A mixed methods approach for investigating residential building thermostat interfaces in the context of use, usability, and user understanding

by

Ruth Tamas

A thesis submitted to the Faculty of Graduation and Postdoctoral Affairs in partial fulfilment of the requirements for the degree of

Master of Applied Science

in

Civil Engineering

Carleton University

Ottawa, Canada

Abstract

Residential energy consumption is largely influenced by thermostat configuration and user interaction. Smart thermostats differ from other devices, offering new features and control. Little comparison or analysis of state-of-the-art systems is available, contributing to a gap in the literature. This thesis examines usability and user understanding related to smart thermostats, emphasizing comparison with literature. Two interview-based studies employing human-computer interaction methods are presented. Chapter two investigates usability, reporting quantitative metrics and user feedback. Smart thermostats demonstrated higher usability compared to programmable devices, while enabling the same or more functionality. Chapter three investigates users' understanding related to smart thermostat operation. Users demonstrated a relatively accurate conceptualization of their system and most employed temperature setbacks, suggesting that smart thermostats effectively communicate the function of the device and enable energy saving behaviour. Concrete objectives for future work are presented. This research brings the smart thermostat literature up-to-date, relative to manual and programmable thermostats.

Preface

The following thesis consists of two journal articles, one published and one under review. However, readers who wish to refer to materials from this document should cite this thesis. The following articles are contained within this thesis:

Article 1: Tamas, Ruth, William O'Brien, and Mario Santana Quintero. 2021. "Residential Thermostat Usability: Comparing Manual, Programmable, and Smart Devices." *Building and Environment.* 203: 108104.

Article 2: Tamas, Ruth, William O'Brien, and Mario Santana Quintero. 2022. "Evolving interaction: A qualitative investigation of user mental models for smart thermostat users." *Architectural Science Review.* [Under review]

To deliver a cohesive thesis, minor adjustments have been made to the articles. Use of copyrighted material from the published articles is acknowledged as per the corresponding publisher's permission guidelines with respect to the author's rights.

Ruth Tamas is the lead researcher and primary contributor to the research methodology, data collection and analysis, tables and figures, and preparation of written materials. The work was completed under the supervision of Dr. William O'Brien and Dr. Mario Santana Quintero.

Acknowledgements

This work was supported by NSERC CREATE Heritage Engineering and Mitacs. I'd like to extend gratitude to my supervisors, Liam O'Brien and Mario Santana Quintero, for their time and support, and to Professor Philip Agee for his advice. Further gratitude is due to the participants, without whom this research would not have been possible. Finally, a special thanks to my family and friends: Dora, Michaela, Lorena, and Danielle for their support during this degree.

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter 1: Introduction

Thermostat adjustment represents the primary method for residential occupants to interact with and influence the energy demand of their heating, ventilation, and air-conditioning (HVAC) system. In Canada, heating and cooling accounts for roughly 65% of the total energy use in households [1]. In America, thermostats control approximately 9% of the country's total energy demand [2]. Research with programmable thermostats has demonstrated that energy savings are primarily determined by the occupant(s), and consumption is highly variable and dependant on the inputted thermostat settings [3]–[5]. Similarly, the probability of achieving the projected savings in high performance buildings was linked to the occupants' understanding of and ability to operate the system, as well as their comfort preferences [6]. Unlike other user interfaces, thermostats do not have requirements for interface configuration, and no standardized process for usability testing [7]. Thermostat users' interactions (or lack thereof) have a significant impact on consumption. This has been emphasized in government initiatives, such as the Home Winterproofing Program in Ontario, Canada, where eligible households receive a free smart thermostat, with the hopes of increasing the use of temperature setbacks through automatic adjustments [8], [9].

As one of the easiest measures to implement, thermostats present a unique opportunity for improving building performance. As infrastructure ages and increases in cultural value, methods to improve building performance can become increasingly limited and complex. Ottawa, Ontario recognizes the conservation of cultural heritage as a priority, with over 300 individually designated heritage buildings and 20 conservation districts [10], where alterations (e.g., window replacement) require approval from the City of Ottawa [11]. Thermostat replacement and upgrade is one of the least invasive and expensive measures for improving building performance and energy use, making it especially suitable for existing structures that require low intervention to preserve heritage elements.

Thermostats present the opportunity to communicate with occupants and shape behaviour; however, the variety of thermostat capabilities and configurations imposes varying constraints. Thermostats can be divided into three categories with differing control opportunities: manual, programmable, and smart. Manual thermostats represent the most simplified and traditional thermostats. The setpoint is manually adjusted by the user and changes can only be realized through direct interface manipulation. Manual devices are simple to use but require consistent interaction from occupants to achieve setpoint changes and temperature setbacks, limiting their reliability in yielding energy savings. Programmable thermostats attempted to reduce the reliance on occupants to achieve temperature setbacks, allowing users to program a schedule of automated changes. However, the literature has documented extensive instances of poor interface usability [3], [5], [12]–[19] that discouraged users from engaging with programmable features, at times resulting in increased energy consumption [20]–[23]. Smart thermostats extend the features of programmable thermostats, offering new interface layouts, methods of control, interaction opportunities, and increased feedback. They present more complex systems, integrating with other smart home devices, enabling new avenues for interaction (e.g., verbal interaction, remote interaction). A recent comprehensive review of building interfaces [24] emphasized the need for research related to the impacts of interface configuration. Smart thermostats have received positive [25], [26] and negative [27] reviews related to usability, but somewhat limited in comparability with manual and programmable thermostats, as concrete usability metrics are yet to be established.

While important, usability alone does not determine energy consumption associated with thermostats. How users expect and understand their HVAC system to function has been shown to drive occupant behaviour and usage, with implications for energy [2], [28]–[31]. User understanding and conceptualization has been investigated in relation to manual [29] and programmable thermostats [2], [28], yielding insights for usage behaviour. Researchers employed user mental model (UMM) diagramming exercises to produce visual representations of users' conceptualizations for their HVAC system and identified different approaches to thermostat use. For example, Kempton [29] identified two approaches to manual thermostat operation from UMM investigation. Referred to as feedback and valve theory, each UMM had specific implications relating to the likelihood of temperature setbacks and energy savings. UMMs for smart thermostat users have yet to be introduce to the literature, limiting the contextualization of smart thermostats as energy saving devices.

There appears to be consensus among researchers that programmable thermostats have not met their potential as energy-saving devices [3], [12]–[17], [32]–[34]. At the same time, manual thermostats' reliance on constant user interaction limits the reliability of energy savings practices. Given the lack of literature related to smart thermostat interfaces, in terms of usability testing and UMM investigation, the same definitive arguments cannot be made. Therefore, to better understand the impacts of the adoption of smart thermostats, research is needed to further investigate them as energy savings devices and contextualize them against manual and programmable counterparts.

This thesis seeks to introduce research that will contribute to the current body of literature such that smart thermostats can be effectively compared to manual and programmable thermostats in terms of usability and user understanding, informing the likelihood of energy savings through occupant usage. A series of quantitative and qualitative investigations are employed to address the existing gaps in the literature and deliver comparable results. To promote research consistency, the methodologies employed to evaluate usability and UMMs follow that of previous studies featuring programmable and manual thermostats.

To provide a comprehensive assessment, methods from human-computer interaction and design are employed. First, smart thermostats are evaluated in terms of usability, alongside manual and programmable devices. Chapter two aims to establish usability metrics for smart thermostats, like those found in literature for programmable and manual devices. The usability testing is contextualized with qualitative and self-reported feedback from thermostat users, related to their usage experience. Once smart thermostat usability is established, relative to programmable and manual counterparts, a qualitative analysis is performed. Chapter three focuses on a sample of smart thermostat users and provides a deeper investigation of what smart thermostat interfaces communicate to users and how users conceptualize their thermostat relative to heating system components. UMMs are evaluated for the presence of shared theories previously identified in programmable and manual thermostat users. Related energy outcomes are discussed. Both chapters seek to capture the characteristics and elements of smart thermostat usage and shed light on the experience of users.

1.1. Motivation & research questions

Considering that the current understanding of usability, user experience and human-computer interaction with respect to smart thermostat interfaces is somewhat limited, this thesis presents an investigation with a sample of smart thermostat users to address these gaps. Results are contextualized through comparison with established metrics and theories from previous research with manual and programmable devices. Two studies, focusing on quantitative and qualitative outputs, form the subsequent chapters. Accordingly, each chapter attempts to address the following research questions:

- **Chapter 2: A quantitative usability evaluation for manual, programmable, and smart thermostat interfaces**
  - How does the usability of manual, programmable, and smart thermostat interfaces compare?
  - What types of HVAC and temperature feedback are provided in homes?
  - What thermostat interface features do users desire/dislike, that should be considered for future design?
- **Chapter 3: A qualitative evaluation of smart thermostat users**
  - Which user mental models/shared theories exist in the group of smart thermostat users?
  - What attitudes are held towards smart thermostat devices and how does their adoption relate to energy consciousness/consumption?
  - How effective is the use of online software and remote strategies in the investigation of user mental models?

1.2. **Document structure**

This integrated thesis presents two chapters, each of which describe a research study performed to evaluate smart thermostats – either in the context of usability metrics and user feedback or user conceptualization and understanding. Both chapters are informed by previous literature and contextualize results accordingly. The following summarizes the contents of each chapter:

- **Chapter 2:** This chapter presents an interview-based study that examines the relationship between thermostat usability and interface characteristics in Ottawa, Canada, with a focus on devices currently installed in homes. To compare manual, programmable, and smart interfaces, human-computer interaction methods, including participant interviews and think-aloud analysis, are employed. A sample of 51 participants were interviewed in their homes and attempted six usability tasks on their thermostat interface(s). Interviews were recorded and each thermostat type (manual, programmable, smart) was evaluated based on the time to complete, required interactions, and qualitative observations. An averaged time and success metric was assigned to each category and contrasted with previous literature. In general, smart thermostats were found to be significantly more usable than their programmable counterparts. Participants expressed a desire for more feedback regarding energy consumption and half (51%) noted their either wanted or enjoyed thermostat control through a smartphone application.
- **Chapter 3:** This chapter investigates how smart thermostat users imagine their home heating system to work and provides insight into what is communicated by the smart thermostat interface. It employs novel research methods to remotely interview participants and construct user mental models (representations of users' internal conceptualizations) using video-calls

and diagramming software. Ten user mental model diagrams are produced, and three detailed case studies are presented. Results are contextualized with previous work through the comparison of cognitive structures and usability issues. All participants produced a user mental model diagram that follows an established cognitive theory (feedback theory), which correlates to an effective understanding of the heating system, enabling effective usage. Three case studies highlight misconceptions and misunderstandings among users. Many participants preferred the design of the smart thermostat; however, a steep learning curve was required to effectively use the features. Overall, smart thermostat users appear to understand how to use their heating system effectively, contrasting previous research with programmable thermostat users. However, users are frequently overwhelmed by the complexity of smart thermostats, limiting engagement with energy-saving features.

The final chapter (4. Conclusions) summarizes the two preceding chapters and restates their findings, contributions, and recommendations for future work.

Chapter 2

This chapter has been published as:

Tamas, Ruth, William O'Brien, and Mario Santana Quintero. 2021. "Residential Thermostat Usability: Comparing Manual, Programmable, and Smart Devices." *Building and Environment.* 203: 108104.

Chapter 2: A quantitative usability evaluation for manual, programmable, and smart thermostat interfaces

2.1. Introduction

Thermostats have been a well-established component of the home since the late 1920s [35]. They were originally very simple electromechanical devices with a temperature setting and switch. Energy savings were heavily dependent on occupants, requiring diligent manual adjustments to achieve energy-saving strategies, such as nighttime setbacks.

Previous research has shown that residential thermostats control approximately 9% of the total energy consumption in the United States of America [2]. Because temperature control is dictated by user interaction (e.g., selecting a setpoint), the usability of thermostat interfaces is critical. The International Organization for Standardization (ISO 9241-11:2018) defines usability as the "extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use." The literature has emphasized the need for careful and consistent user interface design [36]; however, many of us have soon realized that the various thermostats we encounter – in our homes, offices, hotels, and friends' homes – vary greatly in appearance, functionality, accessibility and ultimately, usability.

Starting in 1995, Energy Star endorsed programmable thermostats as means to save energy, suggesting that users could save almost $200 a year [37]–[39]. Programmable devices allow the user to schedule setpoint adjustments, enabling setbacks and related energy savings without continuous interaction. The development was promising, as thermostat replacement represents one of the least invasive and inexpensive measures for improving building performance. Various building codes and government programs endorsed programmable thermostats and required their installation because of the assumed energy savings. However, later studies [3], [5], [18], [19] showed that many programmable thermostats were overly complex, offering low usability and no significant savings, and sometimes higher consumption, compared to manual thermostats. Many researchers have documented instances of poor usability in programmable thermostats [12]–[17], noting that interfaces were too intricate for effective use. This evidence prompted Energy Star to terminate its endorsement in 2009 [37].

Ecobee introduced the first smart Wi-Fi compatible thermostat in 2007 [40]. As these new thermostat interfaces become increasingly mainstream, the question of usability resurfaces. Previous work [24] has suggested that the widening gap between building systems and occupants can reduce perceived control, often at the detriment of occupant satisfaction, productivity, and comfort. To realize the potential energy savings, it is paramount that users understand how to operate their thermostat interface(s) to prevent unintended discomfort and empower energy saving behaviour. But has this occurred for smart thermostats?

There is a growing variety of thermostat interfaces with diverse control layouts, features, and degrees of automation. While elaborate devices engage some users, they alienate others. Unlike other interfaces, such as a car dashboard, thermostats are not subject to detailed design standards [7].

At best, feature-rich interfaces, such as the Google Nest, estimate an average of 10%-15% energy savings [41]. At worst, "advanced" devices overwhelm users and reduce perceived control. Occupants who struggle to understand their building interfaces often experience frustration, discomfort and in many cases, increased energy use [4], [42]–[44].

Occupants rarely receive training or instructions for thermostat operation, making walk-up usability (i.e., the ability for user to rapidly learn and understand how to use the interface) critical for promoting energy savings. However, it is clear from previous work that many interfaces, especially programmable, do not meet this requirement. While manual thermostats offer a high degree of simplicity, they require diligent adjustment from occupants to achieve energy-saving schedules. Smart thermostats have the potential to address the downfalls of programmable and manual devices; however, their success will ultimately be dictated by the interface design. The issue at hand is driven by a lack of design focus on interface usability and the increasing integration of greater levels of automation.

Given the nature of Canadian homes, and North American homes in general, which tend to have central forced-air HVAC systems, the results of this study may not be applicable to other regions or systems. This work investigates the usability of manual, programmable and smart thermostats in Canadian homes and applies the following research questions:

RQ1: How does the usability of manual, programmable and smart thermostat interfaces compare?

This question is addressed through quantitative comparison of six tasks performed on manual, programmable and smart thermostats. Usability is determined through various metrics, including the task duration, interactions taken, instances of confusion and user-assigned usability ratings. A time and success metric was applied to allow for comparisons with previous literature.

RQ2: What types of HVAC and temperature feedback are provided in homes?

Participants were asked, "when you adjust your thermostat, how do you know the HVAC system is working?" Respondents noted the various types of tangible, auditory and visual feedback provided by thermostats and/or HVAC systems.

RQ3: What thermostat interface features do users desire/dislike, that should be considered for future design?

Participants were asked to list features that they would include in their "dream thermostat". Responses were not necessarily feasible and reflect the desires of users based on their current thermostat device.

2.1.1. Defining manual, programmable, and smart thermostats

For this study, the definition of manual thermostats includes all non-programmable thermostat interfaces. Manual thermostats offer basic controls, often limited to a few buttons or a dial, and must be readjusted to achieve a new setpoint. Figure 2.1 shows a sample of manual thermostat interfaces included in the study.

