempted to change a setting. Eighty percent stated they would be more likely to use the thermostat regularly if given a choice between the thermostat and smartphone app interface.

FEEL AND SOUND

Some participants complained about the sound as being too loud when they pushed buttons. A few discovered the sound settings and turned the sound off, so some participants experienced the Venstar with the sound and some experienced it without it. Participants liked the feel of the touchscreen but a couple of them complained that the menu buttons were too close to the frame, so the plastic frame interfered with their touch.

APPEARANCE

Most participants liked the size and the look of the thermostat but a couple thought it looked outdated and was too thick. One participant thought Venstar looked cheap. Many mentioned they liked the large print, bright colorful screen, and color indicators for heat and cool. A few participants didn't like the inconsistency between the smartphone app and the thermostat, as this required them to learn two different interfaces.

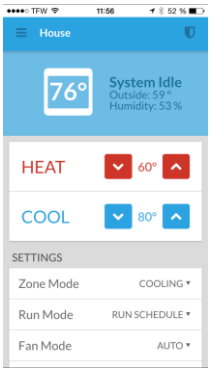
REPORT CARD

SUS	71
Ease of Use	79%
Feel and Sound	76%
Appearance	81%
Task Efficiency	56%
Preference	65%
Overall Grade	A

7. Trane XL824

Ease of Use

The Trane XL824 brought up different reactions from participants. Some thought it was easy to use while others complained that it was "hard to locate things," and one wished there were instructions on how to use the thermostat and the smartphone app.



Most participants were able to discover the touchscreen input mechanism right away but 36% were unable to increase the setpoint by 1 degree within 30 seconds.

Nearly all participants (90%) said they would be more likely to use the thermostat regularly if given a choice between the thermostat and the smartphone app interface.

Feel and Sound

Participants liked the touchscreen, but thought it was not responsive enough to their touch. The Trane received the lowest Overall Feel and Sound rating when compared to all other thermostats.

Appearance

The Trane also received the lowest Appearance ratings. Some participants thought it was clunky and looked antiquated. One mentioned he didn't like the overall design of the thermostat. A few participants liked different colors for cooling and heating. Several participants said that the screen was too dark and the print was too small for their eyes. One participant said the home screen was too busy and another participant thought there was a lot of wasted space on the screen.

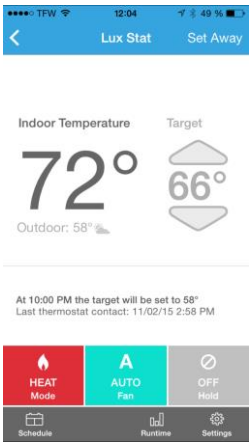
REPORT CARD

SUS	42
Ease of Use	62%
Feel and Sound	57%
Appearance	62%
Task Efficiency	49%
Preference	24%
Overall Grade	C+

8. Lux GEO

Ease of Use

Most participants found the Lux GEO difficult to use. They especially disliked the dial. A few were unable to figure out the push part of the push-dial mechanism needed to navigate through the menu and change settings. While there was one participant who mentioned he liked the ability to change target temperature using the dial, he found no use for the push function.



A few participants mentioned that the smartphone app was easier to use, and it received a slightly higher Ease of Use rating than did the thermostat, but there was no significant difference between the number of participants that said they would be more likely to use the thermostat or app regularly. Most participants had trouble identifying the scheduled cooling temperature, because they didn't realize that they had to switch between the schedules.

Feel and Sound

Some thought the Lux GEO thermostat was made of poor quality materials and found dial difficult to turn. One participant mentioned he disliked Lux GEO because it didn't have a touchscreen.

Appearance

Some thought Lux GEO was too big and bulky and looked outdated. Many liked the big number for current temperature in the middle of the screen.

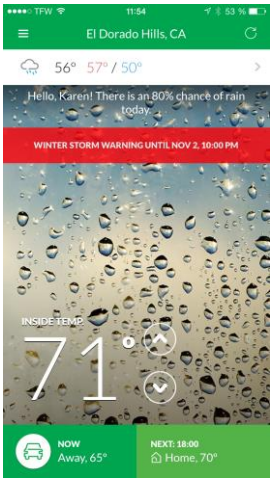
REPORT CARD

SUS	43
Ease of Use	55%
Feel and Sound	62%
Appearance	71%
Task Efficiency	35%
Preference	32%
Overall Grade	C+

9. Schneider Wiser Air

Ease of Use

Most participants found Schneider Wiser Air easy to use, with only a few complaining the thermostat was complicated or hard to navigate. A couple people stated they would do much better if they were provided with instructions. None of the participants were able to complete the task of identifying the scheduled temperature using the smartphone app because the "Schedule" screen showed preset names and times, but not temperatures. Scheduled temperatures were provided on a different screen, and none of the participants were able to figure this out.



Eighty-six percent of participants said they would be more likely to use the thermostat interface regularly if they had to choose between the thermostat and the smartphone app. It was unclear whether participants were frustrated with not being able to complete "schedule" task or if the smartphone app was indeed hard to navigate.

Some participants landed in the presets menu screen while exploring the thermostat and didn't realize they were adjusting preset temperatures rather than target temperatures, therefore not completing the tasks correctly. Two participants mentioned that it was hard to go back to the home screen.

Feel and Sound

The soundless Schneider Wiser Air took fifth place for the Overall Feel and Sound rating. One participant mentioned he wished there was a confirmation sound.

Appearance

Participants liked the modern "digital look and feel" of the Schneider Wiser Air. Most participants were fine with the size of the thermostat but one mentioned the small text was too hard to see.

REPORT CARD

SUS	53
Ease of Use	64%
Feel and Sound	72%
Appearance	84%
Task Efficiency	39%
Preference	38%
Overall Grade	B

10. Carrier Cor

Ease of Use

Some participants found the Carrier Cor easy to use and navigate while others thought the opposite. Those who found Carrier Cor hard to navigate mentioned there were too many choices and buttons. A few participants thought symbols and icons were too confusing. Several participants mentioned needing an instruction manual. Video review indicated that none of those who thought the instruction manual would be useful clicked the help button available in every screen. Some participants said there were too many steps to change a simple setting. A few participants mentioned the thermostat was hard to use and seemed complicated at first but they were able to figure it out.

Feel and Sound

Most participants liked the feel of the unit with the exception of three who thought the touchscreen needed too much pressure. One of those three participants couldn't complete any of the tasks except for identifying current temperature, because no matter how hard or light she was pushing, the screen didn't register her touch.

Appearance

Most participants liked the design of the Carrier Cor. Many liked that the thermostat and smartphone app screens were consistent, so they didn't have to learn how to navigate two different devices. While there were a few participants who thought there were too many choices and buttons, one thought the thermostat had a simple interface and another said there was just enough information on the screen. A few participants liked the red and blue colors for heating and cooling. Some participants thought the screen and the thermostat itself were too small.

REPORT CARD

SUS	52
Ease of Use	59%
Feel and Sound	70%
Appearance	78%
Task Efficiency	46%
Preference	43%
Overall Grade	B

11. Allure Eversense

Ease of Use

Feelings for the Allure Eversense were mixed. Some thought it was very simple to use, easy to navigate, very intuitive and had an "easy menu," but a few found the thermostat hard to use and navigate. Those who thought the Eversense was not easy to use mentioned they would do a lot better if they had instructions. A few participants discovered the help button and pressed it, but it was unclear from the videos whether the Help screen helped them to complete the tasks or confused them even more. Most participants were unable to complete Task 8 using the smartphone app because they needed to connect the Proximity button located in the top right corner of the main screen to the Away and Home preset modes.

Feel and Sound

Participants liked the feel of the touchscreen but most found the Allure Eversense slow and unresponsive. Most had to push the touchscreen buttons several times before the thermostat would react.

Appearance

Participants liked the look of the thermostat but a few thought it was too big. Some said the Allure Eversense had a pretty display and liked the background picture. One participant had the opposite opinion and thought the background interfered with the data on the screen. Some liked the idea of being able to play the music and view pictures on the thermostat but thought these were unnecessary features for a thermostat. One participant thought the thermostat looked like "a little boombox."

REPORT CARD

SUS	57
Ease of Use	63%
Feel and Sound	69%
Appearance	86%
Task Efficiency	51%
Preference	52%
Overall Grade	B+

12. RADIOSTAT CT-80

EASE OF USE

Many participants found the RadioStat CT-80 difficult to use. The main confusion came from the mixed input mechanism of touchscreen and buttons. Once participants figured out they had to push the buttons on the right side to change target temperatures or to get to the main menu, it took some time for them to discover that the screen was a touchscreen. Some tasks, like changing the mode, required the use of the touchscreen. A few participants said that the terminology used for the labels was hard to understand. They also didn't like the amount of information of the screen.

FEEL AND SOUND

Some participants discovered the setting to turn the sound on or off, so some participants experienced the test with sound and some experienced it without. Feedback on this was mixed, with some who had the sound wished it was completely silent or at least had volume control, while those who had no sound complained that there was no auditory feedback.

APPEARANCE

Many liked the screen size of the RadioStat but a few complained the thermostat itself was too big and bulky. Some participants mentioned that the screen was too dull and hard to read unless they pushed a power button for the backlight to turn on, but even then they complained the light went off too quickly.

REPORT CARD

SUS	45
Ease of Use	52%
Feel and Sound	59%
Appearance	63%
Task Efficiency	45%
Preference	28%
Overall Grade	C+

5 Conclusions and Recommendations

The goal of this study was to link thermostat features and participant characteristics with metrics for thermostat usability and likeability for use in procurement of smart thermostats. During the 3-day lab study, 155 participants each performed identical tasks on 2 different thermostats, filled out surveys, and participated in focus group discussion sessions. Each of 12 thermostats was tested by between 24 and 28 participants, roughly evenly distributed by age and home ownership.

Surveys collected user ratings for test thermostats' ease of use, feel and sound, and appearance, along with participant preference for one of the two thermostats tested and ratings for 15 potential advanced features. Videos of individual thermostat tests were used to determine time-on-task for each task.

Thermostat Preference was calculated as the percentage of participants that chose that thermostat from the two they tested. Time-on-task values were used to calculate an Efficiency metric for each task and thermostat. Preference and Efficiency scores were used as the dependent variables in separate linear regression models that included thermostat features, participant characteristics, and thermostat feature satisfaction ratings as independent variables.

The main findings of this study are as follows.

Preference scores were similar across participants of differing age, gender, home ownership, and technology IQ, but were significantly higher for thermostats with:

1. Larger menu text size
2. Highly rated thermostat Ease of Use
3. Highly rated thermostat Appearance

Efficiency scores were significantly higher for thermostats with:

1. Easy-to-use smartphone apps (as rated by participants)
2. Larger screens
3. Fewer steps to change to heating or cooling mode

Efficiency was also influenced by the characteristics of the user. In particular, this study showed significantly higher Efficiency scores for:

1. Younger users
2. Users with higher computer confidence
3. Users with higher smartphone confidence

Table 13 provides the average scores and grades for all 12 thermostats tested in this study.

TABLE 13. SMART THERMOSTAT USABILITY RESULTS SUMMARY

	Satisfaction Scores							
Thermostat	Ease of Use	Feel & Sound	Appearance	Preference	Efficiency	Usability Test Score*	Grade	SUS Score
Honeywell 9320	77%	76%	78%	85%	54%	74%	**A**	73
Venstar T7900	79%	76%	81%	65%	56%	71%	**A**	81
Allure Eversense	63%	69%	86%	52%	51%	64%	**B+**	62
Ecobee 3	60%	72%	72%	54%	43%	60%	**B**	51
Emerson Sensi	72%	78%	62%	31%	53%	59%	**B**	61
Schneider Wiser Air	64%	72%	84%	38%	39%	59%	**B**	65
Carrier Cor	59%	70%	78%	43%	46%	59%	**B**	58
Nest 3rd generation	48%	67%	78%	54%	39%	57%	**B-**	50
Trane XL-824	62%	57%	62%	24%	49%	51%	**C+**	48
Lux GEO	55%	62%	71%	32%	35%	51%	**C+**	52
Honeywell Lyric	44%	64%	73%	25%	42%	50%	**C**	45
Radiostat CT-80	52%	59%	63%	28%	45%	49%	**C**	43

**Average of Satisfaction, Preference, and Efficiency scores.*

This study found three readily measurable thermostat characteristics to be significantly related to Efficiency or Preference score: screen size, text menu size, and number of steps to change the heating or cooling mode. While these characteristics can be used in procurement to specify requirements or desirable characteristics, they are not sufficient to ensure overall usability.

The high correlation of SUS scores with the Usability Test scores (see Figure 23) provides the possibility of obtaining reasonably reliable, valid and discriminating scores for procurement purposes for thermostats that have not been subjected to the full study. Power analysis using the standard deviation of the SUS scores collected in this study indicates that at least 22 SUS surveys should be collected to achieve 90% confidence that the average SUS score is within 10 points of the average of the 22 recorded scores. The research team recommends that SMUD establish as a prerequisite for future thermostat purchases a minimum average SUS score of 50, which corresponds to at least a B grade.

6 REFERENCES

Enerson, M. 2012. "Paired Comparison Usability Testing." *Software Magazine*. November.

Gunshinan, J. "Energy Star Changes Approach to Programmable Thermostats." *Home Energy*, March/April 2007.

Herter, K. and Y. Okuneva, 2013. J. *SMUD's 2012 Residential Precooling Study - Load Impact Evaluation*. Prepared by Herter Energy Research Solutions for the Sacramento Municipal Utility District.

Herter, Karen, and Yevgeniya Okuneva. 2014a. *SMUD's Load Impact Calculator (SLIC) for Smart Grid Cross-Pilot Evaluation*. Prepared by Herter Energy Research Solutions for the Sacramento Municipal Utility District, Sacramento, California.

Herter, K. and Y. Okuneva. 2014b. *SMUD's Residential Summer Solutions Study: 2011-2012.* Prepared by Herter Energy Research Solutions for the Sacramento Municipal Utility District.

Herter, K. and Y. Okuneva. 2014c. *SMUD's Smart Thermostat Pilot – Load Impact Evaluation*. Prepared by Herter Energy Research Solutions for the Sacramento Municipal Utility District.

Herter, K. and Y. Okuneva. 2014d. *SMUD's Low Income Weatherization & Energy Management Pilot – Load Impact Evaluation*. Prepared by Herter Energy Research Solutions for the Sacramento Municipal Utility District.

Herter, K. and Y. Okuneva. 2014e. *SMUD's Communicating Thermostat Usability Study*. Prepared by Herter Energy Research Solutions for the Sacramento Municipal Utility District.

Meier, A., C. Aragon, T. Peffer, D. Perry, M. Pritoni. 2011. Usability of residential thermostats: Preliminary investigations. Building and Environment 46, pp. 1891-1898.

National Institute for Standards and Technology (NIST). 2001. Common industry format for usability test reports, version 2.0. Gaithersburg: NIST.

Nielsen, J. 2005. "Time Budgets for Usability Sessions." Available online at http://www.nngroup.com/articles.

Norman, D.A. (1983). Design rules based on analyses of human error. Communications of the ACM, 26(4), 254-258.

Perry, D., C. Aragon, T. Peffer, A. Meier, M. Pritoni. 2011. Making Energy Savings Easier: Usability Metrics for Thermostats. *Journal of Usability Studies*, Vol. 6, Issue 4, pp. 226-244.

Rubin, J. and D. Chisnell. 2008. *Handbook of Usability Testing*. Wiley Publishing Inc., Indianapolis, Indiana.

Tullis, T. S., & Stetson, J. N. 2004. A comparison of questionnaires for assessing website usability. Proceedings of UPA 2004 Conference. Minneapolis, Minnesota. June 7-11.

U.S. Environmental Protection Agency. *ENERGY STAR® Program Requirements for Residential Climate Controls: Partner Commitments DRAFT 3 Version 1*. Available at https://www.energystar.gov/products/specs/node/161. April 9, 2012.

APPENDIX A –SURVEY RESPONSE DETAILS

AVERAGE FEATURE RATINGS FOR THERMOSTATS

TABLE 14. AVERAGE EASE OF USE RATINGS FOR THERMOSTATS

ID	Thermostat	Number of Tests	Information on the home screen	Input mechanisms	Meanings of words & symbols	Menu navigation	Overall ease of use
1	Nest 3rd generation	26	2.7	2.6	2.8	2.5	2.4
2	Ecobee 3	26	3.3	3.2	3.2	2.9	3.0
3	Honeywell Lyric	24	2.2	2.6	2.6	1.8	2.2
4	Honeywell 9320	26	3.9	4.0	3.6	3.6	3.8
5	Emerson Sensi	26	3.8	3.9	3.9	3.5	3.6
6	Venstar T7900	26	4.0	3.9	4.0	3.9	4.0
7	Trane XL-824	25	2.9	3.1	3.2	2.8	3.1
8	Lux GEO	25	3.1	2.7	2.8	2.6	2.7
9	Schneider Wiser Air	26	3.2	3.4	3.3	3.1	3.2
10	Carrier Cor	28	3.2	3.4	3.3	2.8	3.0
11	Allure Eversense	27	3.5	3.3	3.4	3.0	3.2
12	Radiostat CT-80	25	3.0	2.8	2.9	2.6	2.6

TABLE 15. AVERAGE FEEL AND SOUND RATINGS FOR THERMOSTATS

ID	Thermostat	Number of Tests	Pressure needed for input	Responsive-ness to input	Quality of materials	Audible feedback	Overall feel & sound
1	Nest 3rd generation	26	3.5	3.8	3.8	2.8	3.3
2	Ecobee 3	26	3.9	3.7	3.7	2.7	3.6
3	Honeywell Lyric	24	3.6	3.5	3.5	2.9	3.2
4	Honeywell 9320	26	3.8	3.8	4.0	3.3	3.8
5	Emerson Sensi	26	3.9	4.1	3.9	2.8	3.9
6	Venstar T7900	26	4.1	4.1	3.8	3.7	3.8
7	Trane XL-824	25	3.3	3.2	3.4	2.0	2.9
8	Lux GEO	25	3.2	3.1	3.5	2.6	3.1
9	Schneider Wiser Air	26	4.0	3.8	4.0	2.3	3.6
10	Carrier Cor	28	3.6	3.5	3.9	3.2	3.5
11	Allure Eversense	27	2.7	2.8	4.0	3.3	3.4
12	Radiostat CT-80	25	3.0	3.0	3.1	2.4	3.0

TABLE 16. AVERAGE APPEARANCE RATINGS FOR THERMOSTATS

ID	Thermostat	Number of Tests	Size & shape	Screen size	Color & style	Readability	Overall appearance
1	Nest 3rd generation	26	3.8	3.7	3.7	3.8	3.9
2	Ecobee 3	26	3.8	3.3	3.6	3.4	3.6
3	Honeywell Lyric	24	3.7	3.5	3.6	3.7	3.6
4	Honeywell 9320	26	4.0	4.1	3.8	3.8	3.9
5	Emerson Sensi	26	3.5	3.2	2.8	3.0	3.1
6	Venstar T7900	26	4.2	4.1	3.8	4.0	4.0
7	Trane XL-824	25	3.6	3.3	3.3	2.5	3.1
8	Lux GEO	25	3.3	3.4	3.3	4.0	3.6
9	Schneider Wiser Air	26	4.4	4.4	4.3	4.0	4.2
10	Carrier Cor	28	3.9	3.7	4.0	3.7	3.9
11	Allure Eversense	27	3.6	4.1	4.3	4.1	4.3
12	Radiostat CT-80	25	3.4	3.6	2.9	3.0	3.2

Average Feature Ratings for Smartphone Apps

Table 17. Average Ease of Use ratings for Smartphone Apps

ID	Smartphone Application	Number of Tests	Information on the home screen	Input mechanisms	Meanings of words & symbols	Menu navigation	Overall ease of use
1	Nest 3rd generation	26	2.75	2.92	2.88	2.75	2.71
2	Ecobee 3	26	3.17	3.46	3.22	3.17	3.21
3	Honeywell Lyric	24	2.82	2.65	2.55	2.04	2.35
4	Honeywell 9320	26	3.52	3.52	3.56	3.40	3.44
5	Emerson Sensi	26	3.12	2.92	3.00	2.36	2.48
6	Venstar T7900	26	3.38	3.42	3.40	2.71	2.84
7	Trane XL-824	25	2.48	2.62	2.46	2.29	2.29
8	Lux GEO	25	3.29	3.35	3.38	3.25	3.29
9	Schneider Wiser Air	26	2.88	2.96	2.81	2.31	2.38
10	Carrier Cor	28	3.19	3.38	3.15	2.81	2.89
11	Allure Eversense	27	3.29	3.28	3.16	2.96	2.96
12	Radiostat CT-80	25	3.43	3.43	3.17	3.04	3.09

Table 18. Average Appearance Ratings for Smartphone Apps

ID	Smartphone Application	Number of Tests	Color & style	Readability	Overall appearance
1	Nest 3rd generation	26	3.26	2.96	3.29
2	Ecobee 3	26	3.74	3.74	3.74
3	Honeywell Lyric	24	3.71	3.86	3.71
4	Honeywell 9320	26	3.71	3.40	3.88
5	Emerson Sensi	26	3.17	2.88	3.24
6	Venstar T7900	26	3.83	3.52	3.80
7	Trane XL-824	25	3.00	3.13	3.13
8	Lux GEO	25	3.79	3.42	3.71
9	Schneider Wiser Air	26	3.46	3.23	3.58
10	Carrier Cor	28	3.88	3.69	3.96
11	Allure Eversense	27	4.08	3.96	4.20
12	Radiostat CT-80	25	3.65	3.45	3.72

PARTICIPANT COMMENTS

TABLE 19. PARTICIPANT COMMENTS FROM SURVEYS - PARAPHRASED

ID	Device	Pros	Cons
1	Nest 3rd generation	• Style (4) • Ease to use (4) • Smart system • Dial (4) • Bold temp number • Size • App easy to use (2)	• Small • Difficult to find wifi settings • Difficult to change settings • Felt cheap • Don't like how it feels when pressed (3) • hard to press • Not touchscreen • Confusing icons • Complicated, hard to navigate, not intuitive (4) • Confusing colors • Need instructions • App hard to use (5)
2	Ecobee 3	• Responsive • Consistent app/tstat • Sleek look (3) • Slide temperature • Menu symbols at the bottom • Outside temperature • Easy to navigate (3) • Consistent app/tstat	• No instructions (4) • Difficult to program, navigate (5) • Symbols not clear • Dark, Black screen (3) • Small (5) • Slide temperature a bit cubersome • No time displayed • Need more colors
3	Honeywell Lyric	• Basic • Easy to learn • Easy to adjust • Look • Screen size • Sleek design • Dial	• Need instructions (5) • Not user friendly, hard to figure out (3) • No main menu button • Poor main menu • Not easy to use • Slow reaction • Humidity • Lack of feedback • Current temp is bigger than target temp

ID	Device	Pros	Cons
4	Honeywell 9320	• Simple to use, easy to navigate, intuitive (7) • Bright Display • Colors (2) • Print Size • Clear arrows to set temps • Menu at top • Sound • Similar to existing thermostat, easy transition • Inside temp • Outside temp • Humidity level • Easy to create schedules • Ability to change color themes (2) • Set to right next to current temp • Schedules (2) • Touchscreen (2) • Screen Design • Lots of info on the home screen • App easy to use (5)	• Hard to find wifi settings • Too thick • Too much info • No directions, instruction manual (2) • App/tstat not consistent (2) • Bad on screen instructions • Extra steps to get to home screen • Extra pressure needed • Send button to confirm on the app • No confirmation sound • App not easy to use (2)
5	Emerson Sensi	• Buttons • Looks classic, familiar (2) • Easy to use (6) • Large numbers • App easy to use (1)	• Pressure needed to push buttons • Hard to access wifi settings • Need instructions, instruction manual (3) • Not consistent app/tstat • Hard to read light coloring • Outdated appearance (2) • Does not stay lit • Small screen (3) • Dark screen • Not intuitive • App not easy to use (4)

ID	Device	Pros	Cons
6	Venstar T7900	• Large Print • Brightness • Background picture (4) • Color indicators • Colorful • Ease to use (9) • Date and time • Size • Simple home screen • Look • App easy to use (4)	• Sound (2) • Too thick (2) • Outdated (2) • App/tstat not consistent (2) • Menu buttons are too close to frame • Too many choices, complicated (2) • Cheap • No instructions • App not easy to use (5)
7	Trane XL-824	• Touchscreen • Ease to use (2) • Large current temp print • Different colors for heating/cooling • Good design • App easy to use (2)	• Dark screen • Wasted space on screen • Print too small (3) • Overall Design • Not responsive (3) • Clunky (2) • Looks antiquated • Hard to locate things • Busy home screen • Need instructions • App not easy to use (3)
8	Lux GEO	• Easy to navigate • Dial • Easy to read • Simple • App easy to use (6)	• Difficult to use, tstat and dial (5) • Looks dated • Poor quality • Too big and bulky • Not touchscreen
9	Schneider Wiser Air	• Easy to use, simple (5) • Modern, great digital look at feel (2) • Size • App easy to use (1)	• Hard to go back to home screen (2) • Small text hard to see • Difficult to understand (2) • No confirmation • Not consistent app/tstat • No instruction manual (2) • No symbol explanation • Complicated • App not easy to use (1)

ID	Device	Pros	Cons
10	Carrier Cor	• Consistent app/tstat • Simple interface • Easy of use (2) • Design • Enough info • Colors for heating cooling	• Too many choices, buttons (2) • Small stat, small screen (2) • Need instructions (2) • Hard to find wifi setting • Need pressure (3) • Hard to navigate, confusing (3) • Confusing symbols, icons (2) • Too many steps to change settings • Grey color
11	Allure Eversense	• Simple to use, easy to navigate, intuitive • Background picture • Pretty display • Design and feel • Easy menu (2) • Touchscreen • App easy to use (2)	• Too big (3) • No instructions (4) • Slow, not responsive (6) • Background interfers with data • Hard to navigate, use (3) • App hard to use (4)
12	Radiostat CT-80	• Screen size (2) • Easy to use (3) • App easy to use (3)	• Confusing, hard to navigate (4) • Not obvious that it's touchscreen • Terminology hard to understand • Too loud (2) • Screen too dull • Too much info on screen • Big and bulky • Light turns off too quick • Hard to read • App easy to use (1)

Survey Response Frequency Tables

Survey Question 1: Test Thermostat

Do you have this thermostat at home?

ID	Thermostat	N	Yes	No	NA
1	Nest 3rd generation	26	0.0	100.0	0.0
2	Ecobee 3	26	0.0	100.0	0.0
3	Honeywell Lyric	24	0.0	100.0	0.0
4	Honeywell 9320	26	0.0	96.2	3.8
5	Emerson Sensi	26	0.0	100.0	0.0
6	Venstar T7900	26	0.0	96.2	3.8
7	Trane XL-824	25	0.0	100.0	0.0
8	Lux GEO	25	4.0	96.0	0.0
9	Schneider Wiser Air	26	0.0	100.0	0.0
10	Carrier Cor	28	0.0	100.0	0.0
11	Allure Eversense	27	0.0	100.0	0.0
12	Radiostat CT-80	25	0.0	96.0	4.0

Would you replace your thermostat at home with this one? (it's free)

ID	Thermostat	N	Yes	No	NA
1	Nest	12	33.3	66.7	0.0
2	Ecobee3	11	36.4	63.6	0.0
3	Lyric	12	16.7	83.3	0.0
4	Honeywell	11	54.5	45.5	0.0
5	Sensi	13	38.5	61.5	0.0
6	Venstar	11	54.5	36.4	9.1
7	Trane	12	16.7	83.3	0.0
8	Lux	12	25.0	75.0	0.0
9	WiserAir	14	42.9	57.1	0.0
10	CarrierCor	12	25.0	75.0	0.0
11	Allure	12	33.3	58.3	8.3
12	RadioStat	12	25.0	66.7	8.3

Survey Question 2: Thermostat vs. Smartphone Application

Which interface would you be more likely to use regularly? (circle one)

ID	Thermostat	N	Thermostat	App	Neither	Both	NA
1	Nest	26	53.8	34.6	0.0	0.0	11.5
2	Ecobee3	26	57.7	30.8	0.0	0.0	11.5

3	Lyric	24	75.0	25.0	0.0	0.0	0.0
4	Honeywell	26	50.0	34.6	0.0	3.8	11.5
5	Sensi	26	69.2	15.4	0.0	0.0	15.4
6	Venstar	26	61.5	15.4	3.8	0.0	19.2
7	Trane	25	76.0	8.0	0.0	0.0	16.0
8	Lux	25	48.0	44.0	0.0	0.0	8.0
9	WiserAir	26	69.2	11.5	0.0	0.0	19.2
10	CarrierCor	28	53.6	35.7	0.0	0.0	10.7
11	Allure	27	51.9	40.7	0.0	0.0	7.4
12	RadioStat	25	48.0	32.0	0.0	0.0	20.0

Survey Question 3: Ease of Use

1 Lousy

3. Fine

5. Great!

Information on the home screen - Thermostat

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	23.1	26.9	19.2	19.2	11.5	0.0
2	Ecobee 3	26	15.4	15.4	23.1	19.2	26.9	0.0
3	Honeywell Lyric	24	37.5	16.7	25.0	16.7	0.0	4.2
4	Honeywell 9320	26	0.0	3.8	30.8	34.6	30.8	0.0
5	Emerson Sensi	26	0.0	3.8	30.8	38.5	23.1	3.8
6	Venstar T7900	26	3.8	11.5	11.5	26.9	46.2	0.0
7	Trane XL-824	25	16.0	16.0	40.0	16.0	12.0	0.0
8	Lux GEO	25	16.0	8.0	32.0	20.0	16.0	8.0
9	Schneider Wiser Air	26	11.5	19.2	23.1	11.5	26.9	7.7
10	Carrier Cor	28	10.7	10.7	35.7	28.6	10.7	3.6
11	Allure Eversense	27	7.4	3.7	33.3	33.3	18.5	3.7
12	Radiostat CT-80	25	16.0	24.0	16.0	36.0	8.0	0.0