Figure 2.1: Examples of manual thermostat interfaces

Programmable thermostats are more complex, enabling users to set and program a schedule for at least two time periods. These devices can also offer additional settings, such as vacation mode (where users can temporarily maintain a setpoint for a given period). The interfaces vary widely by product, offering combinations of touchscreen, button and/or dial interaction. Programmable thermostats are limited to direct interface programming and cannot be adjusted remotely. Figure 2.2 gives examples of observed programmable thermostats.

Figure 2.2: Examples of programmable thermostat interfaces

Smart thermostats have similar scheduling protocols as programmable thermostats, with additional opportunities for control and automation. Connection to the internet, occupancy detection, prediction of time to setpoint, geofencing, integration with other smart home hardware/software and remote adjustment (e.g., through an app) set smart devices apart from programmable interfaces. Some enable real-time tracking of energy expenditures while others simply allow users to program their schedule from their mobile device. Examples of smart thermostat interfaces are shown in Figure 2.3.

Figure 2.3: Examples of smart thermostat interfaces

2.2. **Literature review**

The user plays a large role in determining HVAC energy use, influenced by how they configure and interact with their thermostat. To achieve savings, thermostat setpoints must be set up and set back, either manually or through a programmed schedule. Due to the dependency on occupant behaviour, various studies show conflicting results regarding the impact of programmable thermostats. While some found increased energy savings [20]–[23], others show no significant difference between homes with manual or programmable systems [3], [5], [19]. It is clear that programmable thermostats have the capacity to save energy, but often fall short. So why aren't more people programming their thermostats "properly"?

2.2.1. **Usability issues**

Programmable thermostats are fraught with usability issues. Freudenthal and Mook [15] noted that occupants with programmable thermostats rarely utilize all of their device's functions, even valuable ones, due to poor interface design. Other studies [2], [45] have recognized that the layout of thermostat interfaces is often confusing or illogical. These issues prevent effective use and discourage user engagement. Furthermore, when occupants face barriers to their personal comfort, they create work-around solutions, often at a detriment to energy efficiency. For example, research has shown that occupants frequently override their programmable device to use it as a manual thermostat or on/off switch [37].

2.2.2. **Placement**

Beyond the interface design, the placement of the thermostat(s) within the home directly relates to the usability of the device. Meier et al. [37] found that the majority of household thermostats were obstructed or improperly positioned, reducing interface accessibility and functionality. Anecdotally, in some cases users feed thermostat wires through ducts, which ultimately results in supply air being blown directly at the back of thermostats (B. Huchuk, personal communication, March 3, 2021). Thus, obstructions preventing physical access to the interface, and external influences, such as direct sunlight, make thermostat placement an important component of overall usability.

2.2.3. **Smart thermostats**

Smart thermostats have been proposed as a solution to the downfalls of their programmable counterparts – enabling user engagement through familiar control interfaces (e.g., an iPhone application), minimizing the number of user inputs and providing more useful display for improved comprehension. Web-based interfaces have shown positive results, including more successful instances of schedule completion and user understanding [25], [26], [36].

On the other hand, smart thermostats may widen the conceptual gap between the interface and the HVAC system. Studies have shown that occupants do not understand how their HVAC system works [16], [32], [46] and that more informative control layouts (based on HVAC system components) lead to more effective control decisions [31]. Karjalainen [46] found that thermostat and HVAC usability can influence perceived control, which can affect thermal sensation of occupants. A recent study [27] comparing twelve smart thermostats reported a wide range of usability among devices. Over one-third of the studied sample did not achieve the minimum acceptable degree of usability. To avoid repeating the same mistakes seen in programmable thermostats, smart devices must prioritize information dissemination and effective control layout without overwhelming the user.

2.2.4. **User comprehension**

A recent study concluded that the probability of achieving projected savings in high performance building designs was dependent on the occupants' knowledge of and capacity to operate the system, as well as comfort expectations [6]. As programmable and smart thermostats continue to include more advanced features, research suggests that user comprehension may suffer, especially among elderly users [47]. Many studies have indicated that programmable thermostats are already too complicated for use [3], [12]–[17], [32]–[34]. To support users, thermostat interfaces must be accessible, with special consideration for elderly demographics and people with disabilities or impairments.

2.2.5. **Usability testing**

Thermostat usability has been determined through multiple factors, including, but not limited to, thermostat placement, interface layout and the ability for users to complete important tasks, such as programming a schedule. Unlike packaged consumer products, thermostats are normally positioned within the home relative to the HVAC system by occupants or other installers, which adds diversity of circumstances. While other fields have refined their usability assessments, for example, the testing and development of in-vehicle information systems [48], [49], thermostat interface testing is fairly primitive in the literature. Furthermore, occupants are not necessarily required to engage with their thermostat interface, adding complexity to the interpretation of remote testing; it is difficult to determine whether a user does not know how to operate their thermostat or whether they simply choose not to interact. Valuable insights can be derived through remote research [50], [51]; however, these results are often limited to the study of actual user inputs, opposed to user experience. The lack of contextual information makes it difficult to distinguish user actions, including intent, missteps, confusion, frustration, etc.

Previous research has employed time and button metrics to assess the usability of thermostat interfaces. Perry et al. [36] had participants perform five usability tasks on a sample of five thermostat interfaces. Similarly, Meier et al. [37] had participants complete six tasks on five thermostat interfaces. Both studies evaluated usability using a time and success (TS) metric.

The TS metric has been applied to programmable thermostats to indicate usability. It quantifies consumer understanding by rating devices on a scale of 0 to 1 based on task completion and duration, providing a bounded scale for usability.

2.2.6. Gaps in the literature

Manual, programmable, and smart thermostat interfaces face varying constraints that promote different degrees of usability, contributing to (or preventing) effective use. Manual thermostats offer relatively simple and intuitive interfaces but are limited by their reliance on consistent adjustment to achieve setbacks. Programmable and smart thermostats allow users to automate setbacks but are often complex, overwhelming users and discouraging engagement, which ultimately negates potential energy savings.

A recent extensive review of building interfaces [24] noted the need for further research focused how occupants interact with building interfaces, how interfaces provide feedback to occupants, and the impact of interface configuration on occupants' satisfaction and perceived control. While thermostat usability has been investigated previously, the literature is limited by a lack of in-situ testing, few thermostat interfaces tested, a lack of interdisciplinary methods/approaches and small participant groups (e.g., ranging from 14 to 31 participants [7], [36], [37] [42]). Various studies have tackled these components individually; however, this research attempts to address all limiting factors, presenting novel information through the investigation of residential thermostat usability for a sample of 51 participants, including manual, programmable, and smart interfaces.

2.3. **Methodology**

2.3.1. **Participant recruitment**

A total of 51 participants (29 female, 22 male) ranging from 20 to 80 years old were recruited for the study (see Figure 2.4 for age distribution). Approximately half of the participants (47%) were found after the study was advertised via social media. Respondents were sourced from a wide variety of technical backgrounds, representing a broad demographic of users. The remaining participants (53%) were recruited through word of mouth, often referred by previous participants. The utilization of social media resulted in most respondents falling within the 20-30 age range. After completing the interview, each participant was awarded $10 CAD, though only three-fifths accepted this incentive.

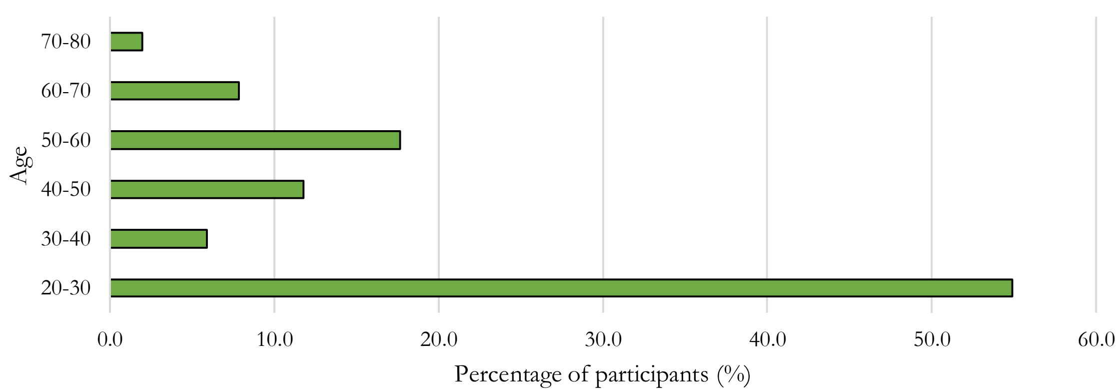


Figure 2.4: Participant age distribution

To be an eligible participant, individuals had to meet the following criteria:

- Must live in Canadian household
- Must have access to the home's heating system (e.g., gas furnace)
- Must have access to thermostat interface(s)

- Must have reliable access to internet connection
- Must have a mobile phone and laptop with a camera for online interview
- Must be able to receive calls from a video-call platform
- Must be an adult (over 18 years old)
- Must be fluent in English

The recruitment and interview process took approximately two months during the heating season, beginning in late-November 2020. Most participants live in the Toronto-Ottawa area. The sample of thermostats includes 11 manual, 33 programmable and 9 smart interfaces. Table A.1 (see Appendix A) lists the brands/models included in the study. While this study does not feature large participant samples for each thermostat category, literature notes that usability testing often identifies 80% of the problems with the first four to five subjects [52].

In some cases, occupants had multiple thermostats in their home (e.g., a participant has a manual and programmable thermostat). This resulted in disparities between the number of participants (51) and number of interfaces evaluated (53).

2.3.2. Interview on thermostat interface usability

A complete list of all survey questions is available in Appendix B. Interviews were conducted by Zoom video call. The questions asked are divided into three parts. A high-level visual description of the methodology is shown in Figure 2.5.

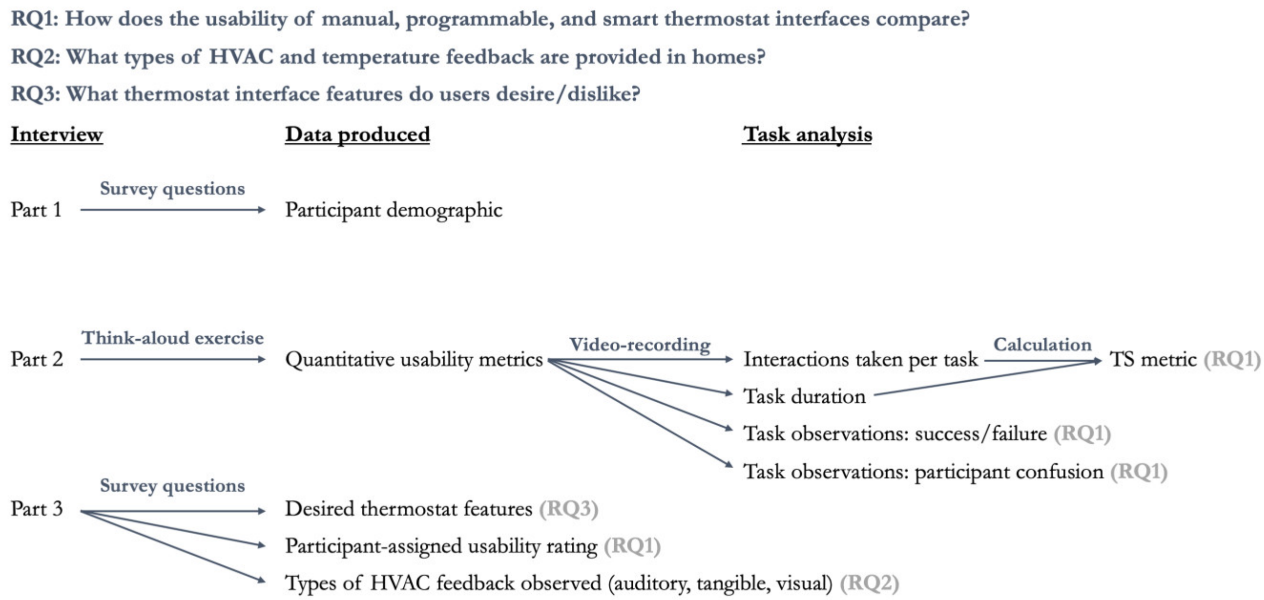


Figure 2.5: Methodology overview

Part 1. Questions to establish participant and household characteristics

To start the interview, participant consent and general information was collected. Once basic information was established, participants showed the researcher their heating/cooling system and thermostat, and then proceeded with usability tasks.

Part 2. Usability tasks

The usability tasks were adapted from previous research [36], [37] to promote consistency with the literature. The six tasks selected were considered the most important functions to operate the thermostat effectively and properly.

Think-aloud protocol, a common design research method [53], was employed to contextualize and enrich usability evaluations. This methodology highlights aspects of interface design that delight, confuse, and frustrate users and contextualizes the quantitative data. Participants were required to verbalize what they were doing and thinking as they completed each task. The researcher explained the purpose and goals of the think-aloud protocol to participants before beginning tasks and reminded them to speak aloud as required.

While think aloud analysis provides valuable insight into user experience, it comes with limitations. The method is likely to increase the time and cognitive load required for each task. Literature notes that thinking aloud while using a device can interfere with reasoning [54], especially in elderly subjects and are susceptible to short-term memory overload [55]. Additionally, it can alter behaviour and may be confusing or difficult for the participant [15], [56]. These limitations should be considered when reviewing quantitative results.

Due to the varying range of features available for manual, programmable and smart thermostats, the required tasks were dependent on the type of thermostat. At minimum, users were expected to be able to complete two of the six tasks for the most limited interfaces (manual thermostats). When a task could not be completed due to the thermostat interface constraints, it was marked as not applicable (N/A).

Some participants had multiple thermostats in their homes. If the thermostats differed, users had the option to complete the usability tasks on both interfaces separately. Table 2.1 shows the number of thermostat interfaces with the capacity to complete each task.

Table 2.1: Number of thermostat interfaces that allow each task

Task	Thermostat sample capable of completing the task (percent of sample)			
	Manual	Programmable	Smart	Total
1: Turn the thermostat off	11 (100%)	33 (100%)	9 (100%)	53 (100%)
2: Change the thermostat's clock to 3:00PM	0	33 (100%)	4 (44%)	37 (70%)
3: Identify the setpoint and current temperatures	11 (100%)	33 (100%)	9 (100%)	53 (100%)
4: Identify the temperature the thermostat is set to reach on Thursdays at 9:00PM	0	14 (42%)	5 (56%)	19 (36%)
5: Put the thermostat into vacation mode	0	4 (12%)	9 (100%)	13 (25%)
6: Program the thermostat	0	33 (100%)	9 (100%)	42 (79%)

Task 1: Turn the thermostat off

This task evaluated the ease at which users were able to adjust the mode of their thermostat. Common modes include HEAT, COOL, OFF or AUTO. Because the interview was conducted during the heating season, it was assumed that most thermostats would already be in heat mode at the time of the interview. Hence, users were asked to turn their thermostats from heat mode to off. The HEAT-OFF switch is a common control opportunity for the majority of thermostats for at least the past 60 years [36].

Upon task completion, participants were directed to return their thermostats to the original setting (e.g., heating) so the remaining tasks could be completed.

Task 2: Change the thermostat's clock to 3:00PM

In this task users were asked to set the thermostat clock to 3:00PM. This time was selected arbitrarily as the objective was not to select the correct time, but to observe the ease at which the clock could be altered.

Task 3: Identify the setpoint and current temperatures

In this task users were asked to identify and read aloud the current and setpoint temperature displayed on the thermostat interface.

Task 4: Identify the temperature the thermostat is set to reach on Thursdays at 9:00PM

This task asked users to view their current schedule and identify the setpoint for a future time period. The period itself is arbitrary. Note that this task was only applicable to thermostats that were programmed prior to the study.

Task 5: Put the thermostat into vacation mode

When applicable, users were asked to schedule a vacation into their thermostat. The details of the vacation, including duration and start time, were arbitrary.