Input mechanisms (buttons, dials, etc.) – Thermostat

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	23.1	19.2	34.6	19.2	3.8	0.0
2	Ecobee 3	26	15.4	11.5	30.8	23.1	19.2	0.0
3	Honeywell Lyric	24	25.0	16.7	37.5	16.7	4.2	0.0
4	Honeywell 9320	26	3.8	3.8	19.2	34.6	38.5	0.0
5	Emerson Sensi	26	0.0	7.7	26.9	34.6	30.8	0.0
6	Venstar T7900	26	3.8	3.8	26.9	26.9	38.5	0.0

7	Trane XL-824	25	16.0	8.0	32.0	28.0	12.0	4.0
8	Lux GEO	25	28.0	8.0	32.0	8.0	16.0	8.0
9	Schneider Wiser Air	26	11.5	23.1	3.8	23.1	30.8	7.7
10	Carrier Cor	28	7.1	17.9	21.4	32.1	17.9	3.6
11	Allure Eversense	27	11.1	14.8	14.8	18.5	22.2	18.5
12	Radiostat CT-80	25	20.0	20.0	28.0	28.0	4.0	0.0

Meanings of words and symbols – Thermostat

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	15.4	30.8	26.9	15.4	11.5	0.0
2	Ecobee 3	26	15.4	19.2	11.5	30.8	19.2	3.8
3	Honeywell Lyric	24	25.0	16.7	33.3	16.7	4.2	4.2
4	Honeywell 9320	26	11.5	11.5	15.4	30.8	30.8	0.0
5	Emerson Sensi	26	0.0	7.7	30.8	23.1	38.5	0.0
6	Venstar T7900	26	0.0	11.5	19.2	15.4	46.2	7.7
7	Trane XL-824	25	12.0	16.0	24.0	36.0	12.0	0.0
8	Lux GEO	25	16.0	24.0	20.0	8.0	16.0	16.0
9	Schneider Wiser Air	26	15.4	7.7	23.1	34.6	15.4	3.8
10	Carrier Cor	28	3.6	25.0	28.6	28.6	14.3	0.0
11	Allure Eversense	27	7.4	11.1	37.0	14.8	22.2	7.4
12	Radiostat CT-80	25	8.0	32.0	24.0	36.0	0.0	0.0

Menu navigation (getting around) – Thermostat

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	30.8	23.1	23.1	15.4	7.7	0.0
2	Ecobee 3	26	19.2	23.1	23.1	19.2	15.4	0.0
3	Honeywell Lyric	24	54.2	20.8	20.8	4.2	0.0	0.0
4	Honeywell 9320	26	7.7	11.5	26.9	19.2	34.6	0.0
5	Emerson Sensi	26	7.7	19.2	23.1	15.4	34.6	0.0
6	Venstar T7900	26	7.7	11.5	11.5	19.2	46.2	3.8
7	Trane XL-824	25	28.0	12.0	28.0	20.0	12.0	0.0
8	Lux GEO	25	24.0	28.0	20.0	4.0	16.0	8.0
9	Schneider Wiser Air	26	15.4	15.4	23.1	26.9	15.4	3.8
10	Carrier Cor	28	17.9	21.4	25.0	32.1	3.6	0.0
11	Allure Eversense	27	22.2	11.1	18.5	33.3	11.1	3.7
12	Radiostat CT-80	25	16.0	44.0	8.0	32.0	0.0	0.0

Overall ease of use – Thermostat

ID	Thermostat	N	1	2	3	4	5	NA

1	Nest 3rd generation	26	23.1	38.5	15.4	19.2	3.8	0.0
2	Ecobee 3	26	19.2	15.4	23.1	30.8	11.5	0.0
3	Honeywell Lyric	24	29.2	29.2	33.3	8.3	0.0	0.0
4	Honeywell 9320	26	0.0	19.2	19.2	19.2	38.5	3.8
5	Emerson Sensi	26	3.8	19.2	19.2	23.1	30.8	3.8
6	Venstar T7900	26	3.8	11.5	15.4	15.4	46.2	7.7
7	Trane XL-824	25	16.0	16.0	20.0	36.0	12.0	0.0
8	Lux GEO	25	20.0	20.0	28.0	12.0	12.0	8.0
9	Schneider Wiser Air	26	15.4	11.5	26.9	23.1	19.2	3.8
10	Carrier Cor	28	14.3	21.4	21.4	39.3	3.6	0.0
11	Allure Eversense	27	11.1	14.8	22.2	37.0	7.4	7.4
12	Radiostat CT-80	25	24.0	24.0	16.0	32.0	0.0	4.0

Information on the home screen – App

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	15.4	30.8	15.4	23.1	7.7	7.7
2	Ecobee 3	26	7.7	23.1	23.1	23.1	15.4	7.7
3	Honeywell Lyric	24	12.5	33.3	16.7	16.7	12.5	8.3
4	Honeywell 9320	26	11.5	15.4	7.7	34.6	26.9	3.8
5	Emerson Sensi	26	7.7	19.2	26.9	30.8	7.7	7.7
6	Venstar T7900	26	15.4	15.4	19.2	15.4	34.6	0.0
7	Trane XL-824	25	32.0	12.0	28.0	12.0	8.0	8.0
8	Lux GEO	25	20.0	8.0	8.0	44.0	16.0	4.0
9	Schneider Wiser Air	26	15.4	42.3	7.7	7.7	26.9	0.0
10	Carrier Cor	28	10.7	10.7	35.7	28.6	10.7	3.6
11	Allure Eversense	27	11.1	7.4	33.3	18.5	18.5	11.1
12	Radiostat CT-80	25	4.0	8.0	36.0	32.0	12.0	8.0

Input mechanisms (buttons, dials, etc.) – App

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	19.2	19.2	19.2	19.2	15.4	7.7
2	Ecobee 3	26	7.7	7.7	26.9	34.6	15.4	7.7
3	Honeywell Lyric	24	12.5	37.5	20.8	20.8	4.2	4.2
4	Honeywell 9320	26	11.5	15.4	11.5	26.9	30.8	3.8
5	Emerson Sensi	26	11.5	19.2	30.8	34.6	0.0	3.8
6	Venstar T7900	26	11.5	15.4	19.2	15.4	30.8	7.7
7	Trane XL-824	25	24.0	20.0	28.0	16.0	8.0	4.0
8	Lux GEO	25	16.0	8.0	12.0	40.0	16.0	8.0
9	Schneider Wiser Air	26	23.1	19.2	19.2	15.4	23.1	0.0
10	Carrier Cor	28	7.1	10.7	32.1	25.0	17.9	7.1

11	Allure Eversense	27	3.7	22.2	29.6	18.5	18.5	7.4
12	Radiostat CT-80	25	4.0	20.0	12.0	44.0	12.0	8.0

Meanings of words and symbols – App

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	11.5	30.8	15.4	26.9	7.7	7.7
2	Ecobee 3	26	7.7	26.9	11.5	23.1	19.2	11.5
3	Honeywell Lyric	24	16.7	20.8	41.7	12.5	0.0	8.3
4	Honeywell 9320	26	11.5	3.8	30.8	19.2	30.8	3.8
5	Emerson Sensi	26	15.4	7.7	38.5	30.8	3.8	3.8
6	Venstar T7900	26	11.5	23.1	11.5	15.4	34.6	3.8
7	Trane XL-824	25	28.0	28.0	20.0	8.0	12.0	4.0
8	Lux GEO	25	12.0	12.0	20.0	32.0	20.0	4.0
9	Schneider Wiser Air	26	19.2	30.8	15.4	19.2	15.4	0.0
10	Carrier Cor	28	10.7	17.9	28.6	25.0	14.3	3.6
11	Allure Eversense	27	11.1	11.1	37.0	18.5	14.8	7.4
12	Radiostat CT-80	25	4.0	28.0	16.0	36.0	8.0	8.0

Menu navigation (getting around) – App

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	23.1	15.4	26.9	15.4	11.5	7.7
2	Ecobee 3	26	11.5	7.7	38.5	23.1	11.5	7.7
3	Honeywell Lyric	24	37.5	29.2	20.8	4.2	4.2	4.2
4	Honeywell 9320	26	15.4	11.5	11.5	34.6	23.1	3.8
5	Emerson Sensi	26	19.2	42.3	15.4	19.2	0.0	3.8
6	Venstar T7900	26	30.8	23.1	7.7	3.8	26.9	7.7
7	Trane XL-824	25	36.0	16.0	32.0	4.0	8.0	4.0
8	Lux GEO	25	12.0	20.0	20.0	20.0	24.0	4.0
9	Schneider Wiser Air	26	23.1	42.3	19.2	11.5	3.8	0.0
10	Carrier Cor	28	14.3	21.4	35.7	17.9	7.1	3.6
11	Allure Eversense	27	14.8	18.5	22.2	22.2	11.1	11.1
12	Radiostat CT-80	25	4.0	28.0	24.0	32.0	4.0	8.0

Overall ease of use – App

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	15.4	30.8	15.4	26.9	3.8	7.7
2	Ecobee 3	26	11.5	15.4	15.4	42.3	7.7	7.7
3	Honeywell Lyric	24	20.8	33.3	29.2	12.5	0.0	4.2
4	Honeywell 9320	26	11.5	15.4	15.4	19.2	26.9	11.5
5	Emerson Sensi	26	19.2	30.8	26.9	19.2	0.0	3.8

6	Venstar T7900	26	19.2	34.6	11.5	3.8	26.9	3.8
7	Trane XL-824	25	28.0	32.0	24.0	4.0	8.0	4.0
8	Lux GEO	25	16.0	12.0	12.0	40.0	16.0	4.0
9	Schneider Wiser Air	26	23.1	42.3	11.5	19.2	3.8	0.0
10	Carrier Cor	28	14.3	21.4	25.0	32.1	3.6	3.6
11	Allure Eversense	27	14.8	18.5	18.5	37.0	3.7	7.4
12	Radiostat CT-80	25	4.0	24.0	28.0	32.0	4.0	8.0

Survey Question 4: Feel and Sound

1 Lousy

3. Fine

5. Great!

Pressure needed for input – Thermostat

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	7.7	11.5	30.8	15.4	30.8	3.8
2	Ecobee 3	26	3.8	7.7	23.1	26.9	38.5	0.0
3	Honeywell Lyric	24	4.2	8.3	33.3	33.3	20.8	0.0
4	Honeywell 9320	26	3.8	11.5	30.8	7.7	46.2	0.0
5	Emerson Sensi	26	3.8	3.8	26.9	23.1	38.5	3.8
6	Venstar T7900	26	0.0	3.8	26.9	19.2	46.2	3.8
7	Trane XL-824	25	8.0	20.0	16.0	44.0	12.0	0.0
8	Lux GEO	25	20.0	4.0	36.0	12.0	24.0	4.0
9	Schneider Wiser Air	26	7.7	3.8	15.4	26.9	42.3	3.8
10	Carrier Cor	28	7.1	17.9	10.7	32.1	28.6	3.6
11	Allure Eversense	27	29.6	14.8	14.8	11.1	18.5	11.1
12	Radiostat CT-80	25	12.0	20.0	44.0	8.0	16.0	0.0

Responsiveness to input – Thermostat

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	0.0	11.5	26.9	30.8	26.9	3.8
2	Ecobee 3	26	7.7	7.7	26.9	26.9	30.8	0.0
3	Honeywell Lyric	24	4.2	12.5	37.5	25.0	20.8	0.0
4	Honeywell 9320	26	7.7	7.7	19.2	23.1	42.3	0.0
5	Emerson Sensi	26	0.0	3.8	26.9	19.2	46.2	3.8
6	Venstar T7900	26	0.0	7.7	23.1	19.2	50.0	0.0
7	Trane XL-824	25	12.0	20.0	16.0	40.0	12.0	0.0
8	Lux GEO	25	20.0	8.0	28.0	20.0	20.0	4.0
9	Schneider Wiser Air	26	7.7	11.5	11.5	26.9	38.5	3.8

10	Carrier Cor	28	7.1	21.4	14.3	32.1	25.0	0.0
11	Allure Eversense	27	29.6	14.8	11.1	18.5	18.5	7.4
12	Radiostat CT-80	25	12.0	20.0	32.0	28.0	8.0	0.0

Quality of the materials – Thermostat

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	3.8	7.7	26.9	15.4	38.5	7.7
2	Ecobee 3	26	11.5	3.8	15.4	42.3	26.9	0.0
3	Honeywell Lyric	24	8.3	8.3	25.0	45.8	12.5	0.0
4	Honeywell 9320	26	0.0	7.7	23.1	19.2	42.3	7.7
5	Emerson Sensi	26	0.0	7.7	30.8	19.2	38.5	3.8
6	Venstar T7900	26	0.0	3.8	26.9	38.5	19.2	11.5
7	Trane XL-824	25	16.0	8.0	20.0	28.0	24.0	4.0
8	Lux GEO	25	12.0	4.0	24.0	24.0	24.0	12.0
9	Schneider Wiser Air	26	7.7	7.7	11.5	15.4	50.0	7.7
10	Carrier Cor	28	0.0	10.7	14.3	46.4	25.0	3.6
11	Allure Eversense	27	3.7	3.7	18.5	29.6	33.3	11.1
12	Radiostat CT-80	25	8.0	20.0	28.0	28.0	8.0	8.0

Audible feedback – Thermostat

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	11.5	11.5	11.5	7.7	7.7	50.0
2	Ecobee 3	26	7.7	0.0	11.5	0.0	3.8	76.9
3	Honeywell Lyric	24	16.7	12.5	16.7	25.0	8.3	20.8
4	Honeywell 9320	26	3.8	7.7	15.4	3.8	11.5	57.7
5	Emerson Sensi	26	7.7	7.7	19.2	0.0	7.7	57.7
6	Venstar T7900	26	0.0	11.5	15.4	30.8	15.4	26.9
7	Trane XL-824	25	24.0	0.0	12.0	0.0	4.0	60.0
8	Lux GEO	25	12.0	12.0	20.0	0.0	8.0	48.0
9	Schneider Wiser Air	26	19.2	7.7	11.5	3.8	3.8	53.8
10	Carrier Cor	28	0.0	10.7	17.9	10.7	3.6	57.1
11	Allure Eversense	27	3.7	0.0	11.1	14.8	0.0	70.4
12	Radiostat CT-80	25	16.0	24.0	8.0	16.0	0.0	36.0

Overall feel and sound – Thermostat

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	7.7	11.5	30.8	26.9	15.4	7.7
2	Ecobee 3	26	3.8	7.7	23.1	38.5	15.4	11.5
3	Honeywell Lyric	24	8.3	12.5	37.5	25.0	12.5	4.2

4	Honeywell 9320	26	0.0	23.1	7.7	26.9	34.6	7.7
5	Emerson Sensi	26	0.0	0.0	38.5	19.2	30.8	11.5
6	Venstar T7900	26	3.8	3.8	19.2	26.9	23.1	23.1
7	Trane XL-824	25	24.0	8.0	24.0	28.0	8.0	8.0
8	Lux GEO	25	12.0	12.0	12.0	20.0	12.0	32.0
9	Schneider Wiser Air	26	7.7	15.4	19.2	15.4	34.6	7.7
10	Carrier Cor	28	3.6	7.1	32.1	39.3	10.7	7.1
11	Allure Eversense	27	7.4	11.1	25.9	29.6	18.5	7.4
12	Radiostat CT-80	25	8.0	28.0	24.0	32.0	4.0	4.0

Pressure needed for input – App

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	3.8	19.2	23.1	23.1	23.1	7.7
2	Ecobee 3	26	0.0	7.7	23.1	30.8	30.8	7.7
3	Honeywell Lyric	24	0.0	4.2	41.7	33.3	16.7	4.2
4	Honeywell 9320	26	11.5	11.5	11.5	11.5	50.0	3.8
5	Emerson Sensi	26	11.5	3.8	30.8	23.1	23.1	7.7
6	Venstar T7900	26	3.8	3.8	19.2	23.1	46.2	3.8
7	Trane XL-824	25	12.0	8.0	24.0	28.0	24.0	4.0
8	Lux GEO	25	12.0	0.0	20.0	28.0	36.0	4.0
9	Schneider Wiser Air	26	7.7	11.5	26.9	19.2	34.6	0.0
10	Carrier Cor	28	3.6	21.4	14.3	17.9	32.1	10.7
11	Allure Eversense	27	3.7	11.1	14.8	22.2	33.3	14.8
12	Radiostat CT-80	25	0.0	8.0	32.0	32.0	20.0	8.0

Responsiveness to input – App

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	3.8	11.5	30.8	30.8	15.4	7.7
2	Ecobee 3	26	3.8	7.7	26.9	30.8	23.1	7.7
3	Honeywell Lyric	24	0.0	8.3	41.7	20.8	25.0	4.2
4	Honeywell 9320	26	15.4	3.8	11.5	15.4	50.0	3.8
5	Emerson Sensi	26	11.5	7.7	30.8	15.4	30.8	3.8
6	Venstar T7900	26	7.7	11.5	19.2	23.1	30.8	7.7
7	Trane XL-824	25	12.0	8.0	24.0	28.0	24.0	4.0
8	Lux GEO	25	8.0	4.0	8.0	40.0	36.0	4.0
9	Schneider Wiser Air	26	3.8	34.6	7.7	15.4	38.5	0.0
10	Carrier Cor	28	7.1	21.4	21.4	21.4	25.0	3.6
11	Allure Eversense	27	7.4	11.1	14.8	22.2	33.3	11.1
12	Radiostat CT-80	25	0.0	8.0	28.0	32.0	24.0	8.0

Quality of the materials – App

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	7.7	11.5	26.9	19.2	23.1	11.5
2	Ecobee 3	26	7.7	3.8	26.9	19.2	26.9	15.4
3	Honeywell Lyric	24	4.2	4.2	33.3	37.5	12.5	8.3
4	Honeywell 9320	26	15.4	3.8	11.5	15.4	42.3	11.5
5	Emerson Sensi	26	7.7	3.8	34.6	19.2	26.9	7.7
6	Venstar T7900	26	3.8	3.8	34.6	19.2	26.9	11.5
7	Trane XL-824	25	16.0	12.0	16.0	20.0	24.0	12.0
8	Lux GEO	25	8.0	0.0	16.0	32.0	32.0	12.0
9	Schneider Wiser Air	26	7.7	23.1	7.7	11.5	42.3	7.7
10	Carrier Cor	28	0.0	14.3	17.9	32.1	28.6	7.1
11	Allure Eversense	27	3.7	7.4	11.1	18.5	37.0	22.2
12	Radiostat CT-80	25	0.0	0.0	40.0	32.0	16.0	12.0

Audible feedback – App

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	11.5	7.7	11.5	19.2	0.0	50.0
2	Ecobee 3	26	7.7	0.0	7.7	3.8	0.0	80.8
3	Honeywell Lyric	24	16.7	8.3	20.8	12.5	8.3	33.3
4	Honeywell 9320	26	15.4	15.4	0.0	3.8	11.5	53.8
5	Emerson Sensi	26	19.2	3.8	19.2	3.8	3.8	50.0
6	Venstar T7900	26	3.8	7.7	19.2	15.4	11.5	42.3
7	Trane XL-824	25	24.0	0.0	12.0	0.0	8.0	56.0
8	Lux GEO	25	8.0	8.0	0.0	8.0	16.0	60.0
9	Schneider Wiser Air	26	23.1	15.4	0.0	7.7	0.0	53.8
10	Carrier Cor	28	0.0	17.9	14.3	7.1	3.6	57.1
11	Allure Eversense	27	7.4	11.1	7.4	7.4	0.0	66.7
12	Radiostat CT-80	25	4.0	8.0	20.0	8.0	8.0	52.0

Overall feel and sound - App

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	7.7	19.2	23.1	23.1	19.2	7.7
2	Ecobee 3	26	0.0	3.8	11.5	46.2	15.4	23.1
3	Honeywell Lyric	24	4.2	12.5	41.7	25.0	12.5	4.2
4	Honeywell 9320	26	15.4	3.8	15.4	19.2	30.8	15.4
5	Emerson Sensi	26	7.7	11.5	38.5	7.7	15.4	19.2
6	Venstar T7900	26	7.7	7.7	19.2	15.4	26.9	23.1
7	Trane XL-824	25	20.0	8.0	12.0	28.0	16.0	16.0
8	Lux GEO	25	4.0	0.0	16.0	20.0	28.0	32.0

9	Schneider Wiser Air	26	7.7	26.9	23.1	7.7	26.9	7.7
10	Carrier Cor	28	0.0	17.9	17.9	39.3	14.3	10.7
11	Allure Eversense	27	11.1	3.7	25.9	25.9	18.5	14.8
12	Radiostat CT-80	25	4.0	12.0	28.0	48.0	0.0	8.0

Survey Question 5: Appearance

1 Lousy

3. Fine

5. Great!

Size and shape of the unit – Thermostat

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	0.0	3.8	46.2	19.2	30.8	0.0
2	Ecobee 3	26	7.7	3.8	26.9	26.9	34.6	0.0
3	Honeywell Lyric	24	12.5	0.0	16.7	45.8	25.0	0.0
4	Honeywell 9320	26	0.0	3.8	26.9	30.8	38.5	0.0
5	Emerson Sensi	26	0.0	19.2	34.6	26.9	19.2	0.0
6	Venstar T7900	26	3.8	0.0	19.2	26.9	46.2	3.8
7	Trane XL-824	25	12.0	12.0	16.0	28.0	32.0	0.0
8	Lux GEO	25	8.0	12.0	32.0	24.0	16.0	8.0
9	Schneider Wiser Air	26	0.0	7.7	3.8	26.9	57.7	3.8
10	Carrier Cor	28	0.0	7.1	25.0	39.3	28.6	0.0
11	Allure Eversense	27	3.7	22.2	18.5	18.5	33.3	3.7
12	Radiostat CT-80	25	0.0	20.0	40.0	24.0	16.0	0.0

Size of the screen – Thermostat

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	0.0	7.7	46.2	11.5	34.6	0.0
2	Ecobee 3	26	7.7	15.4	38.5	19.2	19.2	0.0
3	Honeywell Lyric	24	16.7	4.2	20.8	33.3	25.0	0.0
4	Honeywell 9320	26	0.0	7.7	23.1	23.1	46.2	0.0
5	Emerson Sensi	26	3.8	26.9	30.8	23.1	15.4	0.0
6	Venstar T7900	26	3.8	3.8	19.2	23.1	46.2	3.8
7	Trane XL-824	25	12.0	16.0	24.0	24.0	24.0	0.0
8	Lux GEO	25	8.0	16.0	20.0	24.0	24.0	8.0
9	Schneider Wiser Air	26	0.0	7.7	3.8	26.9	57.7	3.8
10	Carrier Cor	28	0.0	21.4	17.9	32.1	28.6	0.0
11	Allure Eversense	27	3.7	7.4	14.8	14.8	51.9	7.4

12	Radiostat CT-80	25	0.0	12.0	44.0	20.0	24.0	0.0

Color and style – Thermostat

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	3.8	11.5	30.8	15.4	38.5	0.0
2	Ecobee 3	26	7.7	11.5	26.9	23.1	30.8	0.0
3	Honeywell Lyric	24	8.3	12.5	20.8	29.2	29.2	0.0
4	Honeywell 9320	26	3.8	7.7	26.9	30.8	30.8	0.0
5	Emerson Sensi	26	3.8	26.9	53.8	3.8	7.7	3.8
6	Venstar T7900	26	3.8	3.8	30.8	30.8	30.8	0.0
7	Trane XL-824	25	8.0	20.0	24.0	20.0	20.0	8.0
8	Lux GEO	25	12.0	12.0	28.0	16.0	24.0	8.0
9	Schneider Wiser Air	26	0.0	7.7	11.5	19.2	57.7	3.8
10	Carrier Cor	28	0.0	3.6	25.0	39.3	32.1	0.0
11	Allure Eversense	27	0.0	3.7	18.5	22.2	51.9	3.7
12	Radiostat CT-80	25	16.0	20.0	28.0	20.0	12.0	4.0

Readability of smallest text – Thermostat

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	3.8	3.8	38.5	15.4	34.6	3.8
2	Ecobee 3	26	11.5	11.5	23.1	30.8	23.1	0.0
3	Honeywell Lyric	24	4.2	8.3	33.3	20.8	33.3	0.0
4	Honeywell 9320	26	7.7	7.7	23.1	23.1	38.5	0.0
5	Emerson Sensi	26	7.7	26.9	38.5	11.5	15.4	0.0
6	Venstar T7900	26	3.8	3.8	23.1	23.1	46.2	0.0
7	Trane XL-824	25	32.0	16.0	28.0	8.0	12.0	4.0
8	Lux GEO	25	0.0	4.0	24.0	32.0	28.0	12.0
9	Schneider Wiser Air	26	3.8	7.7	15.4	23.1	42.3	7.7
10	Carrier Cor	28	3.6	7.1	28.6	39.3	21.4	0.0
11	Allure Eversense	27	3.7	11.1	7.4	25.9	48.1	3.7
12	Radiostat CT-80	25	12.0	20.0	36.0	16.0	12.0	4.0

Overall appearance – Thermostat

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	0.0	3.8	38.5	23.1	34.6	0.0
2	Ecobee 3	26	7.7	11.5	26.9	23.1	30.8	0.0
3	Honeywell Lyric	24	12.5	4.2	20.8	33.3	29.2	0.0
4	Honeywell 9320	26	0.0	11.5	23.1	26.9	38.5	0.0
5	Emerson Sensi	26	0.0	19.2	57.7	15.4	7.7	0.0

6	Venstar T7900	26	3.8	0.0	19.2	42.3	34.6	0.0
7	Trane XL-824	25	16.0	12.0	36.0	12.0	20.0	4.0
8	Lux GEO	25	8.0	8.0	28.0	20.0	28.0	8.0
9	Schneider Wiser Air	26	0.0	7.7	7.7	34.6	42.3	7.7
10	Carrier Cor	28	0.0	3.6	32.1	35.7	28.6	0.0
11	Allure Eversense	27	0.0	0.0	25.9	14.8	55.6	3.7
12	Radiostat CT-80	25	4.0	20.0	48.0	12.0	16.0	0.0

Color and style – App

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	11.5	15.4	19.2	23.1	19.2	11.5
2	Ecobee 3	26	0.0	3.8	34.6	30.8	19.2	11.5
3	Honeywell Lyric	24	0.0	8.3	25.0	37.5	16.7	12.5
4	Honeywell 9320	26	7.7	7.7	11.5	42.3	23.1	7.7
5	Emerson Sensi	26	7.7	15.4	34.6	23.1	11.5	7.7
6	Venstar T7900	26	7.7	7.7	15.4	19.2	38.5	11.5
7	Trane XL-824	25	12.0	20.0	24.0	20.0	12.0	12.0
8	Lux GEO	25	4.0	8.0	24.0	28.0	32.0	4.0
9	Schneider Wiser Air	26	7.7	23.1	19.2	15.4	34.6	0.0
10	Carrier Cor	28	0.0	7.1	28.6	25.0	32.1	7.1
11	Allure Eversense	27	0.0	7.4	22.2	14.8	44.4	11.1
12	Radiostat CT-80	25	4.0	4.0	24.0	48.0	12.0	8.0

Readability of smallest text – App

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	26	19.2	19.2	15.4	23.1	15.4	7.7
2	Ecobee 3	26	3.8	3.8	26.9	30.8	23.1	11.5
3	Honeywell Lyric	24	0.0	12.5	16.7	29.2	29.2	12.5
4	Honeywell 9320	26	15.4	7.7	19.2	30.8	23.1	3.8
5	Emerson Sensi	26	11.5	30.8	26.9	11.5	15.4	3.8
6	Venstar T7900	26	3.8	23.1	15.4	26.9	26.9	3.8
7	Trane XL-824	25	12.0	12.0	32.0	24.0	12.0	8.0
8	Lux GEO	25	16.0	8.0	20.0	24.0	28.0	4.0
9	Schneider Wiser Air	26	11.5	26.9	19.2	11.5	30.8	0.0
10	Carrier Cor	28	0.0	17.9	17.9	32.1	25.0	7.1
11	Allure Eversense	27	7.4	3.7	11.1	33.3	37.0	7.4
12	Radiostat CT-80	25	4.0	4.0	36.0	36.0	8.0	12.0

Overall appearance – App

ID	Thermostat	N	1	2	3	4	5	NA

1	Nest 3rd generation	26	15.4	7.7	26.9	19.2	23.1	7.7
2	Ecobee 3	26	0.0	3.8	38.5	23.1	23.1	11.5
3	Honeywell Lyric	24	4.2	4.2	25.0	33.3	20.8	12.5
4	Honeywell 9320	26	3.8	7.7	19.2	30.8	34.6	3.8
5	Emerson Sensi	26	7.7	19.2	26.9	26.9	15.4	3.8
6	Venstar T7900	26	3.8	7.7	26.9	23.1	34.6	3.8
7	Trane XL-824	25	12.0	16.0	28.0	20.0	16.0	8.0
8	Lux GEO	25	8.0	12.0	16.0	24.0	36.0	4.0
9	Schneider Wiser Air	26	3.8	23.1	23.1	11.5	38.5	0.0
10	Carrier Cor	28	0.0	7.1	21.4	32.1	32.1	7.1
11	Allure Eversense	27	0.0	3.7	14.8	33.3	40.7	7.4
12	Radiostat CT-80	25	0.0	0.0	40.0	32.0	16.0	12.0

Survey Question 6: System Usability Score (Version 1)

1 Strongly Disagree

3 Neutral

5 Strongly Agrees

I think that I would like to use this thermostat for my home.