Task 6: Program the thermostat

In this task users were asked to program their thermostats to align with the schedule shown in Table 2.2 for every day of the week. The schedule time periods and setpoint temperatures were arbitrarily determined to establish a consistent goal for all participants. Upon completion of the task, the researcher assisted participants in restoring their thermostat to its original settings.

Table 2.2: Task 6 programming instructions

Time period	Setpoint temperature
8:00AM – 10:00PM	24°C
10:00PM – 8:00AM	22°C

2.3.3. Video recordings

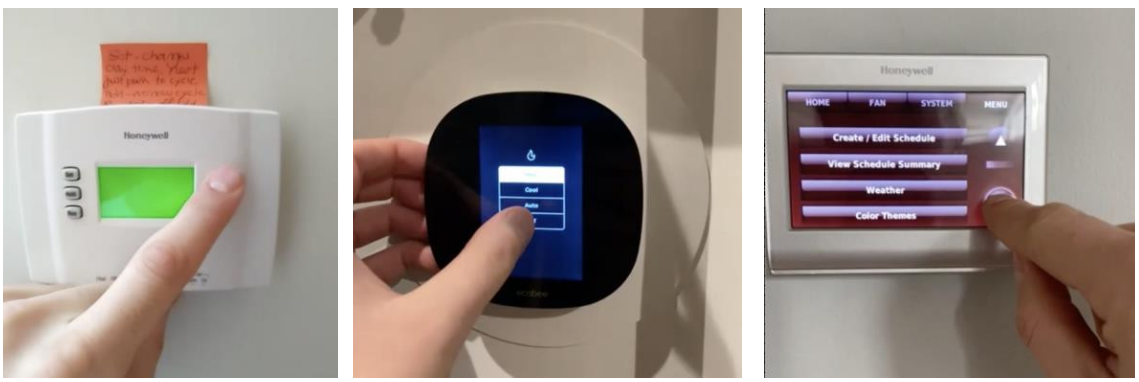


Figure 2.6: Screenshots from video recordings of participants interacting with their thermostat

When processing the video recordings for task analysis (example recordings shown in Figure 2.6), the researcher categorized the context for each task outcome (shown in Table 2.3). Additionally, instances of user confusion and frustration were flagged. To observe confusion, participants had to say the phrase, “I don’t know” or express a similar sentiment. Note that Figure 2.6 shows a note on top of the thermostat. While this note was not utilized during usability testing, the fact that the user felt the need to write additional instructions indicates a lack of walk-up usability and suggests that users require further guidance from their thermostat devices.

Table 2.3: Possible observations for task analysis

Attempt outcome	Observation	Description
Successful	Successful completion	User understands instructions and successfully completed the task
Inapplicable	N/A due to thermostat type	User understands task but cannot complete due to thermostat constraints
Unsuccessful	User resigns from task	User understands instructions but resigns from task
Unsuccessful	User thinks they can't accomplish the task	User understands instructions but (wrongfully) believes they cannot accomplish the task
Unsuccessful	User attempts the wrong task	User understands instructions but attempts the wrong task
Unsuccessful	User wrongfully thinks they were successful	User understands the task and thinks they have accomplished it but has not

The time to complete each task was recorded starting immediately after instructions were delivered to the participant until the instant that the task was either achieved (successfully or otherwise) or the participant resigned.

2.3.4. Questions to evaluate user understanding, misconceptions, and attitudes towards thermostats

Participants were prompted to explain their understanding and attitudes towards their thermostat. Feelings or behaviour related to the thermostat, such as confidence walking up to make an adjustment, was noted.

Participants were asked to note the brand/model of their thermostat interface(s) and if they found the interface aesthetically pleasing. The researcher established if the participants had bought the thermostat themselves and if it was being considered for replacement. The interview aimed to identify any common misconceptions in users, participants were asked about how their thermostat worked in relation to their HVAC system.

Further questions addressed household preferences; occupants' preferred setpoint temperature(s), whether they felt their thermostat displayed accurate information and to what degree the thermostat was able to regulate temperatures.

Finally, the priorities and affinity for energy savings in relation to thermostat use were established. Users were asked what features their "dream thermostat" would/wouldn't include and rated the usability of their current thermostat interface(s) on a 1-5 scale (Table 2.4)

Table 2.4: Participant-assigned usability scale

Rating	Description of usability
5	Easy to use and intuitive, no need for instructions
4	Easy to use but requires some direction from instructions
3	Minor difficulty, needed to consult some instructions
2	Moderate difficulty, even with instructions
1	Difficult to use even with instructions, could not complete tasks

Once the interview was completed, participants had the opportunity to ask the researcher questions or provide feedback.

2.3.5. Time and success (TS) metric

To align with previous work [7], [27], [36], [37], the TS metric was calculated for manual, programmable and smart devices. To account for success rates for individual tasks, completion rate, *s*, indicates whether the task has been accomplished. This combined with the time metric produces the "*M*" statistic, calculated as follows:

$$M = \frac{2s}{1 + e^{x}}$$

where,

$$x = \frac{t}{k}$$

$$s = \begin{cases} 0, if\ subject\ failed\ to\ complete\ task \\ 1, if\ subject\ completed\ task \end{cases}$$

$$t = time\ for\ subject\ to\ complete\ task\ (seconds)$$

$$k = 50\ (emperically\ determined\ constant)$$

Note that M is normalized between 0 and 1. The equation maps time $[0,\infty)$ to the interval $[1,0)$, where values closer to 1 have a shorter duration. Note that the success rate variable, s, also falls between 0 and 1. The TS metric is averaged for all tasks to produce a single value. Further explanation of the metric is provided in previous publications [36], [37].

2.4. **Results**

2.4.1. **Participant and household characteristics**

The sample included households built as early as the 1850s and as late as 2018. Figure 2.7 shows the distribution. Smart thermostats were observed in newer and older homes, but not those built between 1940-2000. Older homes are more likely to require HVAC system upgrades, likely resulting in more recent thermostat replacement/installation. Similarly, newer homes may be more likely to be built with more recent (smart) thermostats. While the sample is limited, the data might suggest that new builds and recently renovated homes are more likely to adopt a smart thermostat interface. In other words, occupants may be more likely to introduce smart thermostats in tandem with household construction or maintenance, rather than seeking out thermostat replacement opportunities individually.

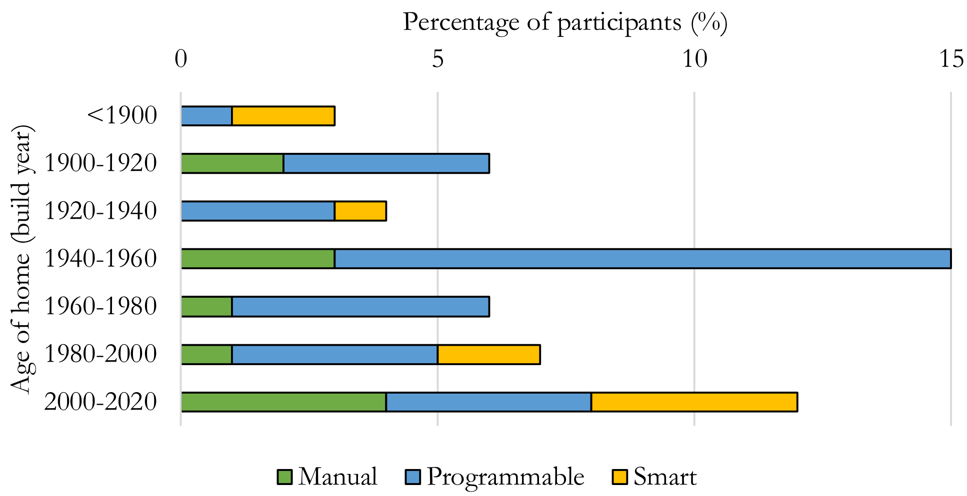


Figure 2.7: Age of home

The majority (75%) of participants lived in homes owned by themselves or their families. Note that this study occurred during the COVID-19 pandemic and that many students (the majority of which would fall within ages 20-30) had relocated to a family home at the time of the interview. This resulted in more participants living in owned, rather than rented, homes at the time of the study.

Most homes (71%) featured a forced air gas furnace. Other methods of heating include hot water systems (boiler and radiant heating) (16%), electric baseboards (8%), heat pumps (4%) and a propane fireplace.

Since previous experience operating the thermostat is likely to contribute to the user's knowledge of and familiarity with the interface, participants were asked if they were the main user of the device. The preponderance of respondents (88%) indicated that they were either the sole thermostat user or shared usage with other household members.

Most (76%) households included in the study were controlled by one thermostat. Homes with multiple heating systems (e.g., both radiant floor heating and forced air gas) often had more than one interface. The highest number of devices (11) was observed in a household with programmable and manual interfaces which controlled baseboard heaters and a propane fireplace.

Most devices (86%) were unobstructed and did not prevent user interaction. Some devices were located behind moveable objects, such as a wall-mounted bicycle. In this case, the user noted that the obstruction was irrelevant to the smart thermostat's usability, as they preferred to make setpoint adjustments on their phone application. This suggests that devices offering remote interaction may alleviate issues associated with thermostat placement.

While all participants reported previous experience and familiarity with thermostat operation, at least one user was unaware that they could replace their thermostat for a more favourable interface. Others noted that they had not considered the option of changing/choosing their thermostat interface. Most participants had not selected their current thermostat themselves – interfaces were frequently chosen by another party or already installed when the occupants moved in (Table 2.5)

Table 2.5: Thermostat selection

Thermostat acquisition	
Thermostat was:	**Participants** (percent of total)
Purchased	20 (38%)
Already there	25 (48%)
Suggested to the occupant during HVAC renovation/replacement	7 (13%)

When asked how often participants make a manual adjustment (e.g., alter setpoint, change program, etc.) on their thermostat interface, most respondents (60%) indicated less than once per day. Figure 2.8 shows the response, categorized by thermostat type. Many participants who made one to two adjustments per day noted that they adjust for nighttime setbacks. Over half (53%) of the respondents reported performing setbacks, however many noted that they do so for comfort, rather than energy savings. Manual thermostat users interacted with their interface more frequently, with 73% adjusting the setpoint at least once per day.

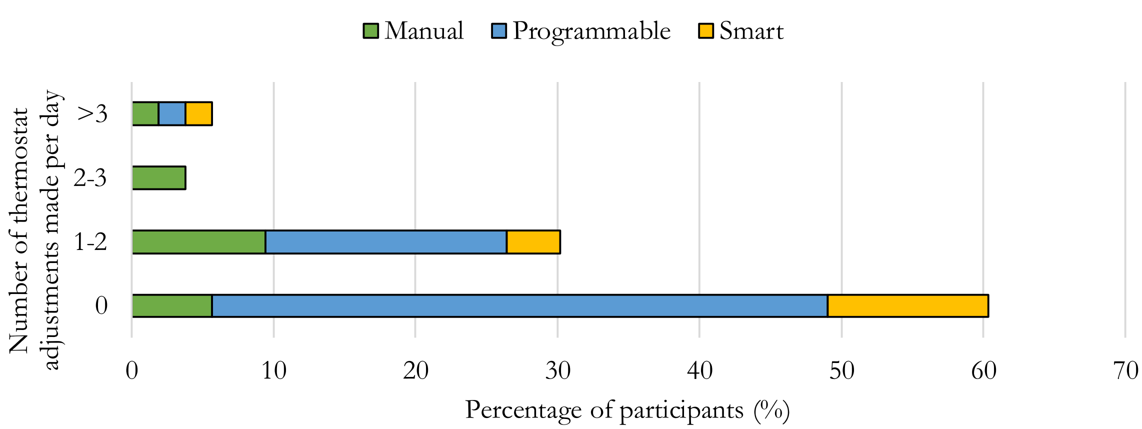


Figure 2.8: Thermostat adjustments per day, self-reported

Almost three quarters (72%) of participants indicated that they did not experience challenges when attempting to adjust their thermostat. This can be contrasted with the 43% that was unable to successfully complete task 6 – programming their smart/programmable interface. Moreover, 60% of devices with programming capability (programmable and smart) were operating in hold mode at the time of the interview. This suggests that many users may have never attempted to program their thermostats, but are familiar and comfortable making simple adjustments, such as turning the thermostat from heat to off, increasing/decreasing a setpoint, etc.

2.4.2. RQ1: How does the usability of manual, programmable, and smart thermostat interfaces compare?

Usability indicators for each task include the time to complete, interactions taken and the user-assigned usability rating. Table 2.6 shows the percentage of participants that successfully completed each task, based on the number of thermostats that could perform the task (e.g., 88% of all programmable thermostats with the capacity to have the task completed, did). Note that over half (59%) of the thermostats with the capacity for scheduling (i.e., programmable and smart devices) were operating in long-term hold mode at the time of the interview.

Table 2.6: Percentage of participants who were able to successfully complete each task, where N/A indicates that the thermostat type did not allow for the task to be completed

Thermostat	Task					
	1: Heat to off	**2: Change time**	**3: Read setpoint and current temperature**	**4: Check schedule settings**	**5: Set a vacation**	**6: Set a schedule**
Manual	100%	N/A	100%	N/A	N/A	N/A
Programmable	88%	70%	94%	64%	75%	52%
Smart	100%	75%	100%	100%	89%	78%

2.4.2.1. **Participant-rated usability**

At the end of the interview, participants were asked to rate their thermostat's usability with a five-point scale, described in Table 2.4 Results from this inquiry are shown in Figure 2.9. Manual thermostats received the highest usability rating. This is unsurprising as they offer users the fewest and most simplistic features. Programmable interfaces received the lowest usability rating – none of these devices received a rating of five (easy to use and intuitive) and 15% of participants with programmable thermostats indicating they could not use their interface, even with instructions. Smart devices performed better, with most users (78%) assigning a rating of four or five (easy to use). Note that the smart thermostat sample is significantly smaller than the programmable, and these results should be interpreted accordingly.

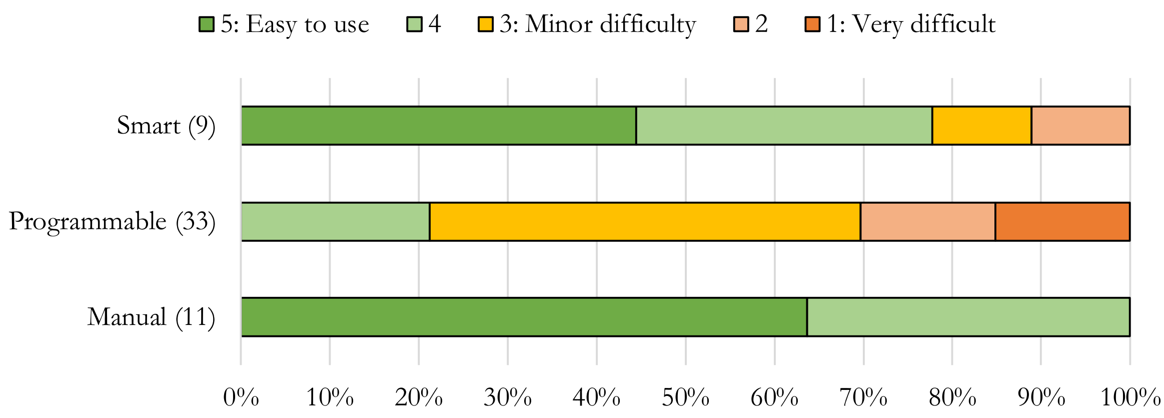


Figure 2.9: Participant-assigned usability rating for their thermostat interface

2.4.2.2. Usability task analysis: time and interactions

Tasks 1 (turning the device from heat to off) and 3 (reading the setpoint and current temperature) are considered simple and straightforward operations, requiring very little interaction with the interface. All thermostat types (manual, programmable and smart) had the capacity to complete these basic functions. The remaining tasks (2, 4, 5 and 6) are more complex, requiring more user interaction with the interface. The following discussion will refer to simple and complex tasks, accordingly.