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	14	14.3	21.4	35.7	14.3	14.3	0.0
2	Ecobee 3	15	33.3	13.3	26.7	20.0	6.7	0.0
3	Honeywell Lyric	12	33.3	8.3	25.0	16.7	16.7	0.0
4	Honeywell 9320	15	20.0	13.3	6.7	20.0	40.0	0.0
5	Emerson Sensi	13	15.4	7.7	46.2	7.7	23.1	0.0
6	Venstar T7900	15	6.7	0.0	20.0	26.7	46.7	0.0
7	Trane XL-824	13	30.8	15.4	38.5	7.7	7.7	0.0
8	Lux GEO	13	46.2	7.7	23.1	15.4	0.0	7.7
9	Schneider Wiser Air	12	16.7	25.0	8.3	8.3	41.7	0.0
10	Carrier Cor	16	31.2	6.2	25.0	25.0	12.5	0.0
11	Allure Eversense	15	13.3	13.3	26.7	40.0	6.7	0.0
12	Radiostat CT-80	13	30.8	23.1	15.4	23.1	7.7	0.0

I found the thermostat unnecessarily complex.

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	14	7.1	21.4	28.6	14.3	28.6	0.0
2	Ecobee 3	15	13.3	20.0	26.7	20.0	20.0	0.0
3	Honeywell Lyric	12	0.0	16.7	41.7	25.0	16.7	0.0

4	Honeywell 9320	15	53.3	6.7	26.7	6.7	6.7	0.0
5	Emerson Sensi	13	30.8	23.1	30.8	0.0	15.4	0.0
6	Venstar T7900	15	33.3	26.7	20.0	13.3	6.7	0.0
7	Trane XL-824	13	15.4	46.2	7.7	23.1	7.7	0.0
8	Lux GEO	13	15.4	23.1	23.1	7.7	15.4	15.4
9	Schneider Wiser Air	12	33.3	16.7	25.0	16.7	8.3	0.0
10	Carrier Cor	16	37.5	18.8	31.2	6.2	6.2	0.0
11	Allure Eversense	15	26.7	20.0	46.7	0.0	6.7	0.0
12	Radiostat CT-80	13	7.7	23.1	7.7	15.4	46.2	0.0

I thought the thermostat was easy to use.

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	14	21.4	7.1	35.7	21.4	14.3	0.0
2	Ecobee 3	15	33.3	6.7	20.0	40.0	0.0	0.0
3	Honeywell Lyric	12	33.3	8.3	33.3	25.0	0.0	0.0
4	Honeywell 9320	15	13.3	0.0	26.7	20.0	40.0	0.0
5	Emerson Sensi	13	15.4	0.0	23.1	30.8	30.8	0.0
6	Venstar T7900	15	6.7	0.0	0.0	46.7	46.7	0.0
7	Trane XL-824	13	15.4	23.1	23.1	30.8	7.7	0.0
8	Lux GEO	13	30.8	7.7	23.1	30.8	0.0	7.7
9	Schneider Wiser Air	12	25.0	0.0	16.7	33.3	25.0	0.0
10	Carrier Cor	16	18.8	12.5	18.8	37.5	12.5	0.0
11	Allure Eversense	15	6.7	20.0	33.3	33.3	6.7	0.0
12	Radiostat CT-80	13	38.5	23.1	15.4	7.7	15.4	0.0

I think that I would need the support of a technical person to be able to use this thermostat.

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	14	35.7	7.1	28.6	14.3	14.3	0.0
2	Ecobee 3	15	40.0	6.7	13.3	26.7	13.3	0.0
3	Honeywell Lyric	12	16.7	33.3	25.0	16.7	8.3	0.0
4	Honeywell 9320	15	53.3	6.7	33.3	0.0	6.7	0.0
5	Emerson Sensi	13	46.2	23.1	23.1	0.0	7.7	0.0
6	Venstar T7900	15	73.3	20.0	0.0	0.0	6.7	0.0
7	Trane XL-824	13	23.1	23.1	30.8	15.4	7.7	0.0
8	Lux GEO	13	23.1	7.7	15.4	7.7	38.5	7.7
9	Schneider Wiser Air	12	25.0	33.3	16.7	8.3	16.7	0.0
10	Carrier Cor	16	50.0	18.8	0.0	6.2	25.0	0.0
11	Allure Eversense	15	53.3	13.3	20.0	13.3	0.0	0.0

12	Radiostat CT-80	13	30.8	15.4	30.8	0.0	23.1	0.0

I found the various functions in this thermostat were well integrated.

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	14	0.0	35.7	42.9	7.1	14.3	0.0
2	Ecobee 3	15	13.3	6.7	26.7	20.0	33.3	0.0
3	Honeywell Lyric	12	25.0	16.7	58.3	0.0	0.0	0.0
4	Honeywell 9320	15	6.7	6.7	20.0	26.7	40.0	0.0
5	Emerson Sensi	13	7.7	15.4	30.8	38.5	7.7	0.0
6	Venstar T7900	15	6.7	6.7	26.7	20.0	40.0	0.0
7	Trane XL-824	13	7.7	15.4	61.5	7.7	7.7	0.0
8	Lux GEO	13	23.1	7.7	30.8	23.1	7.7	7.7
9	Schneider Wiser Air	12	8.3	8.3	41.7	8.3	25.0	8.3
10	Carrier Cor	16	0.0	31.2	37.5	18.8	12.5	0.0
11	Allure Eversense	15	6.7	6.7	40.0	46.7	0.0	0.0
12	Radiostat CT-80	13	30.8	15.4	38.5	7.7	7.7	0.0

I thought there was too much inconsistency in this thermostat.

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	14	14.3	7.1	64.3	7.1	7.1	0.0
2	Ecobee 3	15	20.0	13.3	40.0	13.3	13.3	0.0
3	Honeywell Lyric	12	8.3	25.0	58.3	8.3	0.0	0.0
4	Honeywell 9320	15	46.7	26.7	20.0	0.0	0.0	6.7
5	Emerson Sensi	13	30.8	30.8	30.8	0.0	7.7	0.0
6	Venstar T7900	15	73.3	20.0	6.7	0.0	0.0	0.0
7	Trane XL-824	13	15.4	30.8	38.5	7.7	7.7	0.0
8	Lux GEO	13	30.8	15.4	15.4	7.7	23.1	7.7
9	Schneider Wiser Air	12	33.3	33.3	25.0	0.0	8.3	0.0
10	Carrier Cor	16	31.2	31.2	25.0	6.2	6.2	0.0
11	Allure Eversense	15	26.7	33.3	26.7	13.3	0.0	0.0
12	Radiostat CT-80	13	23.1	7.7	30.8	15.4	15.4	7.7

I would imagine that most people would learn to use this thermostat very quickly.

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	14	7.1	14.3	21.4	35.7	21.4	0.0
2	Ecobee 3	15	20.0	0.0	6.7	53.3	13.3	6.7
3	Honeywell Lyric	12	16.7	25.0	16.7	25.0	8.3	8.3
4	Honeywell 9320	15	0.0	20.0	13.3	33.3	33.3	0.0
5	Emerson Sensi	13	7.7	0.0	30.8	30.8	30.8	0.0
6	Venstar T7900	15	0.0	6.7	0.0	40.0	53.3	0.0

7	Trane XL-824	13	15.4	15.4	46.2	15.4	7.7	0.0
8	Lux GEO	13	7.7	15.4	23.1	23.1	23.1	7.7
9	Schneider Wiser Air	12	0.0	25.0	25.0	16.7	33.3	0.0
10	Carrier Cor	16	6.2	12.5	43.8	18.8	18.8	0.0
11	Allure Eversense	15	6.7	13.3	33.3	26.7	20.0	0.0
12	Radiostat CT-80	13	30.8	0.0	30.8	30.8	7.7	0.0

I found the thermostat very cumbersome to use.

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	14	14.3	7.1	57.1	14.3	7.1	0.0
2	Ecobee 3	15	20.0	13.3	20.0	20.0	26.7	0.0
3	Honeywell Lyric	12	8.3	33.3	41.7	16.7	0.0	0.0
4	Honeywell 9320	15	60.0	13.3	13.3	13.3	0.0	0.0
5	Emerson Sensi	13	46.2	15.4	23.1	0.0	15.4	0.0
6	Venstar T7900	15	40.0	40.0	20.0	0.0	0.0	0.0
7	Trane XL-824	13	15.4	15.4	38.5	23.1	7.7	0.0
8	Lux GEO	13	23.1	15.4	30.8	7.7	15.4	7.7
9	Schneider Wiser Air	12	16.7	33.3	16.7	25.0	8.3	0.0
10	Carrier Cor	16	18.8	43.8	31.2	0.0	6.2	0.0
11	Allure Eversense	15	40.0	6.7	46.7	6.7	0.0	0.0
12	Radiostat CT-80	13	15.4	23.1	15.4	30.8	15.4	0.0

I felt very confident using the thermostat.

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	14	14.3	42.9	14.3	14.3	14.3	0.0
2	Ecobee 3	15	26.7	6.7	6.7	46.7	13.3	0.0
3	Honeywell Lyric	12	25.0	16.7	58.3	0.0	0.0	0.0
4	Honeywell 9320	15	13.3	6.7	13.3	26.7	40.0	0.0
5	Emerson Sensi	13	15.4	7.7	23.1	23.1	30.8	0.0
6	Venstar T7900	15	0.0	6.7	6.7	33.3	53.3	0.0
7	Trane XL-824	13	15.4	15.4	30.8	30.8	7.7	0.0
8	Lux GEO	13	23.1	23.1	7.7	15.4	23.1	7.7
9	Schneider Wiser Air	12	16.7	25.0	8.3	0.0	50.0	0.0
10	Carrier Cor	16	18.8	12.5	18.8	43.8	6.2	0.0
11	Allure Eversense	15	0.0	20.0	40.0	33.3	6.7	0.0
12	Radiostat CT-80	13	30.8	23.1	15.4	7.7	23.1	0.0

I needed to learn a lot of things before I could get going with this thermostat.

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	14	14.3	14.3	28.6	21.4	21.4	0.0
2	Ecobee 3	15	40.0	0.0	13.3	6.7	40.0	0.0
3	Honeywell Lyric	12	16.7	8.3	16.7	16.7	41.7	0.0
4	Honeywell 9320	15	53.3	20.0	13.3	6.7	6.7	0.0
5	Emerson Sensi	13	38.5	23.1	15.4	0.0	23.1	0.0
6	Venstar T7900	15	66.7	20.0	13.3	0.0	0.0	0.0
7	Trane XL-824	13	30.8	30.8	7.7	7.7	23.1	0.0
8	Lux GEO	13	30.8	15.4	15.4	0.0	30.8	7.7
9	Schneider Wiser Air	12	33.3	33.3	8.3	0.0	25.0	0.0
10	Carrier Cor	16	31.2	6.2	25.0	18.8	18.8	0.0
11	Allure Eversense	15	26.7	20.0	20.0	26.7	6.7	0.0
12	Radiostat CT-80	13	30.8	7.7	7.7	23.1	30.8	0.0

Survey Question 6: System Usability Score (Version 2)

1 Strongly Disagree

3 Neutral

5 Strongly Agrees

I think that I would like to use this thermostat/app system for my home.

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	12	50.0	16.7	8.3	8.3	16.7	0.0
2	Ecobee 3	11	18.2	18.2	36.4	18.2	9.1	0.0
3	Honeywell Lyric	12	41.7	16.7	25.0	8.3	8.3	0.0
4	Honeywell 9320	11	9.1	9.1	27.3	18.2	36.4	0.0
5	Emerson Sensi	13	30.8	15.4	23.1	15.4	15.4	0.0
6	Venstar T7900	11	27.3	0.0	18.2	18.2	27.3	9.1
7	Trane XL-824	12	58.3	25.0	8.3	8.3	0.0	0.0
8	Lux GEO	12	33.3	8.3	33.3	8.3	16.7	0.0
9	Schneider Wiser Air	14	35.7	21.4	0.0	14.3	28.6	0.0
10	Carrier Cor	12	25.0	25.0	16.7	8.3	16.7	8.3
11	Allure Eversense	12	50.0	0.0	8.3	8.3	33.3	0.0
12	Radiostat CT-80	12	41.7	16.7	33.3	0.0	8.3	0.0

I found the thermostat/app system unnecessarily complex.

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	12	25.0	25.0	0.0	16.7	33.3	0.0
2	Ecobee 3	11	18.2	36.4	18.2	27.3	0.0	0.0

3	Honeywell Lyric	12	8.3	8.3	16.7	16.7	50.0	0.0
4	Honeywell 9320	11	45.5	0.0	9.1	45.5	0.0	0.0
5	Emerson Sensi	13	23.1	23.1	7.7	23.1	23.1	0.0
6	Venstar T7900	11	18.2	18.2	18.2	18.2	18.2	9.1
7	Trane XL-824	12	8.3	25.0	8.3	16.7	41.7	0.0
8	Lux GEO	12	16.7	16.7	8.3	16.7	41.7	0.0
9	Schneider Wiser Air	14	21.4	21.4	7.1	14.3	35.7	0.0
10	Carrier Cor	12	16.7	0.0	33.3	8.3	33.3	8.3
11	Allure Eversense	12	41.7	16.7	8.3	8.3	25.0	0.0
12	Radiostat CT-80	12	16.7	16.7	16.7	41.7	8.3	0.0

I thought the thermostat/app system was easy to use.

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	12	33.3	25.0	8.3	16.7	16.7	0.0
2	Ecobee 3	11	9.1	36.4	18.2	36.4	0.0	0.0
3	Honeywell Lyric	12	50.0	16.7	25.0	8.3	0.0	0.0
4	Honeywell 9320	11	0.0	18.2	27.3	27.3	27.3	0.0
5	Emerson Sensi	13	46.2	7.7	30.8	7.7	7.7	0.0
6	Venstar T7900	11	18.2	9.1	18.2	27.3	18.2	9.1
7	Trane XL-824	12	41.7	25.0	8.3	16.7	8.3	0.0
8	Lux GEO	12	41.7	16.7	16.7	8.3	16.7	0.0
9	Schneider Wiser Air	14	35.7	7.1	14.3	35.7	7.1	0.0
10	Carrier Cor	12	25.0	25.0	16.7	8.3	16.7	8.3
11	Allure Eversense	12	50.0	0.0	0.0	16.7	33.3	0.0
12	Radiostat CT-80	12	0.0	41.7	25.0	16.7	16.7	0.0

I think that I would need the support of a technical person to be able to use this thermostat/app system.

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	12	16.7	8.3	8.3	16.7	50.0	0.0
2	Ecobee 3	11	36.4	18.2	9.1	18.2	18.2	0.0
3	Honeywell Lyric	12	16.7	25.0	16.7	0.0	41.7	0.0
4	Honeywell 9320	11	45.5	0.0	27.3	27.3	0.0	0.0
5	Emerson Sensi	13	38.5	7.7	7.7	15.4	30.8	0.0
6	Venstar T7900	11	36.4	9.1	9.1	9.1	27.3	9.1
7	Trane XL-824	12	8.3	0.0	16.7	16.7	58.3	0.0
8	Lux GEO	12	41.7	8.3	0.0	0.0	50.0	0.0
9	Schneider Wiser Air	14	14.3	14.3	14.3	14.3	42.9	0.0

10	Carrier Cor	12	25.0	8.3	8.3	16.7	33.3	8.3
11	Allure Eversense	12	58.3	0.0	0.0	25.0	16.7	0.0
12	Radiostat CT-80	12	33.3	8.3	33.3	0.0	25.0	0.0

I found the various functions in this thermostat/app system were well integrated.