To address the research question, the time required to successfully complete each task was recorded. Figure 2.10 compares the time required for each task for manual, programmable and smart thermostat interfaces.

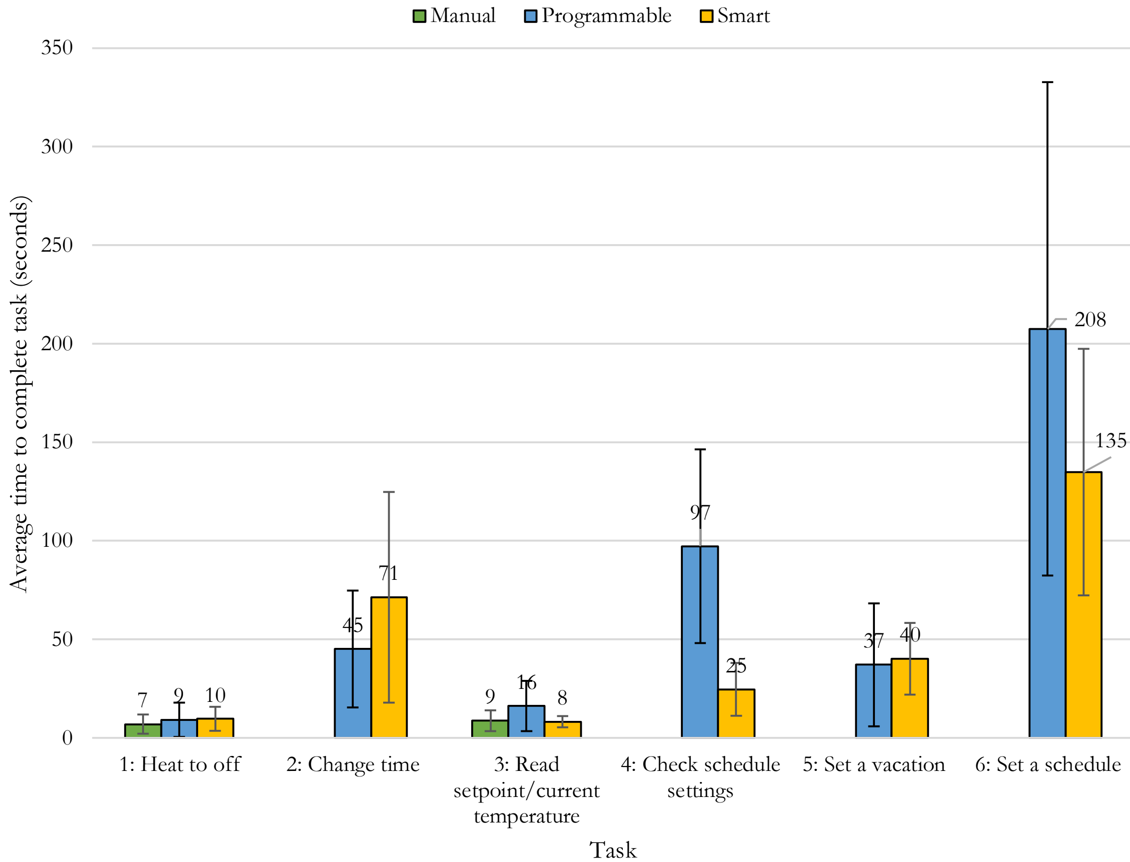


Figure 2.10: Average time to complete tasks 1-6 for each thermostat type. Error bars show one standard deviation.

Participants with manual thermostats completed the simple tasks (one and three) in the least time, however the variability between thermostat types was relatively small. When changing the clock (task 3) and setting a vacation (task 5), users with programmable thermostats took less time compared to those with smart devices. However, when considering activities related to scheduling (checking the schedule and programming the schedule – tasks 4 and 6, respectively), users with programmable thermostats required more time compared to those with smart interfaces.

Figure 2.11 shows the average number of interactions taken by users when completing each task for each thermostat type. Interactions refer to a count of the number of physical or digital (touch screen) buttons were pressed, sliders slid, and/or knobs turned, while attempting each task.

If the participant pushed a non-button area of their touchscreen, this was not included in the count. Furthermore, up/down selections, required when setting temperature or time, were counted as one interaction as the number of incremental adjustments does not directly relate to the navigation pathway required to complete each task.

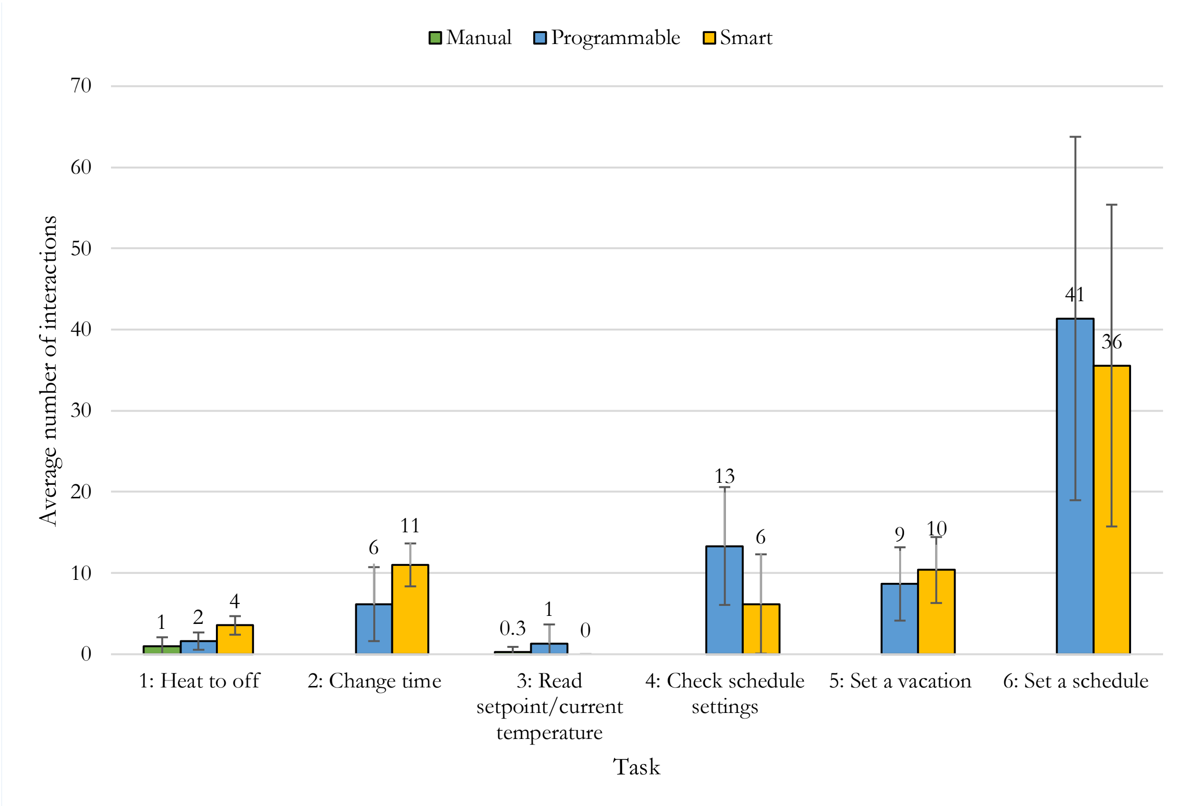


Figure 2.11: Average interactions required to complete tasks 1-6 for each thermostat type. Error bars showing one standard deviation.

The average number of interactions required to complete each task followed a similar trend as the average time to complete tasks. Manual thermostats required very few interactions from users. For tasks 2 (change time) and 5 (set a vacation), programmable thermostats required slightly fewer interactions. However, for tasks associated with programming, smart thermostats outperformed their programmable counterparts, requiring less user interactions.

2.4.2.3. **User confusion**

Figure 2.12 shows instances of confusion observed for each task. Approximately one third (32%) of users experienced confusion when attempting to program their thermostat. Note that only four programmable interfaces could set a vacation (task 5), likely contributing to the lack of confusion observed. Smart thermostat users experienced fewer instances of confusion compared to those with programmable interfaces. However, confusion was observed for a small number of users in tasks 2, 4 and 5. Manual thermostat users did not experience confusion when attempting usability tasks, however, manual devices were limited to tasks 1, 2 and 3.

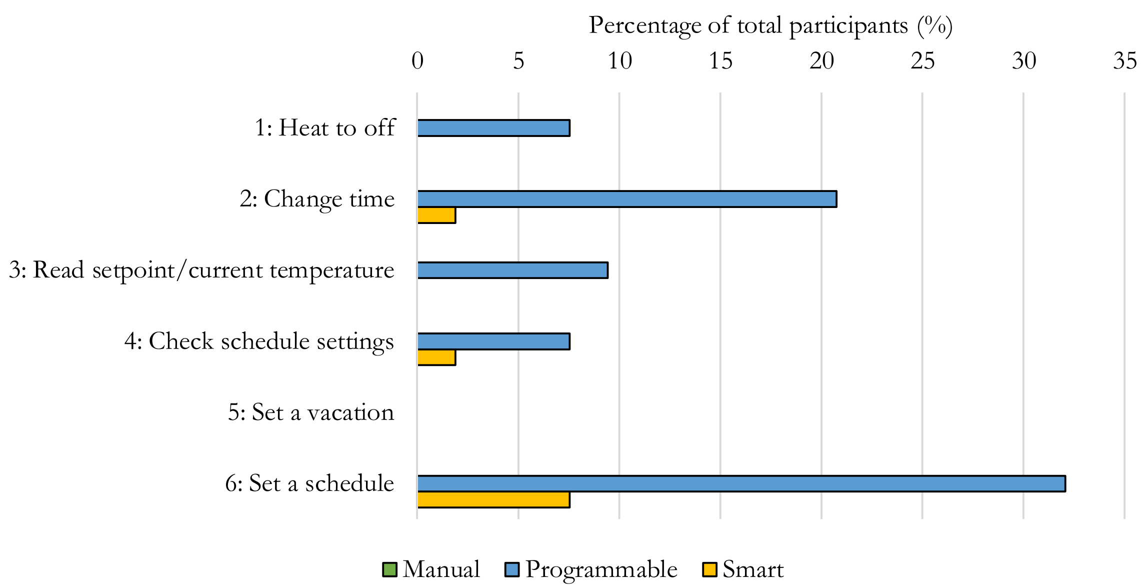


Figure 2.12: Observations of confusion for each task

2.4.2.4. **Observations for each task**

Observations, describing the context of incomplete task attempts, were noted for all task outcomes. Attempts were classified as defined by Table 2.3. Figure 2.13 shows the summary of observations based on task and thermostat type.

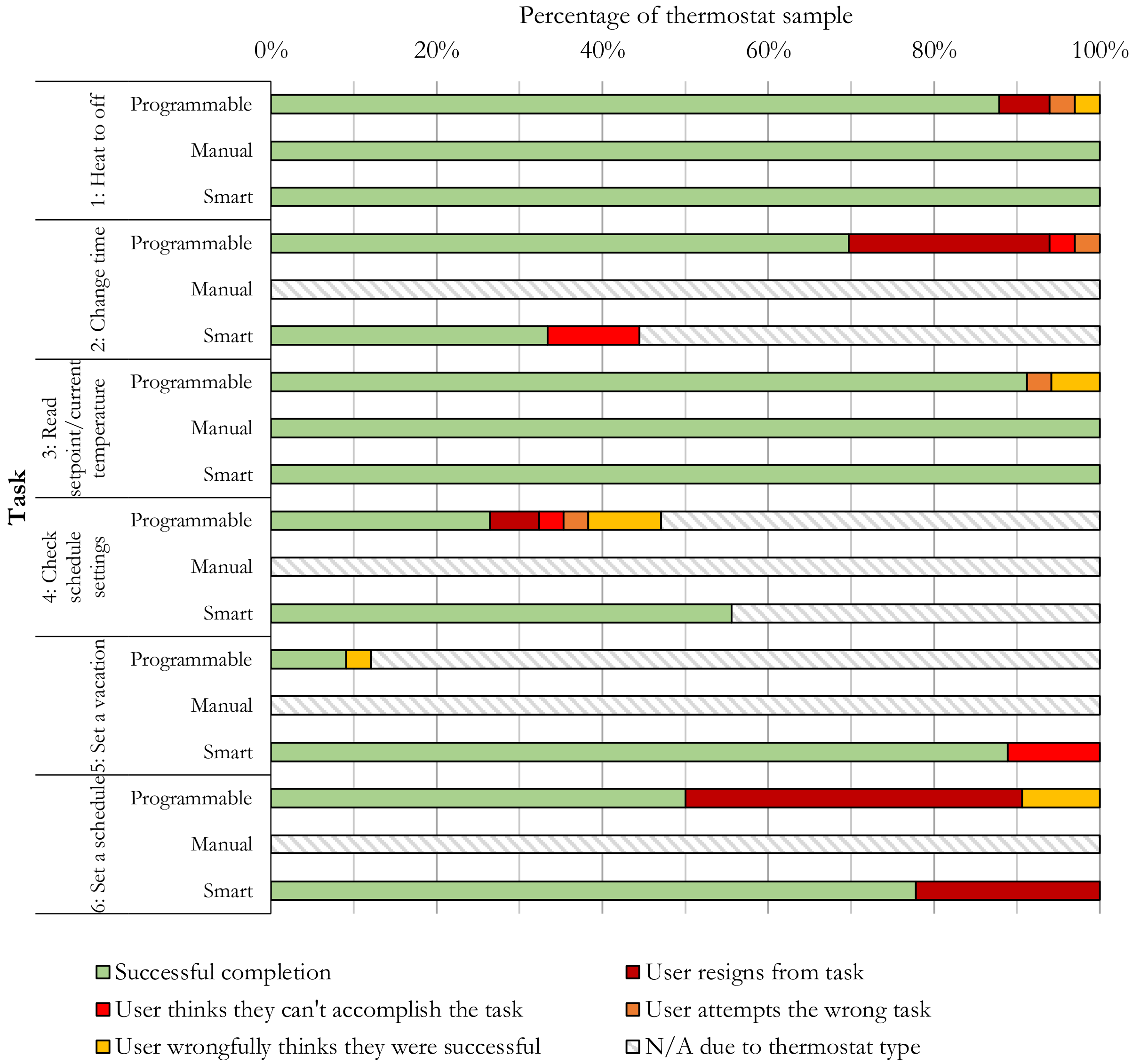


Figure 2.13: Observations for all thermostat types, organized by task

Task failure is defined as any outcome where the user did not successfully complete the task, assuming that the task was applicable to the thermostat. Significantly more instances of task failure were observed for programmable interfaces. Furthermore, programmable thermostats were the only device where users wrongfully thought they had successfully completed a task. Programming the thermostat can be considered the most difficult task, as it saw the greatest number of participant resignations.

2.4.2.5. **Time and success (TS) metric**

The TS metric was calculated for each task and thermostat type. The average TS values for all tasks, categorized by manual, programmable, and smart thermostats were 0.92 (for tasks 1 and 3), 0.46, and 0.61, respectively.

2.4.3. **RQ2: What types of HVAC and temperature feedback are provided in homes?**

Participants were asked how they know their heating system is working when they adjusted the thermostat. They could choose zero or more of: see, feel, or hear (refer to interview questions in the Appendix B). Figure 2.14 shows the response.