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	12	16.7	16.7	50.0	8.3	8.3	0.0
2	Ecobee 3	11	0.0	9.1	36.4	45.5	9.1	0.0
3	Honeywell Lyric	12	25.0	25.0	25.0	25.0	0.0	0.0
4	Honeywell 9320	11	0.0	0.0	27.3	45.5	27.3	0.0
5	Emerson Sensi	13	7.7	0.0	53.8	30.8	7.7	0.0
6	Venstar T7900	11	9.1	18.2	27.3	18.2	18.2	9.1
7	Trane XL-824	12	25.0	25.0	41.7	8.3	0.0	0.0
8	Lux GEO	12	25.0	33.3	25.0	8.3	8.3	0.0
9	Schneider Wiser Air	14	7.1	28.6	21.4	35.7	7.1	0.0
10	Carrier Cor	12	16.7	16.7	41.7	0.0	16.7	8.3
11	Allure Eversense	12	41.7	0.0	25.0	25.0	8.3	0.0
12	Radiostat CT-80	12	8.3	33.3	41.7	8.3	8.3	0.0

I thought there was too much inconsistency in this thermostat/app system.

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	12	8.3	8.3	41.7	25.0	16.7	0.0
2	Ecobee 3	11	27.3	36.4	18.2	9.1	9.1	0.0
3	Honeywell Lyric	12	25.0	16.7	16.7	8.3	33.3	0.0
4	Honeywell 9320	11	45.5	18.2	9.1	27.3	0.0	0.0
5	Emerson Sensi	13	15.4	7.7	46.2	15.4	15.4	0.0
6	Venstar T7900	11	27.3	27.3	27.3	9.1	0.0	9.1
7	Trane XL-824	12	0.0	16.7	41.7	25.0	16.7	0.0
8	Lux GEO	12	16.7	8.3	25.0	8.3	41.7	0.0
9	Schneider Wiser Air	14	14.3	42.9	28.6	14.3	0.0	0.0
10	Carrier Cor	12	8.3	16.7	25.0	16.7	25.0	8.3
11	Allure Eversense	12	41.7	8.3	8.3	16.7	25.0	0.0
12	Radiostat CT-80	12	8.3	16.7	50.0	8.3	16.7	0.0

I would imagine that most people would learn to use this thermostat/app system very quickly.

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	12	16.7	33.3	0.0	33.3	16.7	0.0
2	Ecobee 3	11	9.1	27.3	0.0	54.5	9.1	0.0
3	Honeywell Lyric	12	25.0	8.3	50.0	8.3	8.3	0.0

4	Honeywell 9320	11	9.1	27.3	27.3	9.1	27.3	0.0
5	Emerson Sensi	13	7.7	15.4	38.5	15.4	23.1	0.0
6	Venstar T7900	11	0.0	27.3	18.2	27.3	18.2	9.1
7	Trane XL-824	12	25.0	33.3	25.0	16.7	0.0	0.0
8	Lux GEO	12	41.7	25.0	16.7	8.3	8.3	0.0
9	Schneider Wiser Air	14	35.7	7.1	21.4	21.4	14.3	0.0
10	Carrier Cor	12	25.0	25.0	0.0	33.3	8.3	8.3
11	Allure Eversense	12	16.7	8.3	33.3	25.0	16.7	0.0
12	Radiostat CT-80	12	8.3	41.7	8.3	41.7	0.0	0.0

I found the thermostat/app system very cumbersome to use.

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	12	16.7	8.3	16.7	33.3	25.0	0.0
2	Ecobee 3	11	54.5	9.1	18.2	9.1	9.1	0.0
3	Honeywell Lyric	12	8.3	41.7	16.7	25.0	8.3	0.0
4	Honeywell 9320	11	45.5	18.2	27.3	9.1	0.0	0.0
5	Emerson Sensi	13	30.8	23.1	23.1	15.4	7.7	0.0
6	Venstar T7900	11	18.2	18.2	18.2	18.2	18.2	9.1
7	Trane XL-824	12	0.0	16.7	25.0	33.3	25.0	0.0
8	Lux GEO	12	8.3	16.7	25.0	8.3	41.7	0.0
9	Schneider Wiser Air	14	14.3	14.3	21.4	28.6	21.4	0.0
10	Carrier Cor	12	16.7	33.3	0.0	25.0	16.7	8.3
11	Allure Eversense	12	25.0	16.7	16.7	16.7	25.0	0.0
12	Radiostat CT-80	12	16.7	16.7	16.7	33.3	16.7	0.0

I felt very confident using the thermostat/app system.

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	12	33.3	33.3	8.3	16.7	8.3	0.0
2	Ecobee 3	11	9.1	18.2	36.4	36.4	0.0	0.0
3	Honeywell Lyric	12	41.7	0.0	41.7	8.3	0.0	8.3
4	Honeywell 9320	11	0.0	36.4	0.0	27.3	36.4	0.0
5	Emerson Sensi	13	23.1	30.8	15.4	7.7	23.1	0.0
6	Venstar T7900	11	27.3	9.1	18.2	9.1	27.3	9.1
7	Trane XL-824	12	41.7	25.0	8.3	16.7	8.3	0.0
8	Lux GEO	12	50.0	0.0	16.7	8.3	25.0	0.0
9	Schneider Wiser Air	14	35.7	14.3	14.3	28.6	7.1	0.0
10	Carrier Cor	12	33.3	25.0	16.7	8.3	8.3	8.3
11	Allure Eversense	12	41.7	16.7	0.0	16.7	25.0	0.0
12	Radiostat CT-80	12	8.3	41.7	16.7	16.7	16.7	0.0

I needed to learn a lot of things before I could get going with this thermostat/app system.

ID	Thermostat	N	1	2	3	4	5	NA
1	Nest 3rd generation	12	8.3	16.7	16.7	8.3	50.0	0.0
2	Ecobee 3	11	18.2	36.4	27.3	0.0	18.2	0.0
3	Honeywell Lyric	12	8.3	8.3	16.7	16.7	50.0	0.0
4	Honeywell 9320	11	45.5	18.2	9.1	18.2	9.1	0.0
5	Emerson Sensi	13	30.8	38.5	0.0	7.7	23.1	0.0
6	Venstar T7900	11	36.4	18.2	9.1	9.1	18.2	9.1
7	Trane XL-824	12	0.0	25.0	8.3	33.3	33.3	0.0
8	Lux GEO	12	16.7	16.7	0.0	16.7	50.0	0.0
9	Schneider Wiser Air	14	14.3	21.4	14.3	21.4	28.6	0.0
10	Carrier Cor	12	8.3	25.0	8.3	16.7	33.3	8.3
11	Allure Eversense	12	50.0	16.7	0.0	0.0	33.3	0.0
12	Radiostat CT-80	12	25.0	8.3	25.0	16.7	25.0	0.0

Survey Question 7: Preference of Two Thermostats Tested

Imagine that the thermostat in your home dies and your mechanic offers a choice between the two thermostats you just reviewed – at the same price. Please circle the thermostat you would choose to have installed.

ID	Thermostat	N	1
1	Nest 3rd generation	26	54%
2	Ecobee 3	26	54%
3	Honeywell Lyric	24	25%
4	Honeywell 9320	26	85%
5	Emerson Sensi	26	31%
6	Venstar T7900	26	65%
7	Trane XL-824	25	24%
8	Lux GEO	25	32%
9	Schneider Wiser Air	26	38%
10	Carrier Cor	28	43%
11	Allure Eversense	27	52%
12	Radiostat CT-80	25	28%

Survey Question 8: Advanced Features

1 No way

3 Maybe

5 Definitely!

Do you think you would find the following features useful on a thermostat in your home?

Feature	N	1	2	3	4	5	NA
Auto Away	155	9.0	3.2	18.7	14.8	50.3	3.9
Auto Schedule	155	11.6	7.7	22.6	20.6	32.9	4.5
Color Display	155	3.9	8.4	23.9	27.1	32.9	3.9
Efficiency Indicator	155	2.6	3.2	17.4	23.9	49.0	3.9
Fresh Air	155	3.2	2.6	8.4	18.7	65.2	1.9
Home Energy Display	155	2.6	3.9	9.7	18.7	62.6	2.6
Online Account	155	10.3	7.1	15.5	17.4	47.1	2.6
Outdoor Temperature	155	3.2	5.2	7.7	26.5	55.5	1.9
Parental Controls	155	12.3	12.3	20.6	21.9	27.1	5.8
Precool	155	4.5	5.2	14.8	29.7	43.2	2.6
Price Response	155	4.5	6.5	18.7	23.2	42.6	4.5
Proximity	155	14.2	11.6	25.2	12.9	30.3	5.8
Smartphone App	155	9.7	7.1	12.9	12.3	55.5	2.6
Time to Temperature	155	5.8	5.2	20.0	25.8	40.6	2.6
Touchscreen	155	3.2	3.2	16.1	21.9	49.0	6.5

APPENDIX B – REGRESSION MODEL DETAILS

DATA DICTIONARY

TABLE 20. DATA DICTIONARY FOR VARIABLES INCLUDED IN THE REGRESSION MODELS

	Variable	Description	Data type
A	**Preference**	Thermostat was chosen as the better of 2 units tested	Boolean
B	**Efficiency**	Task Efficiency (see Equation 1)	Continuous
1	Tstat_Size	Size of the thermostat (cm^2)	Continuous
2	Tstat_Screen_Size	Screen size (cm^2)	Continuous
3	Tstat_Screen_Color	Color display screen (more than 2 colors)	Boolean
4	Tstat_Input_Buttons	Thermostat has differentiated buttons for input	Boolean
5	Tstat_Input_Touchscreen	Thermostat has a touchscreen for input	Boolean
6	Tstat_Input_Dial	Thermostat has a dial for input	Boolean
7	Tstat_Button.Size_Menu	Size of the thermostat menu button/indicator (cm^2)	Continuous
8	Tstat_Text.size_Menu	Height of the main menu text (mm)	Continuous
9	Tstat_Text.size_Smallest	Height of the smallest text (mm)	Continuous
10	Tstat_Buttons_Help	Help button is available	Boolean
11	Tstat_Buttons_Home	Home button is available	Boolean
12	Tstat_Buttons_Back	Back button is available	Boolean
13	Tstat_Buttons_Done	Button to confirm chosen setting	Boolean
14	Tstat_Symbols_Labels	Symbols are labeled with text	Boolean
15	Tstat_Volume	Thermostat has input sounds and volume control	Boolean
16	Tstat_Steps_Mode	Number of steps needed to set thermostat to cool	Integer
17	App_Consistent	App looks and functions just like the thermostat	Boolean
18	Participant_Age	Participant age (years)	Integer
19	Participant_Renter	Participant rents their home	Boolean
20	Participant_ComputerIQ	Participant self-rated confidence using a computer	5-point Likert
21	Participant_SmartphoneIQ	Participant self-rated confidence using a smartphone	5-point Likert
22	Tstat_Rated_Ease.of.Use	Thermostat survey question 3e: Overall ease of use	5-point Likert
23	Tstat_Rated_Feel.Sound	Thermostat survey question 4e: Overall feel and sound	5-point Likert
24	Tstat_Rated_Appearance	Thermostat survey question 5e: Overall appearance	5-point Likert
25	App_Rated_Ease.of.Use	App survey question 3e: Overall ease of use	5-point Likert
26	App_Rated_Appearance	App survey question 5e: Overall appearance	5-point Likert

CORRELATIONS

TABLE 21. CORRELATIONS OF INDEPENDENT AND DEPENDENT VARIABLES

	Variable	Preference Correlation	Preference P-value	Efficiency Correlation	Efficiency P-value
A	**Preference**	1.00	--	0.16	**0.0040**
B	**Efficiency**	0.16	**0.0040**	1.00	--
1	**Tstat_Size**	-0.02	0.7159	0.16	**0.0057**
2	**Tstat_Screen_Size**	0.13	**0.0186**	0.20	**0.0004**
3	**Tstat_Screen_Color**	0.16	**0.0039**	0.04	0.4880
4	**Tstat_Input_Buttons**	-0.19	**0.0009**	0.03	0.6611
5	**Tstat_Input_Touchscreen**	0.12	**0.0289**	0.12	**0.0350**
6	**Tstat_Input_Dial**	-0.08	0.1515	-0.19	**0.0006**
7	**Tstat_Button.Size_Menu**	-0.06	0.3262	-0.12	**0.0339**
8	**Tstat_Text.size_Menu**	0.12	**0.0426**	-0.08	0.1635
9	Tstat_Text.size_Smallest	-0.01	0.7954	-0.10	0.0823
10	**Tstat_Buttons_Help**	0.18	**0.0018**	0.11	**0.0470**
11	Tstat_Buttons_Home	0.04	0.4945	0.07	0.2126
12	Tstat_Buttons_Back	0.09	0.1221	0.03	0.5523
13	**Tstat_Buttons_Done**	0.09	0.1032	0.12	**0.0407**
14	**Tstat_Symbols_Labels**	0.08	0.1515	0.16	**0.0054**
15	Tstat_Volume	-0.01	0.8154	-0.01	0.8083
16	**Tstat_Steps_Mode**	0.05	0.3657	-0.14	**0.0134**
17	App_Consistent	0.07	0.2533	-0.09	0.0965
18	**Participant_Age**	-0.07	0.1926	-0.60	**0.0000**
19	**Participant_Renter**	-0.01	0.9219	0.22	**0.0001**
20	**Participant_ComputerIQ**	0.06	0.2668	0.46	**0.0000**
21	**Participant_SmartphoneIQ**	0.08	0.1729	0.57	**0.0000**
22	**Tstat_Rated_Ease.of.Use**	0.38	**0.0000**	0.46	**0.0000**
23	**Tstat_Rated_Feel.and.Sound**	0.34	**0.0000**	0.30	**0.0000**
24	**Tstat_Rated_Appearance**	0.36	**0.0000**	0.09	0.1352
25	**App_Rated_Ease.of.Use**	0.21	**0.0003**	0.41	**0.0000**
26	**App_Rated_Appearance**	0.13	**0.0269**	0.22	**0.0001**

PREFERENCE MODEL

TABLE 22. CORRELATIONS FOR VARIABLES IN THE PREFERENCE MODEL

	Preference	Tstat_Input_Touchscreen	Tstat_Screen_Color	Tstat_Screen_Size	Tstat_Input_Buttons	Tstat_Buttons_Help	Tstat_Text.size_Menu	App_Rated_Ease.of.Use	Tstat_Rated_Ease.of.Use	Tstat_Rated_Feel.Sound	Tstat_Rated_Appearance	App_Rated_Appearance
Preference	1.00											
Tstat_Input_Touchscreen	0.12	1.00										
Tstat_Screen_Color	0.16	0.42	1.00									
Tstat_Screen_Size	0.13	0.81	0.22	1.00								
Tstat_Input_Buttons	-0.19	-0.41	-0.57	-0.37	1.00							
Tstat_Buttons_Help	0.18	0.41	0.34	0.35	-0.34	1.00						
Tstat_Text.size_Menu	0.12	-0.14	-0.32	-0.13	-0.11	0.18	1.00					
App_Rated_Ease.of.Use	0.21	0.06	-0.05	0.05	-0.09	0.12	0.22	1.00				
Tstat_Rated_Ease.of.Use	0.38	0.17	0.04	0.27	-0.11	0.11	-0.10	0.44	1.00			
Tstat_Rated_Feel.Sound	0.34	0.01	0.04	0.02	-0.04	0.07	-0.03	0.33	0.63	1.00		
Tstat_Rated_Appearance	0.36	0.10	0.22	0.04	-0.21	0.18	0.07	0.17	0.37	0.55	1.00	
App_Rated_Appearance	0.13	0.11	0.06	0.05	-0.06	0.18	0.10	0.58	0.26	0.36	0.41	1.00

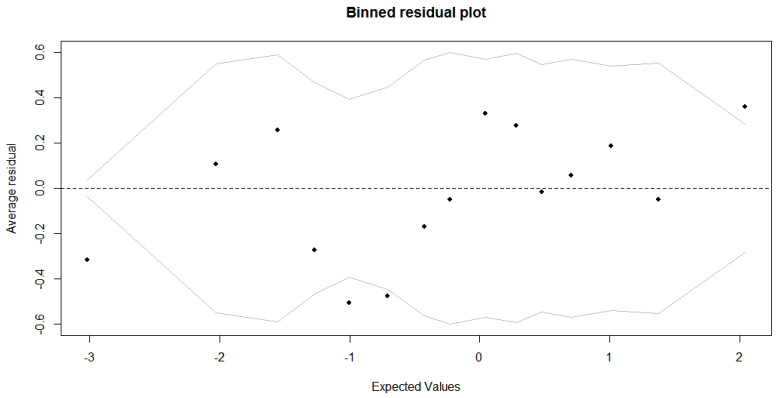
FIGURE 28. PREFERENCE MODEL COEFFICIENTS AND SIGNIFICANCE

Variable	Coefficient	Std. Error	z value	Pr(>|z|)	[1]
(Intercept)	-6.2344	1.1754	-5.30	0.0000	***
Tstat_Screen_Color	0.5324	0.4231	1.26	0.2083	
Tstat_Screen_Size	0.0078	0.0138	0.56	0.5734	
Tstat_Buttons_Help	0.3689	0.3847	0.96	0.3376	
Tstat_Text_Menu.Size	0.7012	0.2762	2.54	0.0111	*
Tstat_Rated_Ease.of.Use	0.4032	0.1729	2.33	0.0197	*
Tstat_Rated_Feel.and.Sound	0.1734	0.1842	0.94	0.3465	
Tstat_Rated_Appearance	0.5766	0.1840	3.13	0.0017	**
App_Rated_Ease.of.Use	0.2017	0.1653	1.22	0.2222	
App_Rated_Appearance	-0.2561	0.1787	-1.43	0.1520	

[1] Statistical Significance codes: ***: α=0.001; **: α=0.01; *: α=0.05; .: α=0.1

AIC: 297.67
Number of Fisher Scoring iterations: 4

FIGURE 29. PREFERENCE MODEL – BINNED RESIDUALS



EFFICIENCY MODEL

TABLE 23. CORRELATIONS FOR VARIABLES IN THE EFFICIENCY MODEL (PART 1)

	Efficiency	Tstat_Input_Touchscreen	Tstat_Screen_Color	Tstat_Input_Dial	Tstat_Screen_Size	Tstat_Size	Tstat_Buttons_Help	Tstat_Symbols_Labels	Tstat_Buttons_Done	Tstat_Button.Size_Menu
Efficiency	1.00									
Tstat_Input_Touchscreen	0.12	1.00								
Tstat_Screen_Color	0.04	0.42	1.00							
Tstat_Input_Dial	-0.19	**-0.81**	-0.12	1.00						
Tstat_Screen_Size	0.20	**0.81**	0.22	-0.80	1.00					
Tstat_Size	0.16	0.43	-0.3	-0.61	0.60	1.00				
Tstat_Buttons_Help	0.11	0.41	0.34	-0.34	0.35	0.30	1.00			
Tstat_Symbols_Labels	0.16	0.39	0.12	-0.54	0.61	0.40	0.34	1.00		
Tstat_Buttons_Done	0.12	0.31	0.25	-0.25	0.54	-0.00	0.25	0.25	1.00	
Tstat_Button.Size_Menu	-0.12	-0.36	-0.47	0.47	-0.20	0.10	-0.00	-0.50	-0.10	1.00
Participant_Age	-0.60	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.00	-0.02
Participant_Renter	0.22	-0.01	-0.06	-0.02	0.02	0.00	-0.00	0.06	-0.00	-0.02
App_Rated_Ease.of.Use	0.41	0.06	-0.05	-0.02	0.05	0.10	0.12	-0.10	0.01	0.12
Tstat_Rated_Ease.of.Use	0.46	0.17	0.04	-0.27	0.27	0.20	0.11	0.18	0.14	-0.05
Tstat_Rated_Feel.Sound	0.30	0.01	0.04	-0.09	0.02	0.00	0.07	0.05	-0.00	-0.07
Tstat_Rated_Appearance	0.09	0.1	0.22	0.00	0.04	0.00	0.18	0.06	-0.10	-0.01
App_Rated_Appearance	0.22	0.11	0.06	-0.05	0.05	0.10	0.18	-0.00	-0.10	0.04
Tstat_Steps_Mode	-0.14	0.27	0.14	-0.11	0.11	-0.00	-0.1	-0.20	-0.10	0.12
Participant_ComputerIQ	0.46	-0.01	0.00	0.00	-0.00	0.00	0.01	-0.10	-0.00	0.01
Participant_SmartphoneIQ	0.57	0.03	0.02	-0.06	0.09	0.00	0.03	0.11	0.06	-0.03

TABLE 24. CORRELATIONS FOR VARIABLES IN THE EFFICIENCY MODEL (PART 2)

	Participant_Age	Participant_Renter	App_Rated_Ease.of.Use	Tstat_Rated_Ease.of.Use	Tstat_Rated_Feel.Sound	Tstat_Rated_Appearance	App_Rated_Appearance	Tstat_Steps_Mode	Participant_ComputerIQ
Participant_Renter	-0.42	1.00							
App_Rated_Ease.of.Use	-0.37	0.14	1.00						
Tstat_Rated_Ease.of.Use	-0.38	0.16	0.44	1.00					
Tstat_Rated_Feel.Sound	-0.22	0.15	0.33	0.63	1.00				
Tstat_Rated_Appearance	-0.10	0.17	0.17	0.37	0.55	1.00			
App_Rated_Appearance	-0.28	0.17	0.58	0.26	0.36	0.41	1.00		
Tstat_Steps_Mode	0.02	-0.1	0.06	-0.00	0.01	0.10	0.00	1.00	
Participant_ComputerIQ	-0.43	0.08	0.23	0.23	0.15	0.03	0.09	0.00	1.00
Participant_SmartphoneIQ	-0.63	0.22	0.37	0.38	0.22	0.05	0.21	0.10	0.62

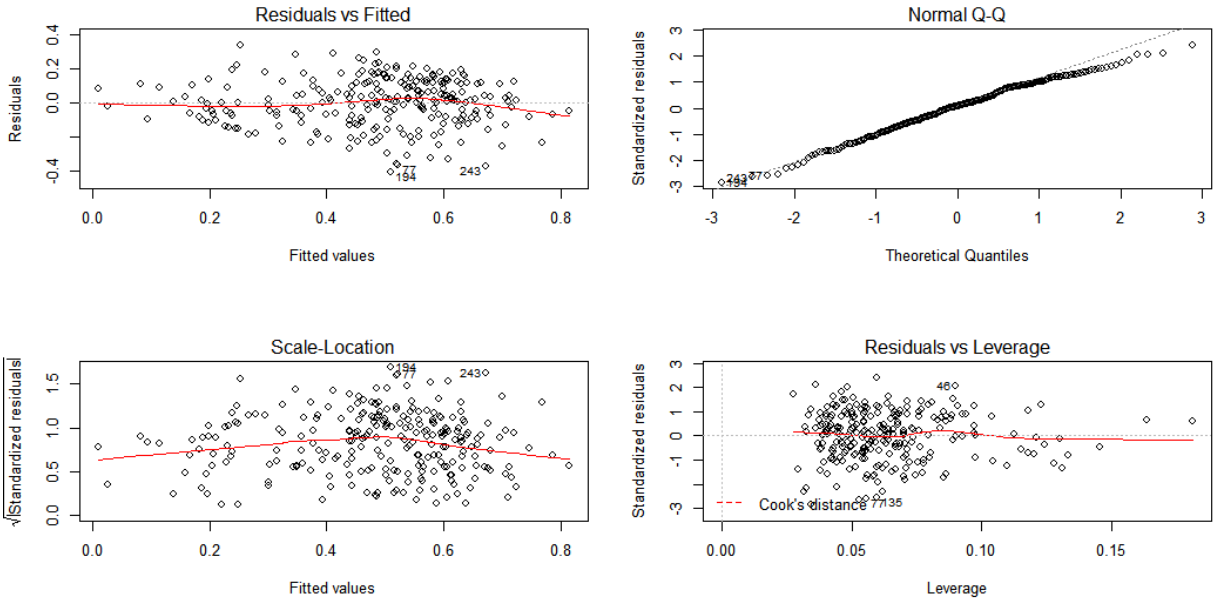
FIGURE 30. EFFICIENCY MODEL.1 – COEFFICIENTS AND SIGNIFICANCE

Efficiency	Estimate	Std. Error	t value	Pr(>|t|)	[1]
(Intercept)	0.1725	0.1009	1.71	0.0886	.
Tstat_Screen_Color	0.0343	0.0249	1.38	0.1696	
Tstat_Screen_Size	0.0038	0.0012	3.21	0.0015	**
Tstat_Buttons_Help	-0.0269	0.0251	-1.07	0.2845	
Tstat_Buttons_Done	-0.0288	0.0322	-0.89	0.3721	
Tstat_Symbols_Labels	0.0042	0.0296	-0.14	0.8877	
Tstat_Steps_Mode	-0.0267	0.0072	-3.72	0.0002	***
Participant_Age	-0.0042	0.0008	-5.07	0.0000	***
Participant_Renter	-0.0166	0.0214	-0.78	0.4385	
Participant_ComputerIQ	0.0441	0.0138	3.20	0.0016	**
Participant_SmartphoneIQ	0.0268	0.0120	2.24	0.0263	*
Tstat_Rated_Ease.of.Use	0.0123	0.0108	1.14	0.2548	
Tstat_Rated_Feel.and.Sound	0.0190	0.0115	1.65	0.1005	
Tstat_Rated_Appearance	-0.0132	0.0112	-1.18	0.2383	
App_Rated_Ease.of.Use	0.0352	0.0103	3.40	0.0008	***
App_Rated_Appearance	-0.0064	0.0113	-0.56	0.5743	

[1] Statistical Significance codes: ***: α=0.001; **: α=0.01; *: α=0.05 .: α=0.1

Residual standard error: 0.1445 on 236 degrees of freedom
Multiple R-squared: 0.54
Adjusted R-squared: 0.51
F-statistic: 18.69 on 15 and 236 DF
p-value: < 2.2e-16

FIGURE 31. EFFICIENCY MODEL – RESIDUAL PLOTS

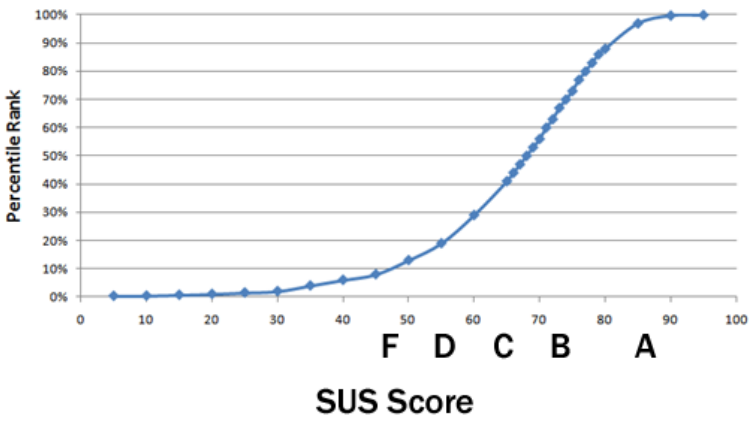


SUS MODEL

[To discuss with Vikki]

Also, for comparative purposes, even though the SUS scores correlate highly with both preference and efficiency scores, regress thermostat characteristics against first 8 sessions' SUS scores? I was thinking that we would want to regress the SUS scores that relate only to the thermostats and not the thermostat/app system against thermostat characteristics.

Add SUS grade (from http://www.measuringu.com/sus.php)



A raw SUS score of a 74 converts to a percentile rank of 70%. A SUS score of 74 has higher perceived usability than 74% of all products tested. It can be interpreted as a grade of a B-.

You'd need to score above an 80.3 to get an A (the top 10% of scores). This is also the point where users are more likely to be recommending the product to a friend. Scoring at the mean score of 68 gets you a C and anything below a 51 is an F (putting you in the bottom 15%).

A=>80
B=70?-80
C=60?-70?
D=50-60?
F=<50

These seem to me to track closest to his distribution above. I've checked several sources and the percentiles vary depending on the population size. As the population of SUS scores has increased the median SUS score has gotten lower.

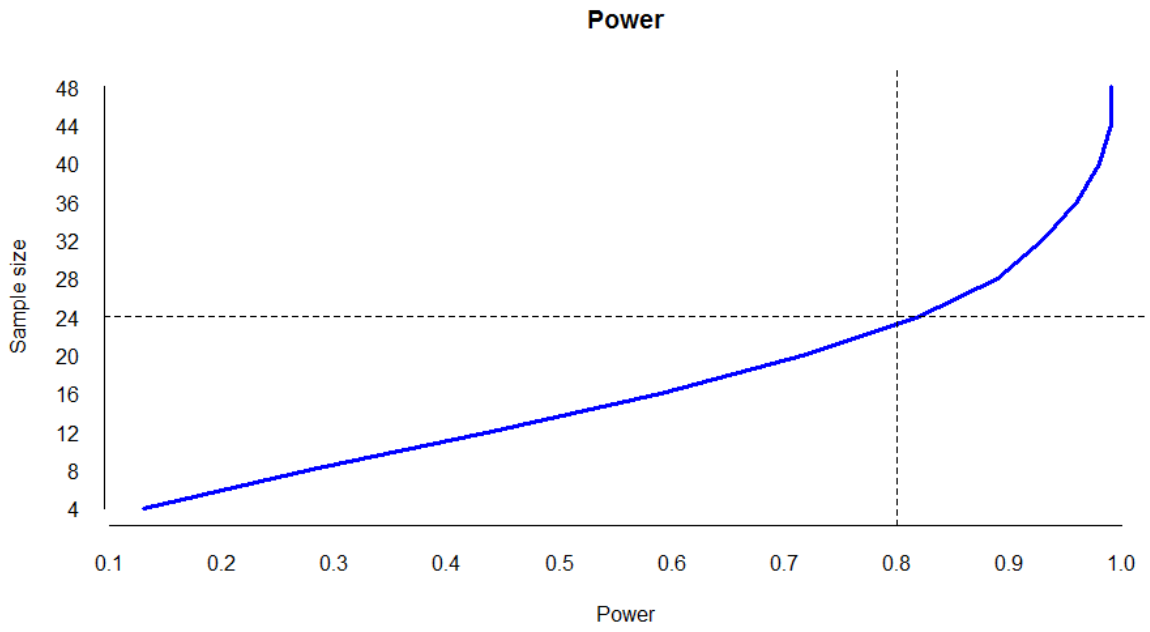
APPENDIX C – SAMPLE SIZE POWER ANALYSIS

Figure 32 plots the sample size requirements as a function of desired statistical power using the following parameters:

- Number of groups = 12 (one group per thermostat)
- Effect size = 0.25 (medium)[6]
- Significance level = 0.05 (95% confidence)

Figure 32 shows that a standard power of 0.8 requires that at least 24 participants test each thermostat. A sample size of 30 would provide a power greater than 0.9, which is generally considered excessive.