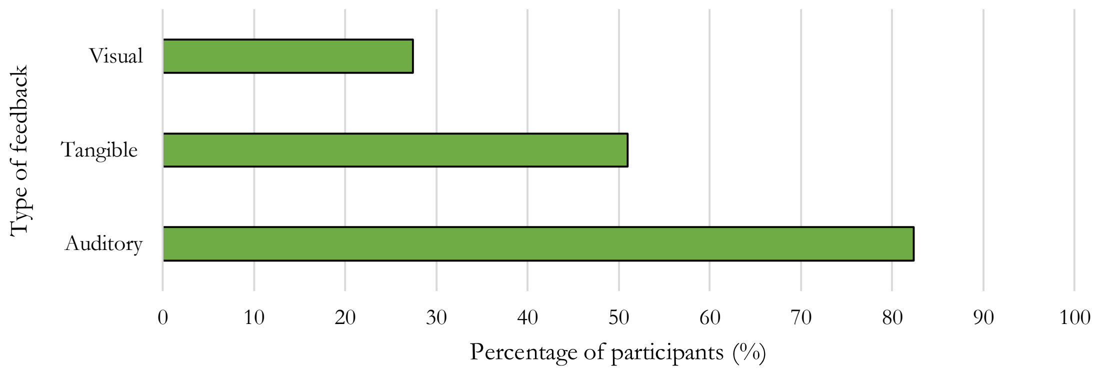


Figure 2.14: Reported feedback

Auditory feedback from the heating system was the most common, reported by 82% of participants. This included the sound of the furnace fan turning on, a "click" from the furnace after an adjustment was made, airflow through the vents and heating coils expanding (for an electric baseboard system). One participant was hearing impaired and noted that they could not observe this type of feedback.

Tangible feedback was the second most reported, noted by 51% of respondents. Participants reported feeling increased airflow from registers, heat emitted from radiators and sensing an overall temperature change. One participant noted that they frequently touch the diffusers to ensure the furnace is responding. Another reported that the furnace causes their living room coffee table to vibrate.

Visual feedback (besides the interface displaying the setpoint and current temperature) was reported by 28% of participants. Some thermostat interfaces indicated the furnace status with icons or phrases. For example, the Ecobee thermostat displays a flame icon while the furnace is supplying heat, another displays the message, "system heating". One participant noted that their thermostat also provides error messages if the furnace stops working.

The visual indicators on thermostat interfaces were reported as the least common method of feedback. The need for physical investigation, such as placing a hand over a radiator, demonstrates that current degree of feedback from thermostats is not adequate to indicate system status.

These results indicate that thermostats cannot easily (and perhaps should not) be the primary means for feedback to occupants regarding system functionality. The results also highlight the importance of HVAC system providing direct feedback. While industry is focusing on quieter and visually subtle HVAC systems [57], the current results suggest that trend should be pursued with caution. Unlike thermostats, which typically require occupants to walk up to them, tangible and auditory feedback from HVAC systems can often be felt throughout a space.

2.4.4. RQ3: What thermostat interface features do users desire/dislike?

Participants listed features they would include in their "dream thermostat" – feasible or otherwise. A comprehensive list of responses is available in Appendix A (Table A.2). It is important to note that responses are based on users' desires which stem from their current interface rather than general desires for a generic thermostat. Responses were categorized as desires for feedback, specific interface features or HVAC system features/functions.

Participants expressed an overall desire for increased feedback from their thermostat interface. This included metrics such as gas/electricity used per day and projections of energy costs per month/year/season. Others wanted real-time estimations of monetary savings, shown in tandem as users adjust setpoints/scheduling to provide an estimation of potential energy savings realized from different schedule configurations. Participants noted a lack of instructions, feedback, and guidance as the primary barriers to saving energy with thermostat usage.

Several users indicated that they preferred a touchscreen display. There was a strong desire for programmable thermostats supported by intuitive scheduling and more flexibility than the four time periods defined by Energy Star (wake, leave, return, sleep) [58]. Some participants wanted the interface to provide suggestions related to energy saving practices and indicate furnace filter status. Others wanted the interface to indicate outdoor conditions (temperature, weather, windspeed, humidity) and even a bus schedule. Participants noted the importance of a 'save' and 'undo' button to assist their navigation and comprehension. Others noted accessibility needs, including a large screen and text for increased readability, as well as the ability to brighten the display.

After the interview, a participant remarked that they would attempt to program their device, as the process made them feel more confident using their thermostat interface. Another user purchased a new thermostat post-interview after realizing their current interface was inadequate for completing usability tasks. These anecdotes suggest that users are willing to engage with energy saving practices but require further support, especially for complex tasks.

The overwhelming sentiment was that participants wanted the ability to adjust their thermostat remotely (e.g., using a smartphone application). Besides this, many users noted a desire for zone control in relation to heat distribution. Other methods of control included the integration of remote sensors and AI learning such that occupants did not need to make adjustments. One participant noted that they would like a text reminder to setback their setpoint when the house was unoccupied.

Overall, participants had extensive feedback for interface design. While some issues stretch beyond the limitations of the thermostat interface (e.g., zone control), there is clear desire for improved systems and potential for increased user engagement.

2.5. Discussion

2.5.1. Time and success (TS) metric

Comparing TS metrics with literature contextualizes the results. However, it is critical to note that this study differs from previous work, as participants performed usability tasks on their own, presumably familiar thermostat interfaces. Because the users' degree of familiarity differs from previous studies[36], [37], comparisons between TS metrics cannot be made without limitations.

While other research [27], [36], [37] has employed similar methodology, this study can only be contrasted with findings for programmable thermostats, as previous studies did not include manual devices and had significantly different tasks for smart thermostat testing [27]. Literature has focused on the comparison of individual TS metrics for a small sample of programmable thermostats with varying interface configurations [27], [36], [37]. Because this study featured 41 different thermostat interfaces, comparison of individual TS metrics for each device configuration is not feasible – thus, the TS metric has been averaged for each thermostat category and literature values are reported as a range.

Direct comparisons between findings and literature are challenging as previous studies omitted some tasks in their TS analysis and users in this study are more familiar with their interface. With these limitations in mind, Table 2.7 compares the averaged TS metrics for programmable and smart devices with values from literature. Manual devices could only achieve tasks 1 and 3 and had an average TS metric of 0.92.

Table 2.7: Time and success (TS) metric comparison based on tasks considered in literature

Tasks considered (Varies to allow comparison with literature)	**Source**	**Device type**	**TS metric** (Current study is averaged)
1, 3, 4	Meier et al. [37]	Programmable	0.2-0.7
	Current study	Programmable Smart	0.6 0.86
1-5	Perry et al. [36]	Programmable	0.1-0.6
	Current study	Programmable Smart	0.54 0.7
1-6	Current study	Programmable Smart	0.46 0.61

The TS metric serves as a valuable benchmark for usability testing; however, the literature is limited by inconsistencies in the tasks considered, the number of devices analyzed and few usability trials.

When considering tasks 1, 3 and 4 the average TS metric for programmable thermostats fell within the range described by Meier et al. [37]. Smart thermostats outperformed programmable devices, achieving a higher TS metric. This was observed again when comparing tasks 1-5 from the current study to previous work by Perry et al. [36]. Finally, the comparison of tasks 1-6 shows smart thermostats achieving a higher TS metric than their programmable counterparts, which is especially notable as task 6 is the most complex task, requiring the longest duration.

Note that the TS metric values from the current study are in the upper range of the comparable previous studies [36], [37]. This comparison suggests that user familiarity with the thermostats improves the TS metric values. Future studies should seek a more controlled experiment to directly compare the impact of familiarity and indicate the ecological validity of lab and home-based thermostat usability studies.

The TS metric values worsened as more complex tasks were included for both programmable and smart thermostats, emphasizing the importance of including complex tasks (e.g., programming the thermostat) in usability analysis.

Overall, user preferences and usage aligned with previous work. The number of participants who reported employing nighttime setbacks (53%) aligned with literature [20]. Participants commented that the buttons and font on their interface were too small, that it was hard to understand abbreviations or icons, which is a well-established problem [42], [45], [59]–[62]. The current research reinforces results from a recent review [24], suggesting that participants would opt for a smart thermostat with remote web or app-based interfaces, which may lead to more successful completion of schedules [25], [36].

2.5.2. Limitations

Although a mixed-methods approach allowed for more qualitative data collection and insight, the think-aloud methodology has the potential to lengthen task durations, as verbalizing actions increases participants' cognitive load [63]. If the goal is to compare multiple usability trials using the TS metric (and similar usability measurements), a different methodology may promote better consistency.

Interviews conducted over video-calls are susceptible to internet connectivity issues, influencing video and audio quality. While they provided an appropriate means to conduct research during the COVID-19 pandemic, recorded, in-person usability testing should be considered for future work.

This study was somewhat inconsistent with previous work [36], [37] for task 5: setting a vacation. In previous studies, participants were asked to put the thermostat into vacation, hold and/or away mode to achieve this task. In this work, participants were asked to put the thermostat into vacation mode, often resulting in confusion as most thermostats did not have a designated "vacation mode". For this reason, task 5 does not align with previous literature and most programmable thermostats were unable to attempt this task. Future usability trials should rephrase this task to reduce user confusion and allow more participants to complete the task.

The literature has employed additional usability metrics [36], including path length, button mash effect and confusion. While these are valuable for deriving insight from usability trials, they were not applied to this study due to the large sample size, variety of interfaces tested and conflicting methods for observing confusion.

The current study involved 51 participants with a majority being young adults. While the sample size is generally larger than previous studies, the overrepresentation of young participants may result in greater comfort with modern interfaces (e.g., touch screens) and technology in general. Future research should seek to expand on the participant sample. Furthermore, this study only considers North American homes and typical technologies. Thus, results may not be applicable to other building systems, such as those found in Europe or Asia.

Although programmable thermostats have been repeatedly recognized for their lack of usability, research focusing on usability testing and the application of usability metrics, such as the TS metric, is very limited. Future work should consider applying and extending the four metrics developed by Perry et al. [36] to a range of thermostats – manual, programmable and smart.

2.6. **Closing remarks**

This study investigated the usability of thermostat interfaces with a mixed-methods approach. When comparing the usability of manual, programmable and smart thermostats, participants using programmable devices took significantly more time and a greater number of interactions to accomplish complex tasks, such as programming a schedule. Programmable thermostat users also had the most user confusion and lowest self-reported usability rating. Manual thermostats had the highest TS metric score but could only accomplish two of the six usability tasks. Furthermore, 60% of their users adjust their thermostat less than once a day, suggesting that manual thermostats alone are not the solution to the downfalls of programmable thermostats. Smart thermostats received a higher TS metric value than their programmable counterparts and offer similar, if not more, functionality. These results cannot definitively address whether increasing levels of automation have a positive or negative effect on usability, however, the smart thermostat sample demonstrated a higher degree of usability in testing (TS metric, observations, confusion) and feedback (participant-assigned usability rating) compared to programmable devices.

The HVAC and temperature feedback observed in homes included auditory, visual, and tangible feedback. Auditory feedback was the most common among participants, and visual, the least. Few thermostat interfaces provided visual cues indicating system status.

Participants were asked what thermostat interface features they desired or disliked and what should be considered in the context of future design. By far the most common feedback suggests that users should have remote thermostat control through web or phone applications. Participants expressed desire for more feedback related to their energy consumption and projected expenditure.

Chapter 3

This chapter is under review as:

Tamas, Ruth, William O'Brien, and Mario Santana Quintero. 2022. "Evolving interaction: A qualitative investigation of user mental models for smart thermostat users." *Architectural Science Review*.

Chapter 3: A qualitative evaluation of smart thermostat users

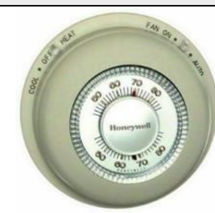
3.1. Introduction & literature review

Thermostat systems have developed significantly since they first emerged in 1906 [64]. Originally very simple devices, manual thermostats allowed users to select a specific setpoint, often limited to an interface with a few buttons or a dial. Overnight setbacks and other energy saving practices required diligent user adjustment. Programmable thermostats emerged in the 1980s [64] remedying the need for constant interaction. Programmable thermostats allow users to input a schedule of setpoint changes, eliminating the need for continuous adjustments to achieve setbacks. Interface design varied greatly, offering combinations of touchscreen, button and/or dial interaction. EnergyStar endorsed programmable thermostats in 1995, suggesting they could save users nearly $200 a year [37]–[39]. However, later studies [3], [5], [18], [19] demonstrated that programmable thermostats rarely met their energy saving potential. Programmable thermostats have been documented in many instances to have low usability, leading to frustration, discomfort, and increased energy use [4], [12], [50], [13]–[17], [42]–[44]. Both manual and programmable thermostats require the user to be at the thermostat interface to interact with their heating system. Ecobee launched the first smart thermostat in 2008 [65] with the goal of addressing this issue, offering Wi-Fi compatibility. Since then, smart thermostats have further developed, offering a wide variety of new features (e.g., Wi-Fi connectivity, remote adjustment, zone control, integration with smart home devices, occupancy detection, schedule learning, energy use feedback, etc.). North America represents the largest market for smart thermostats [66]. Since 2008, the adoption of smart thermostats has rapidly increased. In 2019 alone, the installation of smart home systems increased by 28.5%, amounting to use in nearly a third of North American households [67].

As smart thermostats become commonplace, they present unique opportunities for interaction [31]; most allow remote control (e.g., temperature adjustments through a smartphone application). Other systems use occupancy detection through motion sensing (e.g., Nest) or geo-fencing technology (e.g., Tado) to regulate temperature and setbacks. Some smart thermostats (e.g., Nest) attempt to learn occupants' schedules and routines through adjustment behaviour. While these new features appear promising, they introduce a significant degree of complexity to thermostat operation.

For this thesis, manual, programmable and smart thermostats are considered three distinct device categories. Table 3.1 gives a brief description and shows examples of manual, programmable and smart thermostat devices.

Table 3.1: Thermostat categories

Thermostat type	Description	Sample thermostats
Manual	All non-programmable thermostats.	 
Programmable	Thermostats allowing the user to schedule a program with automated setpoint changes. No further capabilities.	 
Smart	Thermostats that extend beyond the definition of manual and programmable devices (e.g., enables remote control, motion detection, etc.)	 

3.1.1. User mental models

The way users expect their thermostat and home heating system to function appears to dictate how they interact with it [2], [28]–[31], significantly more than how the system actually works. This is especially relevant, given that user behaviour has been shown to significantly influence building performance [6], specifically in relation to thermostat operation [7], [68], [69]. Thermostats represent the primary point of contact between occupants and their heating, ventilation, and cooling (HVAC) system in many buildings; thus, thermostat interfaces contribute significantly to a user's understanding and conceptualization of operation. In relation to thermostat interfaces, the user's understanding has demonstrated implications for temperature setbacks and related energy consumption [29]. In Canada, 65% of a household's energy consumption is attributed to heating and cooling [1]. Therefore, understanding how users imagine their HVAC system to work is critical to anticipating their behaviour and designing effective user interfaces.

To investigate how users understand or imagine their HVAC system to work, previous literature has employed user mental model (UMM) analysis. A UMM describes how a user imagines or perceives a system's components to relate, offering insight into the user's logic and behaviour [70]. Historically, UMMs have been represented by visual diagrams, illustrating the relationships between system components [28], [29]. As an established method in human-computer interaction research, UMMs have been employed in various domains (domestic [29], transport [47], military [71]) to support interface design [70], [72], usability [70], [73], [74], and sustainable behaviour [28], [29], [75]. In relation to thermostat interfaces, UMMs have been used to evaluate manual [29] and programmable [28] thermostat users, with unique results linked to the respective thermostat categories. Extensive work has been performed to explore UMMs in relation to manual and programmable thermostat operation [2], [28], [29], [69], [70], [76], [77] due to the implications for energy consumption.

Kempton pioneered UMM investigation in relation to thermostats, conducting a study with manual interfaces in 1986 [29]. Through this, he confirmed that thermostat interface configuration implicitly communicates the function of the HVAC system. For example, a manual, dial-like interface was demonstrated to produce a 'valve theory' schema in some participants. Here, users operated the thermostat as if increasing the setpoint would increase the rate of heat delivery, like a gas valve. Since then, the literature has expanded to include the UMMs for programmable thermostat users [2], [28], [30], [31], with a focus on how usability issues contribute to the formation of inaccurate UMMs and the elements that prevent users from interacting with their devices as intended. This literature further demonstrates how users' conceptualizations influence the use of thermostat devices, and the disconnect between designer's intentions and users' interpretation.

3.1.2. **Shared theories among thermostat users**

UMM investigation has demonstrated patterns and groupings in how users expect their thermostat and HVAC system to function (e.g., multiple users expecting their thermostat to operate as a valve [29]). When consistent cognitive frameworks are identified, they are denoted as 'shared theories' (e.g., users who expect thermostats to operate as a valve [29] adhere to valve theory). Table 3.2 presents the four shared theories identified in manual and programmable thermostat users and describes for each how the user would imagine the heating system to function. Note that shared theories are discussed in relation to heating and heat delivery but can be applied to both heating and cooling.

Table 3.2: Summary of shared theories previously identified in the literature.

Shared theory	Description
Feedback [29]	Users believe that the thermostat measures and maintains the user-inputted setpoint. The system delivers heat until the setpoint temperature is achieved. Once the setpoint is achieved, the system stops delivering heat until a differential between setpoint and current temperature is measured.
Valve [29]	Users believe that thermostat adjustment regulates the intensity of heat delivery from the heating system. Turning the thermostat to its maximum temperature will deliver heat at a higher rate. The onus of temperature control and maintaining the setpoint is on the user.
Timer [70]	Users believe that the thermostat operates on a timer-like setting. When the user wants the heating system to run for a short period, they will select a setpoint slightly higher than their current temperature. For a longer period, the user will select a setpoint significantly higher than the current temperature. The user does not adjust the setpoint with target temperature in mind. The onus of temperature control and maintaining the setpoint is on the user.
Switch [2]	Users believe the thermostat is used as an on/off switch for the heating system. When the user wants to activate or deactivate the heating system, they select a setpoint above or below the current temperature, respectively. The user does not adjust the setpoint with target temperature in mind. The onus of temperature control and maintaining the setpoint is on the user.

In contrast with shared theories, institutional theory denotes how the system truly functions. With respect to thermostat systems, feedback theory best reflects institutional theory, presenting a simplified model of the heating system. Valve, timer, and switch theory all diverge from institutional theory, assuming the user measures and regulates the heat delivery. The implications of feedback, valve, timer, and switch theory have been noted in literature, but are limited.

Kempton's work [29] evaluated manual thermostat, often featuring rotating, dial-like interfaces [78]. Using UMMs, he identified feedback and valve theory. While feedback theory was considered more 'correct' in relation to institutional theory, users operating their manual thermostat as a valve were found to be more likely to employ overnight setbacks. This was mostly related to valve theory users having more consistent thermostat interactions and adjustments, as they imagined they were controlling the intensity of heat and were therefore more attentive. More recently, Norman [70] and Peffer et al. [2] identified timer and switch theory, respectively, in programmable thermostat users. Unlike Kempton [29], neither Norman [70] nor Peffer et al. [2] directly refer to the studies that informed their shared theories, or describe specific energy-related behaviour or outcomes.

UMMs and shared theories have been established for manual and programmable thermostats, however, there is a widening gap in the literature in relation to current devices. Thermostat interfaces look very different now than they did in the most recent UMM investigation [2], ten years ago. The first Canadian smart thermostat emerged in 2007 [40]. Since then, smart thermostats have developed significantly, offering a new breadth of features and novel interface layouts. Research has suggested that more informative control layouts lead to more effective operation [31] however, this has yet to be demonstrated in the UMMs of smart thermostat users. Additionally, smart thermostats have increased the available methods of user input, enabling remote interaction through smartphone applications and verbal interaction through smart home systems. Therefore, while UMMs exist for manual and programmable thermostat users, the novelty of smart thermostat interfaces demands a new round of investigation.

This thesis contributes to the existing body of literature by presenting a UMM investigation with smart thermostat users. A series of semi-structured, qualitative interviews were employed to produce and categorize UMMs. The structure of the inquiry closely follows the work of Revell and Stanton [28], who performed a comprehensive investigation of UMMs with programmable thermostat users. Adopting a similar methodology allows the results to be compared not only in terms of shared theories, but also via thermostat category (smart, programmable, manual), usability, and participants' attitudes and preferences. The investigation categorizes UMMs using the shared theories described in Table 3.2.

3.1.3. Thermostat usability issues contributing to inaccurate UMMs

Revell and Stanton performed an extensive evaluation of programmable thermostats [79]. They focused on identifying UMMs [28] and redesigning the thermostat interface to address usability issues and promote more accurate UMMs [80]. Results were positive, demonstrating that interface design influences user comprehension, behaviour, and the related energy consumption of buildings [31], [34]. During this investigation, they argued that usability issues found in programmable thermostats would persist in smart thermostats, and that instead of enabling better UMMs, smart thermostats increase interface complexity while failing to support users [80]. The following usability issues were noted for their sample of programmable thermostats and were deemed to contribute to inaccurate UMMs.

- Programmable thermostat interfaces do not effectively communicate the feedback function of the device.
- Programmable thermostat interfaces do not communicate where the temperature is measured (e.g., at the thermostat).
- Dial control triggers a valve schema [81], resulting in a UMM that reflects valve theory.
- The programming procedure and interface is confusing and discourages engagement.
- The program schedule is hidden.
- The link between thermostat and heating system operation is not clear; users assume that scheduled periods equate to active heating periods.

This thesis contributes to the limited literature related to smart thermostat usability. It applies established human-computer interaction research methods to address the previous claims that smart thermostats do not support effective UMMs. Furthermore, the results and methods of this work contribute to future remote usability studies and provide a framework for the evaluation of other types of smart home systems.

3.1.4. Chapter overview

Building upon previous work featuring manual and programmable thermostats [2], [28], [29], [70], this chapter investigates the UMMs of ten Canadian smart thermostat users, interviewed during the heating season. Three case studies are discussed in depth, and participant responses are presented in the context of user understanding as well as thermostat use and usability. Attitudes and preferences for smart thermostat interface design are disseminated. Thermostat schedules for the ten participants are presented and contextualized with relative heating degree days. All research and data collection was conducted remotely, contrasting traditional UMM interview and diagramming methods (i.e., [28], [29]) with online software and resources. The following research questions are addressed:

RQ1: Which UMMs/shared theories exist in the group of smart thermostat users?

RQ2: What attitudes are held towards smart thermostat devices and how does the adoption of smart thermostats relate to energy consciousness/consumption?

RQ3: How effective is the use of online software and remote strategies in the investigation of UMMs?

3.2. **Methodology**

3.2.1. **Procedure**

The study employed semi-structured interviews and a diagramming exercise to investigate research questions and produce UMMs. Interviews were conducted and recorded using Zoom video calls. Recording was done with permission and supported the accurate interpretation of participant responses/sentiments, specifically in clarifying intention and the interpretation of UMMs. The average interview duration was approximately one hour, ranging between 40 and 90 minutes, determined primarily by participant enthusiasm or need for clarification. The methodology mirrored the process taken by Revell and Stanton [28], utilizing virtual post-it notes to establish participant-initiated terminology for various components of the heating system (e.g., "furnace"), drawing relationships between components and the inclusion of rich contextual information in the form of additional text, notes, etc. An online whiteboard software tool, Miro, was employed to serve the UMM diagramming exercise. The software allowed the researcher and participant to use the same whiteboard space and see each other's cursors in real-time, interacting with, pointing to, and adjusting components of the UMM. An example of the Miro interface is shown in Figure 3.3.

The UMM investigation and categorization followed the methodology of Revell and Stanton [28]. Analysis tables for categorizing participant responses when reviewing interview records were used to interpret and define shared theories within UMMs. If the UMM showed a feedback loop between the thermostat and heating system, where the thermostat measured the temperature and maintained the setpoint (informed by Section III. and IV.), it was considered to display feedback theory. If the UMM and usage scenarios (Sections III. and IV.) did not reflect the thermostat as controlling and monitoring the current temperature, and instead showed the onus of temperature control on the user, the UMM was evaluated for valve, timer, and/or switch theory.

The interview had four sections:

I. **Pre-interview survey for demographic information**

An online survey form, detailing basic demographic information including gender, occupation, country of residence, experience with home heating systems, as well as household type (e.g., single-detached), characteristics (e.g., sub-metered vs. bulk-metered energy), familiarity (e.g., number of years at residence), and number of occupants.

II. Self-reported user behaviour

Interview questions were based upon previous work by Revell and Stanton [28] with additional inquiry unique to smart thermostats and energy consumption/attitudes. Participants were asked how they used their heating system over a typical week during the (current) heating season.

III. Usage scenarios

Participants were presented with four different home-heating scenarios and asked how they would behave or might imagine themselves behaving. For each scenario, participants noted how likely they were to adjust the thermostat and how frequently they encountered each scenario. Questions evaluated users' understanding of thermostat control, with specific probes for shared theories (e.g., increasing the setpoint temperature to increase the rate of heat delivery demonstrates valve theory behaviour, and is probed in scenario 1.a).

Scenario 1: It is winter. You come home to a cold house and want to warm up. What do you do? Do you do anything differently if…

a. You want it to heat up quickly
b. You want the heat to stay on for a long time
c. You want the heat to come on immediately
d. You want to achieve a specific temperature

Scenario 2: You have been working at home and feel cold from sitting at your desk. What do you do?

Scenario 3: The heating is on, and you have been rushing around and now feel uncomfortably warm. What do you do?

Scenario 4: You are sitting at home, and it is comfortably warm. What do you do?

IV. User mental model diagramming exercise

Throughout the interview, participants indirectly discussed the components of their home heating system. The participant-initiated terminology was recorded on virtual post-it notes during the interview, and then related in the UMM diagramming exercise. Heating system components were linked with arrows, defined in terms of location and purpose, and assigned degrees of certainty (i.e., if a participant wasn't sure if two components were linked, the link was given a red marker).

Figure 3.1 shows the key for UMM construction, highlighting components and links that represent points of user interaction.

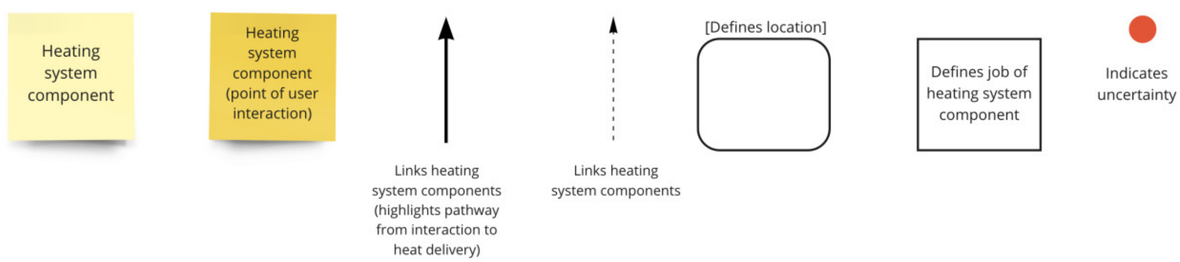


Figure 3.1: Key for UMMs

The diagramming process was semi-structured, allowing participants to recall new elements of the heating system throughout the exercise. For each component (e.g., vents), the participant was asked "what is the job of the [component]?", "what do you think the [component] is connected to?", "what do you think happens when you adjust the [component]?" Once the diagram was complete, the researcher listed each component individually, and paraphrased what had been annotated. The participant confirmed or amended as required, until both parties had an agreed understanding of the UMM.

3.2.2. Recruitment

Participants were recruited through various means; the researcher contacted individuals previously identified as smart thermostat users from an earlier study [82] (yielding 7/10 participants). The remaining were recruited through advertisements via email and social media. Each participant was compensated $30 CAD.

To be included, participants were required to be adults, living in a Canadian home with a smart thermostat, with the ability to receive video-calls and internet access. Participants represent a diverse sample in age and gender, with varying occupations and experience. Except for one respondent (participant D), the sample can be considered non-experts in relation to thermostat and heating system usage. While not intentional, all participants operated forced-air gas furnaces. Gas furnace heating systems are very common in Canadian and Ontario homes, reported in 52% and 80% of dwellings, respectively, in the most recent survey of household energy use (SHEU) [83]. It should be noted that these systems can feature variable rates of heat delivery, depending on if they are single or dual stage.

Data collection occurred during the Canadian heating season over the course of two months, beginning in November 2021. The Eastern Ontario region experiences cold winters. During the three-month investigation, average monthly temperatures ranged from -14 °C (January 2022) to 1 °C [84].

3.2.3. Mitigating Bias

As a qualitative method, UMM investigation is highly vulnerable to bias. Revell and Stanton [77] addressed this thoroughly, producing a methodology to assist researchers identify bias and strategies for mitigation. This includes consideration of participant and analyst background knowledge or experience and communication styles, as well as opportunities for bias when constructing the UMM and within methods of analysis.

Following Revell and Stanton's method for exploring the relationship between UMMs and behaviour (the quick association check [77]), the interviewer asked participants about their background experience to contextualize responses. To reduce anxiety and encourage free dialogue, positioning texts (i.e., a consistent, pre-determined communication outlining expectations) were provided at the start of each interview step, encouraging participants to respond based on their own understanding or conceptualization, rather than what they might anticipate is the "correct" answer. Participants were frequently given the opportunity to amend responses and each component of the UMM was assigned a level of confidence. The diagramming exercise didn't use any pre-prepared templates and terminology for components of the heating system was initiated by the participant. Participants had multiple opportunities to verify UMM components and connections (e.g., the 'thermostat' is related with an arrow drawn towards the ‘furnace’) and were asked to paraphrase their system diagram to ensure accurate interpretation by the analyst.

3.2.4. **Analysis of outputs**

The investigation process produced outputs during and post interview. The entire duration of the video-call was recorded and reviewed to further inform and confirm participant responses(s). Table 3.3 addresses each research question individually, linking it to the supporting outputs from the investigation.

Table 3.3: Research questions, linked to methodology and results

	Research question	Informed by:
1	Which UMMs/shared theories exist in the group of smart thermostat users?	- Self-reported user behaviour: typical thermostat usage and schedule(s). - Usage scenarios: probes for shared theories. - UMM diagrams: clarifying inter-relationship(s) for heating system components. - Interview recording: clarified participants' intention(s), verbalization(s), and terminology.
2	What attitudes are held towards smart thermostat devices, and how does the adoption of smart thermostats relate to energy consciousness/consumption?	- Self-reported user behaviour: direct questions about thermostat usage priorities. Discussion of smart thermostat advantages and disadvantages. - Self-reported thermostat schedules: used to compare participants and calculate relative heating degree days.
3	How effective is the use of online software and remote strategies in the investigation of UMMs?	- Reflection of online interview and UMM diagramming exercise. - Review of limitations encountered while working remotely.

A flowchart of the study procedure is available in Figure 3.2, outlining the data gathered by the interview process, the post-interview analysis, and links to results and research questions.