FIGURE 32. SAMPLE SIZE AND POWER FOR 12 GROUPS WITH 0.25 EFFECT SIZE, 95% CONFIDENCE LEVEL



[6] Effect sizes of 0.10, 0.25, 0.40 are standard for ANOVA power calculation.

SAMPLE SIZE CALCULATION SUS SCORE

Approximation of a sample size required to obtain 90% Confidence

M=1.65σn

n=2.72σ2/$M$2

σ=27.76 (obtained from SUS score, n = 307)

M = 5 which yields n=85

M = 10 which yields n = 22

95% Confidence

M=2σn

n=4σ2/$M$2

σ=27.76 (obtained from SUS score, n = 307)

M = 5 which yields n=124

M = 10 which yields n = 31

APPENDIX D – RECRUITMENT SCRIPT

Hello, May I speak to **(NAME ON LIST)**?

1) Gender **(RECORD BY OBSERVATION)**
   a) Male
   b) Female

2) Hello, my name is _____. I'm calling from Elliott Benson market research on behalf of SMUD to see if you would be interested in evaluating some new residential technologies. The testing will take about 90 minutes. For your participation, you would be compensated $70. Does this sound like something you might be interested in?
   a) Yes----------→ CONTINUE
   b) No----------→ **TERMINATE**

3) During the sessions, you will be given the opportunity to try out two items. After each one, there will be surveys and a short group discussion about your experience. These surveys and discussions will be held entirely in English. Do you feel comfortable speaking and reading in English?
   a) Yes----------→ CONTINUE
   b) No----------→ **TERMINATE**

4) On a scale from 1 to 5 with 5 being very comfortable, how comfortable do you feel using the following devices?

a) Computer	1	2	3	4	5
b) Smart phone	1	2	3	4	5
c) Thermostat	1	2	3	4	5

5) In what year were you born?
   a) Year _____
   b) Refused → **TERMINATE**

6) Have you or has anyone in your household ever worked for any of the following?
   a) Market research company ----------> **If YES, TERMINATE**
   b) News media----------> **If YES, TERMINATE**
   c) Public or private utility company ----------> **If YES, TERMINATE**
   d) The energy industry ----------> **If YES, TERMINATE**
   e) Marketing or advertising company----------> **If YES, TERMINATE**

7) Do you currently own or rent your home or apartment?
   a) Own
   b) Rent
   c) (Refused/Don't Know)

8) There are several sessions available on November 9th, 10th,11th and 12th. Will you be able participate on one of these days?
   a) Yes----------→ CONTINUE
   b) No----------→ **TERMINATE**

SCHEDULE DAY AND TIME

As a courtesy reminder, we'll call you on the day before the session. Is **[PHONE NUMBER]** still the best number to reach you?

We look forward to seeing you on **[DAY & DATE & TIME**].

THANK YOU!

APPENDIX E – DISCUSSION SESSIONS

The following sections were written by Group Works for the study.

FACILITATOR'S GUIDE

November 9, 10, 11, 12

(Seat participants in the chair that matches the first number on their nametag)

A. WELCOME and INTRODUCTIONS – 10 minutes

1) Introduce self and explain research purpose
   a. Thank participants for coming and participating in the research.
   b. I'm ES, an independent researcher who helps organizations tests products and look for ways to improve them.
   c. Purpose of today's study is to have consumers review thermostats both for ease of use and likeability. SMUD is interested in understanding what thermostat features are most important to consumers.
   d. Imagine you are shopping for a thermostat and have the opportunity to try out two models.

2) Describe room set-up and overall process
   a. Booths with twelve different thermostats; each of you have been randomly assigned to try out two of the thermostats.
   b. For each one, we'll ask you to complete a set of common tasks that are listed at each station in the Task Booklet. (Show consumers the task booklet)
   c. After you try out the first thermostat, we will have a brief discussion before you go to another booth to try a second thermostat and then we'll have another brief discussion about your experiences.
   d. To help us analyze the thermostat tests, we will be video-recording the process. You also may have noticed the one-way mirror behind me. Some of the people I'm working with may come to observe some of the testing, but they stay in the back room so it's less crowded and not distracting to the test or our discussion as they come and go.

3) Three key points to emphasize
   a. We value candid opinions. If you do not like the way your thermostat looks, feels, or works, it's important to let us know on your surveys and in the discussions. We didn't design these thermostats ourselves so you won't offend anybody by being completely honest.
   b. Each person has their own preferences about thermostats so differences of opinion are okay and expected.
   c. We are testing the thermostat, not the user! Do not worry about making mistakes or not being able to complete the tasks. If you have difficulty and are not able to complete a task easily, it is the fault of the thermostat design, not you. It's important for us to learn which thermostats are difficult or easy to use and why.

4) Instructions for the test

a. Before I read you some instructions about the test, is anyone left handed?

b. In a few minutes I'll have you turn around in your chair to your assigned station to try your first thermostat.

c. The tasks to perform on your thermostat will be listed in the Task Booklet. There is one task per page.

d. Each station has a thermostat and a smart phone app. Use the thermostat to perform the task unless the task tells you to use the app.

e. As you complete each task, write in the answer or simply mark the box labeled "Done" to indicate that you completed it. (show the checklist) For example, if you are asked to identify a scheduled temperature, write the degrees here in the box. If you are asked to change a setting, simply check the box to indicate when you are done and move on to the next task.

f. If you spend more than 2 or 3 minutes on a single task and don't feel you will be able to complete it, check the box "Not Done" and begin the next task.

g. We will be timing each task in order to see which thermostats are easiest to use The timer will start for each task when you turn the page and will end when you check either the "Done" or "Not Done" box so be sure to check the box one way or another. And if you need to do something other than the task, please do it after all the tasks are completed so it doesn't impact the timing (e.g. cell phone)

h. When you reach the end of the booklet and mark the last box, please fill out the survey.

i. Please stay at your station until I indicate that the 20 minutes is up. Then, I'll invite you back to the table for a brief group discussion before you try the second thermostat.

5) A few other logistics before we get started

a. Also in the room is Jenya, our technical support staff. She will be monitoring the video equipment and handling any technical problems that come up.

b. Jenya and I can only answer questions about the test process or if you think there is a technical malfunction. Since we are testing how easy or difficult it is to use the thermostats, we cannot answer questions about how your thermostat works or help you complete your tasks.

c. It's important that each of you do your best to complete the tasks and survey on your own – without any help from your neighbors. So please no talking at all to each other during the test. If you have questions about the process, please ask me.

d. Does anyone have any questions about the process now before we get started?

B. TEST #1 – 20 minutes

1) Please turn around in your chair to face your assigned thermostat booth.
2) Before we start the tasks in the booklet we'll do a quick test to make sure you are comfortable with the thermostat. This test task is to use the thermostat to increase the temperature by one degree.
3) You will have 30 seconds. At the end of that time, we will stop for a moment to make sure everyone was able to complete the first task, and provide instruction for those of you who had trouble.
4) Reminders:
   a. When I say "begin," please open your booklet and start the first task. After each task, write in the answer or simply mark the box "Done" and move onto the next task until you've completed all of them.
   b. Remember, if you spend more than 2 or 3 minutes on a single task and don't feel you will be able to complete it, it's okay to mark the box "Not Done" and move on to the next task.
   c. When you finish all the tasks, please complete both sides of the survey and stay quietly at your station until I call everyone back to the table.
   d. Please do not remove anything from the wall, and be careful to avoid moving the camera.
   e. Are there any questions?
5) **(Check that technical staff is ready to begin video recording)**
6) Is everyone ready? You may now begin. START TIMER

C. DISCUSSION #1 – 15 minutes

1) Time is up … please remove your checklist from the station, attach it to the back of your survey, and bring your survey back to the table with you.
2) Before we start the discussion and move on to the second thermostat test does anyone have any questions about the process?

As group discusses feedback, create chart of key strengths and weaknesses of thermostats.

3) By show of hands, how many would exchange your unit at home for the unit you just tested if it was free?
   a. If yes, what did you especially like about it? What made it easy to use? What features did you like?
   b. If no, what didn't you like? What was especially difficult or frustrating to use?

(Note: do NOT probe as to which thermostat they used to minimize any potential bias)

D. TEST #2 – 20 Minutes

1) Please proceed to the chair that matches the second number on your nametag and turn around to face your second assigned thermostat booth.
2) Once again we'll do a quick test to make sure you are comfortable with the thermostat. This test task is to use the thermostat to increase the temperature by one degree.
3) After 30 seconds I'll ask you to stop and Jenya will provide instruction for those of you who had trouble.
4) Don't forget – no talking and stay at your booth until I invite the group back for our final discussion.
5) **(Check that technical staff is ready to begin video recording.)**
6) Is everyone ready? You may now begin. START TIMER

E. DISCUSSION #2 – 20 Minutes

1) Time is up. Please remove your checklist from the wall, attach it to the back of your survey, and bring your survey with you to the table. I'll pass around a stapler so you can staple the two surveys together. Please put the first thermostat survey on top.
2) Just a show of hands, how many would exchange your unit at home for the unit you just tested?
2) Let's go around the table and hear from each person which of the two thermostats you tried you liked best and what features made you prefer that one over the other one.
2) If you were going to purchase a new thermostat for your home, which features would be most important to you when making your selection?
3) Look at the list of possible advanced features on your survey:
    a. Which ones are important enough to you to make you want to seek out a thermostat with that feature? What makes that important to you?
    b. Which features would you not want to have on your thermostat? What makes you feel that way?
4) Are there any other aspects of the thermostat design that we haven't talked about yet that are important to you when choosing a thermostat for your home? Or specific improvements you'd like to see?

F. WRAP-UP – 5 minutes

1) Any final advice you would give to thermostat designers to develop a thermostat that is ideal for you?
2) Thank respondents and instruct them to see hostess for their incentive.

FOCUS GROUP DISCUSSION SUMMARY

Focus group participants discussed which thermostat features were most important to them and what advice they would give to companies designing thermostats. Their feedback can be summarized in twelve key themes.

1. Keep it Simple
    - They want the basic functions of heating/cooling and adjusting temperature up and down easily available at the top level.
        - Heating and cooling should not be on separate screens/menus/levels. It should not be difficult to change from heat to cold.
        - Many are used to and prefer up and down arrows for increasing and decreasing heat.
        - Some like a dial that allows them to change multiple degrees more easily than pressing a button multiple times.
        - Many like the heat and cooling options to be color coded with red for heat and blue for cool.
    - Many found the thermostats they used to be too complicated.
        - Many felt there were too many options.
        - Some felt the thermostats they tested were so complicated and difficult to use they would not exchange the new unit for their existing thermostat at home, even if it was free.
        - One said he shouldn't need an advanced degree to operate it. Another said he'd need a class to learn how to use it. Many wished for an instruction manual.
2. Make it Intuitive by Using Commonly Understood and Legible Icons and Labels
    - They want self-explanatory icons. Several testers complained that on some thermostats they were unable to discern the labels and icons.

3. Readability: Well Lit and Large Enough Screen/Font
    - Well lit ... both for better legibility and for seeing in dark hallways at night
        - Some expressed frustration when the light didn't stay on or the screen looked too dark or dim.
    - Large enough screen and large enough font for readability (especially important for seniors and others with limited vision)
        - Several commented that #5's screen was too small.
        - Several complained about too small print on some of the thermostats.
4. Apps
    - Not all consumers, especially among the older age demographic, use smart phones regularly and are familiar with apps. Several said they would prefer to use only the thermostat to perform essential functions. Several were mystified and frustrated trying to use the smart phone app to adjust the thermostat settings. One explained simply, "I don't do computers."
    - Some Android users had trouble with the iPhone app. They suggested the need for both Android and iPhone apps
5. Design Thermostats and Apps to Match

- It's easier for consumers to learn if the thermostat and app look the same (similar graphics, icons, language, organization). Conversely, it adds to their confusion and frustration when they are different.

6. Appearance
   - Many prefer an updated, streamlined look for their thermostat.
   - Several thought the #11 was too bulky (one described as a "mini boom box").
   - Many like the screens with good color and picture options. Several commented positively on #6's photo.
   - A few think the photos in the background are distracting and prefer a plain screen. Ideally, it would be an option.
7. Responsive Touch
   - Many said they want some feedback when pushing on buttons.
   - Several complained that buttons were hard to tap.
   - A few complained that it was too slow to reflect input on some thermostats and it felt like a delayed reaction.
   - Test users could be seen repeatedly jabbing at the screen trying to get it to respond, especially thermostat #11. This looked particularly difficult and frustrating for some older users with shaky hands.
8. Confirmation
   - Consumers want some form of visual or audible confirmation that their setting adjustments have been received. Several expressed frustration when they are uncertain if they were successful in completing the task.
9. Sound
   - The beeping noise of #12 bothered some users. One said she wouldn't want it to wake up sleeping family members if she made adjustments at night. Ideally, sound would be optional.
10. Make it Easy to Get Back Home
    - Several users complained they got lost and couldn't find a way to get back home on a few thermostats and thought they should always have a clear path home.
11. Instructions/Help
    - Many wished they had a manual or instructions. Several suggested a help feature be available.
12. Familiarity
    - Many simply prefer the style of thermostat they are used to.
    - Some complained about the style of #1 Nest as being SO different, it's hard to adjust to turning and pressing. Some said they felt they would get used to it eventually. Many tried swiping and pressing different parts of the screen, but not the center. Several did not find the dial. There were several instances of visible user frustration.

APPENDIX F – DATABASES (SEPARATE FILE)

PARTICIPANT DATABASE

The Participant Database is a matrix containing one record for each participant, or roughly 180 records. Each record contains information about the participant, IDs for the thermostats they tested, and any other data needed for testing.

SURVEY DATABASE

The Survey Database is a matrix containing one record for each thermostat test survey, equal to the number of participants times two, or 310 records. Each record contains the participant ID and all information collected by survey.

THERMOSTAT DATABASE

The Thermostat Database is a matrix with one record for each thermostat, for a total of 12 records. Each record contains information about the thermostat features, average survey-based ratings for ease of use, looks, and sound/feel; efficiency and preference scores; and any other measures used in the final regression analysis.