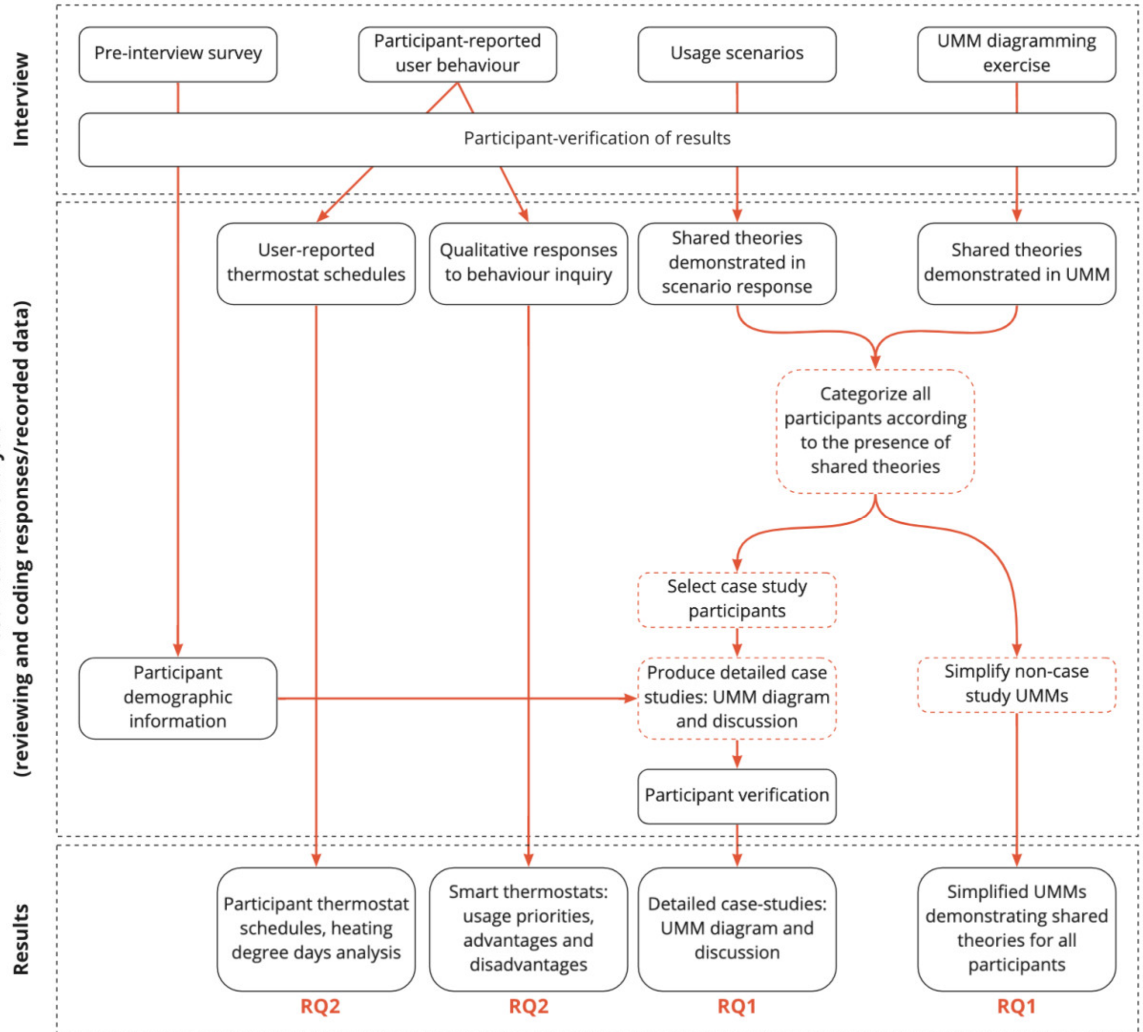


Figure 3.2: Research procedure flowchart

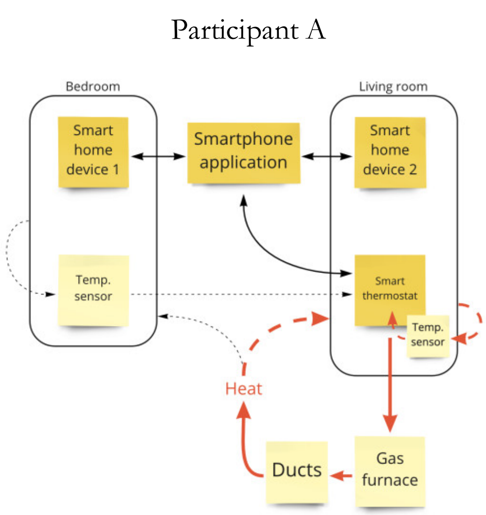
3.3. Results & discussion

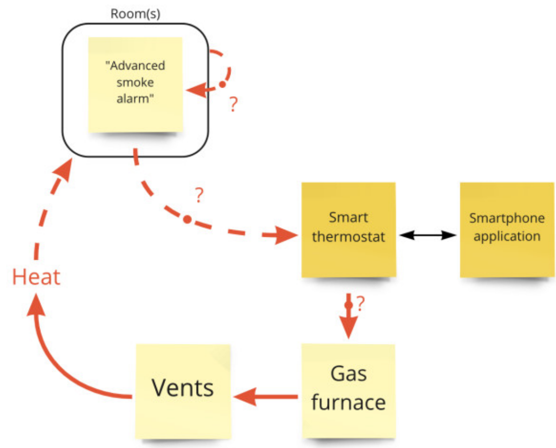
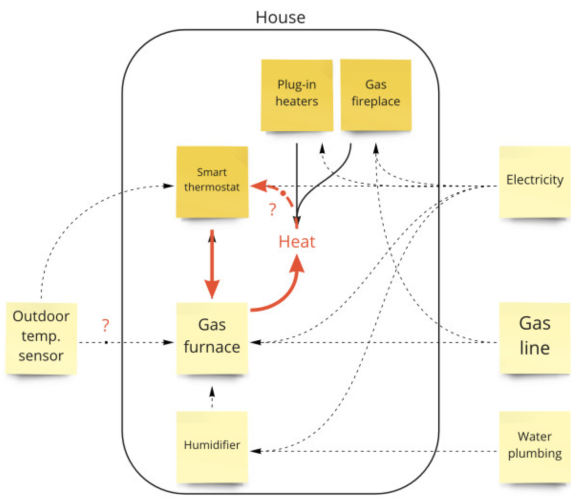
3.3.1. RQ1: Which UMMs/shared theories exist in the group of smart thermostat users?

3.3.1.1. Overview of smart thermostat UMMs

The first research question focuses on the identification of shared theories within the UMMs. Table 3.4 provides an overview of UMMs for all ten participants. The UMMs are simplified (diagrams exclude contextual information and descriptions of heating system components, consistent terminology is presented, question marks are used to mark areas with uncertainty) and the relationship between temperature sensor (thermostat) and heat delivery is emphasized to illustrate the feedback theory observed for all participants. Feedback theory describes a relationship where the thermostat measures current temperature against a setpoint and determines whether to call for heat, forming a feedback loop.

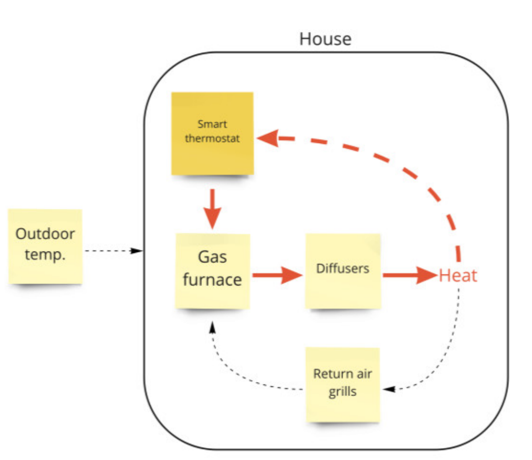
Table 3.4: Simplified UMMs for smart thermostat users. Red pathways illustrate the relationship between temperature measurement and heat delivery.

<table>
<tr><th>UMM diagram</th><th>Description</th></tr>
<tr><td>
</td><td>- Smart home system integrated with heating system.
- Two smart home devices installed, enabling verbal interaction.
- Supplementary temperature sensor in bedroom zone.
- Smartphone application enabling remote control.
- Cycle between temperature measurement (at smart thermostat) and heat production, indicating feedback theory schema.</td></tr>
</table>

Participant B 	- Uncertainty about inter-relationships of heating system components. - Uncertainty about where temperature was measured. - Smartphone application enabling remote control. - Cycle between temperature measurement (at "advanced smoke alarm") and heat production, indicating feedback theory schema.
Participant C 	- Supplementary heating devices (plug-in heaters, gas fireplace). - Large scope (electricity, gas line, water plumbing). - Uncertainty about where temperatures (indoor and outdoor) were measured - No interaction with smartphone. - Cycle between temperature measurement (at smart thermostat) and heat production, indicating feedback theory schema.

<table>
<tr>
<td>

Participant D

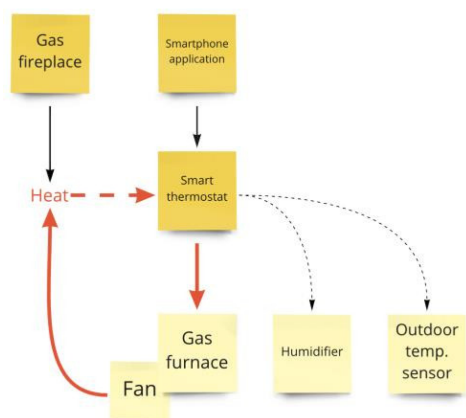


</td>
<td>

- No interaction with smartphone.
- Cycle between temperature measurement (at smart thermostat) and heat production, indicating feedback theory schema.

</td>
</tr>
<tr>
<td>

Participant E

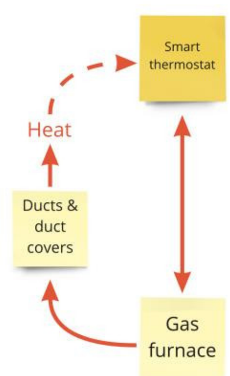


</td>
<td>

- Supplementary heating device (gas fireplace).
- Smartphone application enabling remote control.
- Cycle between temperature measurement (at smart thermostat) and heat production, indicating feedback theory schema.

</td>
</tr>
<tr>
<td>

Participant F



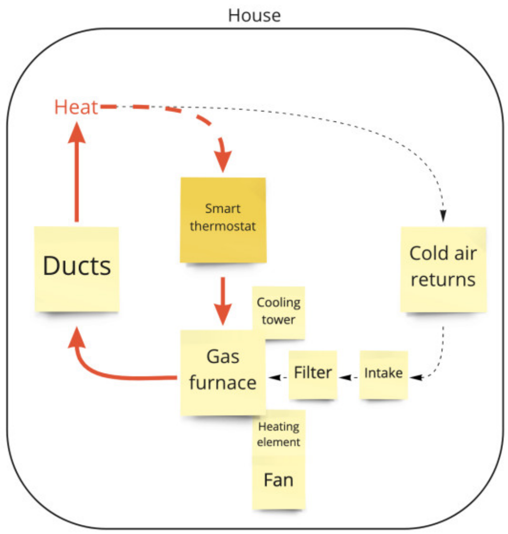
</td>
<td>

- No interaction with smartphone.
- Cycle between temperature measurement (at smart thermostat) and heat production, indicating feedback theory schema.

</td>
</tr>
</table>

<table>
<tr>
<td>

Participant G

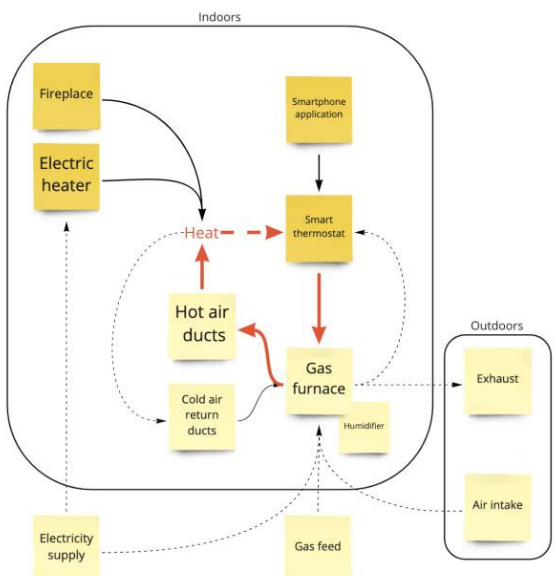


</td>
<td>

- Detailed gas furnace description.
- No interaction with smartphone.
- Cycle between temperature measurement (at smart thermostat) and heat production, indicating feedback theory schema.

</td>
</tr>
<tr>
<td>

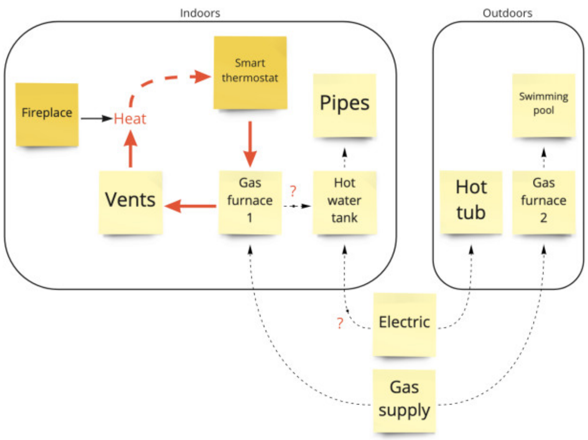
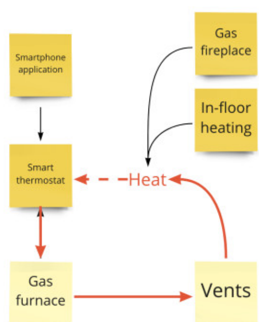
Participant H



</td>
<td>

- Supplementary heating devices (electric heater, fireplace).
- Large scope (electricity supply, gas feed).
- Smartphone application enabling remote control.
- Cycle between temperature measurement (at smart thermostat) and heat production, indicating feedback theory schema.

</td>
</tr>
</table>

<table>
<tr>
<td>Participant I

</td>
<td>- Supplementary heating device (fireplace).
- Large scope (electric, gas supply).
- Uncertainty about whether hot water tank was gas or electric powered.
- Heating system extends to hot tub and swimming pool.
- No interaction with smartphone.
- Cycle between temperature measurement (at smart thermostat) and heat production, indicating feedback theory schema.</td>
</tr>
<tr>
<td>Participant J

</td>
<td>- Supplementary heating devices (gas fireplace, in-floor heating).
- Smartphone application enabling remote control.
- Cycle between temperature measurement (at smart thermostat) and heat production, indicating feedback theory schema.</td>
</tr>
</table>

Adherence to feedback theory was consistently demonstrated in the diagramming exercise and in the responses to usage scenarios. All ten participants noted that it is the thermostat, rather than the user, that regulates the delivery of heat. While all UMMs demonstrated feedback theory, they differed in the scope, certainty, terminology, and description of heating system components.

The energy outcomes related to shared theories has evolved with new devices. From research with manual thermostats, Kempton [29] suggested that users with a valve theory UMM were more likely to employ temperature setbacks, resulting in lower energy consumption. However, this result assumes that users with feedback theory UMMs do not employ setbacks. The sample of smart thermostat UMMs demonstrates feedback theory and thermostat schedules indicate that most (8/10) users employ setbacks, suggesting that smart thermostat users with a feedback theory UMM do not consume more energy due to lack of setbacks. Feedback theory provides users with the most accurate depiction of the heating system, compared to other shared theories, and does not discourage energy saving behaviour, such as overnight setbacks. Thus, smart thermostats appear to elicit an accurate and energy efficient UMM.

The heating system components (e.g., "smartphone application") included in the UMMs reflected the devices users engaged or interacted with. Half of the respondents (participants A, B, E, H, J) included the smartphone application in their UMM diagram, which corresponds to the households that employ remote interaction. Two participants noted they used the different interfaces for different tasks; the remote interface better supported scheduling activities and the wall-mounted interface was used more frequently to make small setpoint adjustments. All but one participant (participant A) lived in sub-metered homes, where occupants are responsible for the energy consumed, rather than paying a flat rate (bulk-metered).

From the group of ten, three participant UMMs are further explored as case studies. The UMMs are presented with greater detail (following the key from Figure 3.1). The first, participant A, was the only participant to have integrated a smart home system with their smart thermostat. They were keen to adapt technology, installing smart home devices in two rooms to enable verbal interaction, as well as an additional temperature sensor to allow targeted zone heating. Participant A represents an interesting UMM case study because of the use of additional smart technology and new opportunities for interaction (e.g., verbal communication with the smart home). The second case study highlights participant B, who was selected due to a high degree of misconception, relative to other participants. Participant B experiences uncertainty related to the communication between heating system components and where conditions (i.e., temperature) are measured. He expressed frustration with his smart thermostat system, specifically when attempting to use the programming and remote-control features. Participant C was selected as a case study because her UMM presents a broad understanding of components in tandem with a high degree of confusion and frustration related to thermostat use. Her UMM includes supplementary heating components (e.g., plug-in heaters) and energy sources (e.g., gas line) but omits remote interaction. Participant C replaced her previous, manual thermostat with her current, smart thermostat, but did not engage with the new features, such as programming or remote interaction, prompting further analysis via case study. Participants A, B and C provide anecdotes and insight into misconceptions, highlighting areas of confusion and demonstrate a range of adoption (or rejection) of smart thermostat control opportunities.

3.3.1.2. **Detailed UMM case studies**

3.3.1.2.1. **Participant A – feedback UMM with switch behaviour, use of smart home system and preference for remote interaction**

Participant A is a male set designer between 25 and 35 years old and the sole occupant of a one-bedroom, bulk-metered loft apartment. He received his smart thermostat as a gift and installed it himself, preferring to use it in place of his previous manual thermostat. Participant A has always resided in Canada and has ten years of experience using baseboard or centralized forced-air gas furnace heating. He reported familiarity with manual (dial-like) thermostats as well as smart thermostat interfaces. Participant A has lived at his current accommodation for one year, using a forced-air gas furnace controlled by a smart thermostat and smart home system. Although he rents his apartment, participant A installed the smart thermostat system in place of the previous, manual interface. The detailed UMM diagram is shown in Figure 3.3 (key available in Figure 3.1).

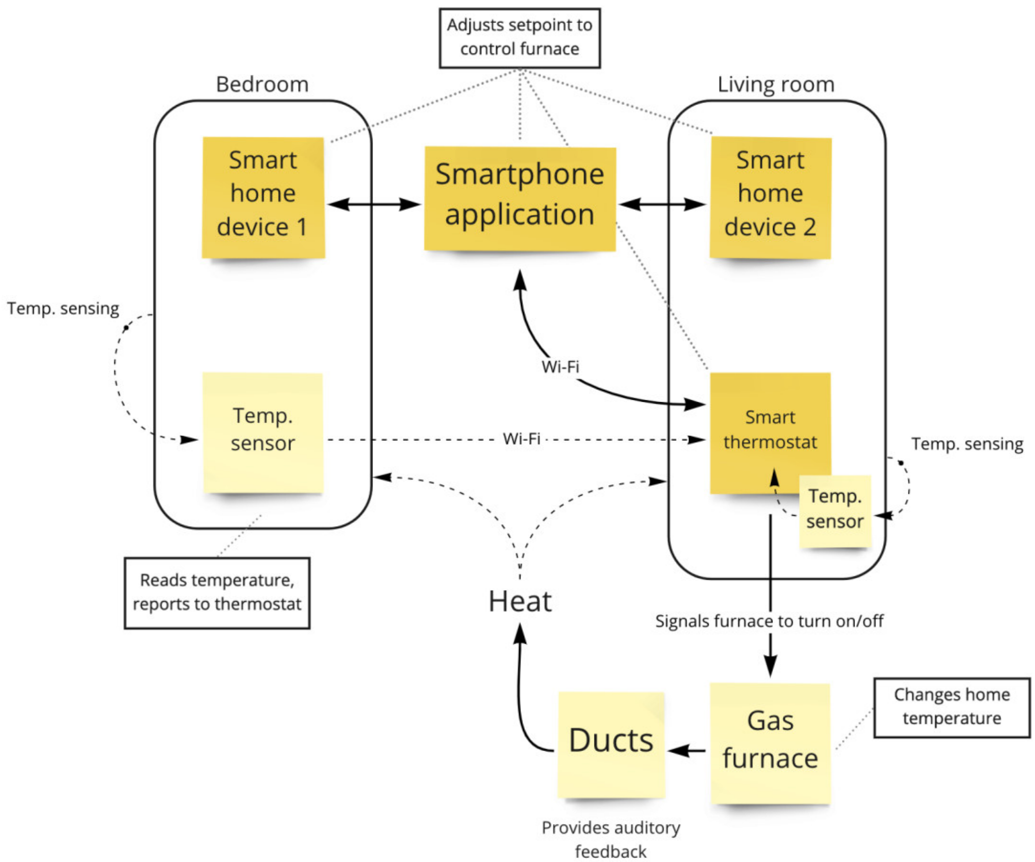


Figure 3.3: UMM description of home heating for participant A, redrawn for clarity.

The UMM (Figure 3.3) shows four opportunities for user interaction: at the thermostat interface, using the smartphone application and at smart home devices 1 and 2. From the points of interaction, a loop exists between where temperature is measured and heat is delivered (highlighted in Table 3.4). When asked to describe the system, participant A noted that the furnace is active until the desired setpoint in the occupied room(s) is achieved. If both rooms are occupied, the thermostat will regulate heat delivery such that the average temperature between the two zones (bedroom and living room) reflects the setpoint. Participant A's UMM for home heating reflects feedback theory, where the thermostat maintains a user-selected setpoint with a feedback loop between thermostat, furnace, and dwelling.

Compared to the sample, participant A employed many smart devices, opting to use a smart thermostat, two smart home devices and an additional temperature sensor. He noted that the devices themselves were easy to use, but that he had trouble installing the smart thermostat interface, due to the required hardware. Participant A had a strong preference for remote interaction – so much that he installed a bike-rack over the wall mounted thermostat interface, obstructing its use. The smart home devices (1 and 2) allowed him to communicate verbally with the thermostat (e.g., "Hey [smart home device], turn up the temperature."). Participant A noted that he preferred to make small temperature adjustments verbally but required his smartphone interface to perform more complex interaction, such as changing the thermostat schedule. The following quotes, transcribed from the interview video recording, express Participant A's sentiments related to thermostat control:

> ***RESEARCHER:*** What is the job of the thermostat?
>
> ***PARTICIPANT A***: To control the furnace.
>
> ***RESEARCHER***: And what is the job of the [smart home device]?
>
> ***PARTICIPANT A***: To control the furnace as well. Which is why I said, the thermostat is irrelevant. The [smart home device] and my phone do the same thing as the thermostat does.

Verbal communication with the thermostat system presents a new and unique method of control. When making tactile adjustments (i.e., on a phone application or wall-mounted interface), users can select specific setpoint temperatures, which further communicates that it is the thermostat, rather than the user, responsible for maintaining a comfortable temperature. While participant A's UMM depicts feedback theory, his verbal adjustments echo switch theory, increasing the setpoint temperature to activate heat delivery:

> ***PARTICIPANT A:*** I never actually ask the [smart home device] what temperature it is because I don't really care. I either know I'm cold or hot. I just tell it to change.

Participant A demonstrates a new instance of UMM and usage. He understands the cyclical relationship of the heating system but is constrained by the nature of verbal control. Participant A's approximate adjustment style (warmer/colder, rather than selecting a setpoint) reflects switch theory behaviour.

While approximate adjustment is likely encouraged by the method of verbal communication, it is important to consider other occupant motivations. Unlike the other respondents, participant A's home was bulk-metered. His primary goal during the heating season was to keep warm. He noted that his thermostat was "scheduled" – meaning that it was programmed to maintain a constant temperature (i.e., not employing overnight setbacks). Whether approximate adjustment is a symptom of the control mechanism or a lack of engagement/energy consciousness due to bulk-metering remains somewhat undetermined.

New smart thermostat systems offer a high degree of feedback, often promoting energy conservation. Participant A receives feedback from his thermostat via email and text message, detailing his home's energy consumption and how his usage compares to other users. However, the user noted that he did not read these emails and was not inclined to employ energy-saving measures as it had no financial advantage. Previous work has identified a causal relationship between bulk-metered dwellings and energy-intensive behaviour [85]–[87]. Participant A's heating strategy and UMM suggests that users may not attempt to save energy in bulk-metered units, regardless of the technology, interface or feedback employed.

Participant A's home presents a novel scenario where the occupant has employed a high degree of smart home technology to operate his heating system. He strongly preferred remote control through verbal communication or smartphone interaction. Verbal communication with the thermostat or smart home device poses unique constraints on interaction. While participant A was aware that temperature setbacks save energy and had access to feedback detailing his usage, he was not motivated to engage with these behaviours, possibly due to a lack of financial incentive.

3.3.1.2.2. Participant B – feedback UMM with user uncertainty, misconceptions, and low engagement with smart features

Participant B is a male sales associate between 25 and 35 years old and one of three roommates renting a two-bedroom, sub-metered apartment. The smart thermostat system was not selected by the user(s) and was installed before their occupancy. One of the roommates (not participant B) had remote thermostat adjustment enabled on their smartphone. All three occupants interact with the thermostat regularly. Participant B has always resided in Canada and has approximately one year of experience using a centralized forced-air gas furnace at his current dwelling. The furnace is shared between four apartments. His experience with heating systems stems solely from his current accommodation, which utilizes a smart thermostat. The detailed UMM diagram is shown in Figure 3.4 (key available in Figure 3.1).

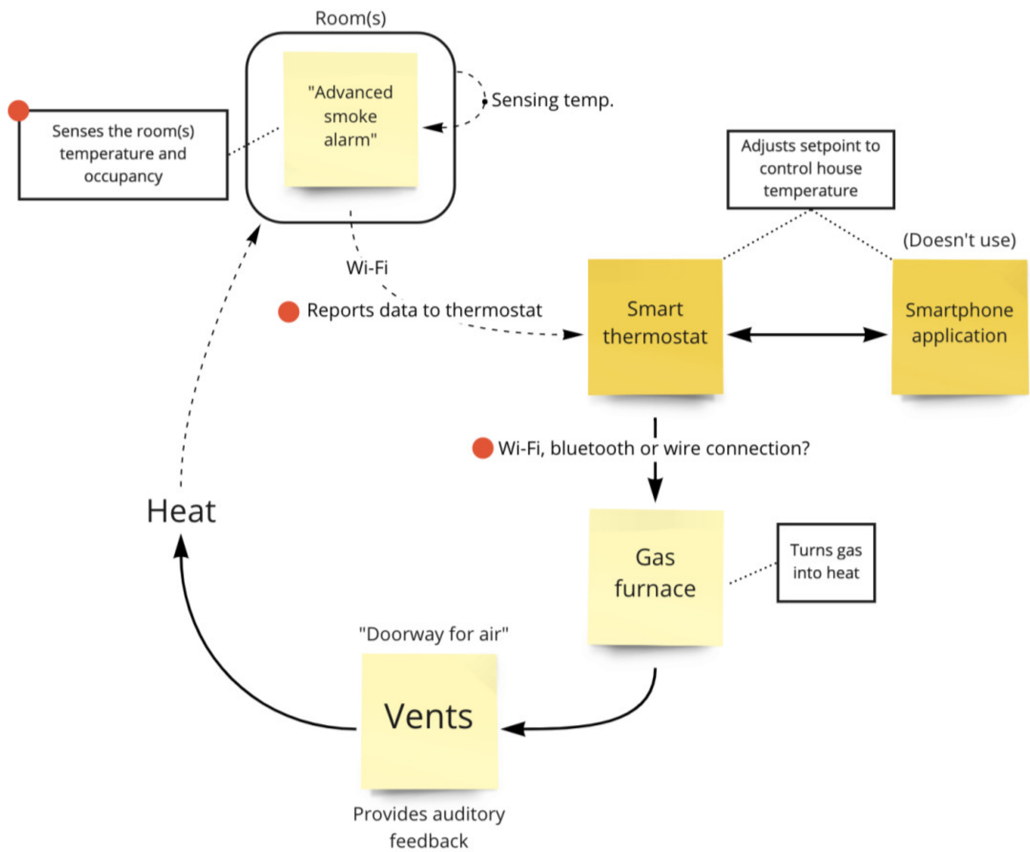


Figure 3.4: UMM description of home heating for participant B, redrawn for clarity.

Participant B's UMM differs greatly from participant A's in terms of complexity and confidence. Two opportunities for user control are presented: at the wall-mounted thermostat and smartphone application. Participant B's UMM presents a cyclical relationship similar to participant A's, reflecting feedback theory. Here, the temperature is sensed in each room at the "advanced smoke detector," then communicated to the thermostat to activate the furnace to until the setpoint temperature is achieved (highlighted in Table 3.4).

When constructing the UMM, participant B expressed confusion and uncertainty, primarily related to where the temperature was sensed and how the thermostat communicated with the furnace. He believed that the "advanced smoke detectors," located on the ceiling of each room measured and reported household characteristics including levels of carbon monoxide or smoke, occupancy, and temperature (i.e., communicates occupancy and temperature data to thermostat). He was confident the thermostat communicated with the furnace but was unsure if the connection was physical (wire) or wireless (Wi-Fi, Bluetooth). The following dialogue illustrates participant B's response and certainty:

> ***RESEARCHER****:* What does the furnace do?
>
> ***PARTICIPANT B***: It turns gas into heat.
>
> ***RESEARCHER***: And how does the furnace know when to turn gas into heat?
>
> ***PARTICIPANT B***: There must be something. Somehow the thermostat is connected to something on the furnace. I don't know.
>
> ***RESEARCHER***: Okay. You mentioned earlier that it was Wi-Fi or Bluetooth?
>
> ***PARTICIPANT B***: Yeah, maybe. I don't even know because there are no wires connected to the [smart thermostat], it's just like, on the wall. So, it must be like, some sort of internet connection… I have no idea.

Participant B was aware that his thermostat system could be controlled remotely and had attempted to set up his smartphone to do so. He reflected on the smartphone set-up procedure, noting that he found it confusing and opted for traditional control when it could not be accomplished quickly. When asked if he might reattempt enabling remote control, he dismissed the opportunity:

PARTICIPANT B: I don't need more things connected to my phone.

Participant B noted that there were too many features included on the thermostat, making the usage experience confusing. He reported feeling no more familiar with his heating system since using the smart thermostat.

Participant B's goal for the heating season was to keep warm. Unlike participant A, he opted to adjust his smart thermostat was manually. In the shared home, two occupants performed routine adjustments (e.g., overnight setbacks) – the other interacted irregularly, based on (dis)comfort. Participant B believed that thermostat interactions, such as adjusting the setpoint, increased his energy consumption. He noted that because the smart thermostat interface was attractive and easy to use, it received more interaction and thus, consumed more energy.

Participant B found manual setpoint adjustments intuitive. However, the scheduling interaction was challenging. When asked if there was anything his thermostat could do better to improve his experience, he remarked that the programming interface was difficult to use, and he would prefer a more intuitive method of inputting his desired schedule. He did not provide specific suggestions for improving usability.

In summary, although participant B, along with his roommates, understood that he could program his smart thermostat to perform routine setbacks, he chose to adjust it manually, daily. His manual overnight setbacks were performed for comfort, rather than energy savings, and he believed this behaviour was more energy-intensive than employing a consistent setpoint. He demonstrated an aversion to remote thermostat control, after an unsuccessful attempt to set it up on his smartphone. This pattern is echoed in his attempt to input a schedule, where he found the procedure confusing and from there, opted for manual adjustment. Participant B's UMM has some areas of confusion and misconception, primarily related to where and how the temperature is measured and the communication between the sensor, thermostat, and furnace. However, his UMM depicts a clear representation of feedback theory.

3.3.1.2.3. **Participant C – feedback UMM, limited engagement with smart features**

Participant C is a retired woman, between 56 and 65 years old, sharing ownership and living with two other occupants in a sub-metered, single-detached home. The occupants purchased the smart thermostat system as it was more compatible with their new forced-air gas furnace. All household members had access to the thermostat and avoided making adjustments. Participant C has always resided in Canada and has lived twenty-two years at her current dwelling, using a forced-air gas furnace. She noted she was most comfortable with a dial-like thermostat interface and had only used her current smart thermostat interface for one year. She had 43 years of experience with home heating systems, including gas and wood fireplaces, forced-air gas furnace, oil, and electric radiator heating. Her detailed UMM diagram is available in Figure 3.5 (key available in Figure 3.1).

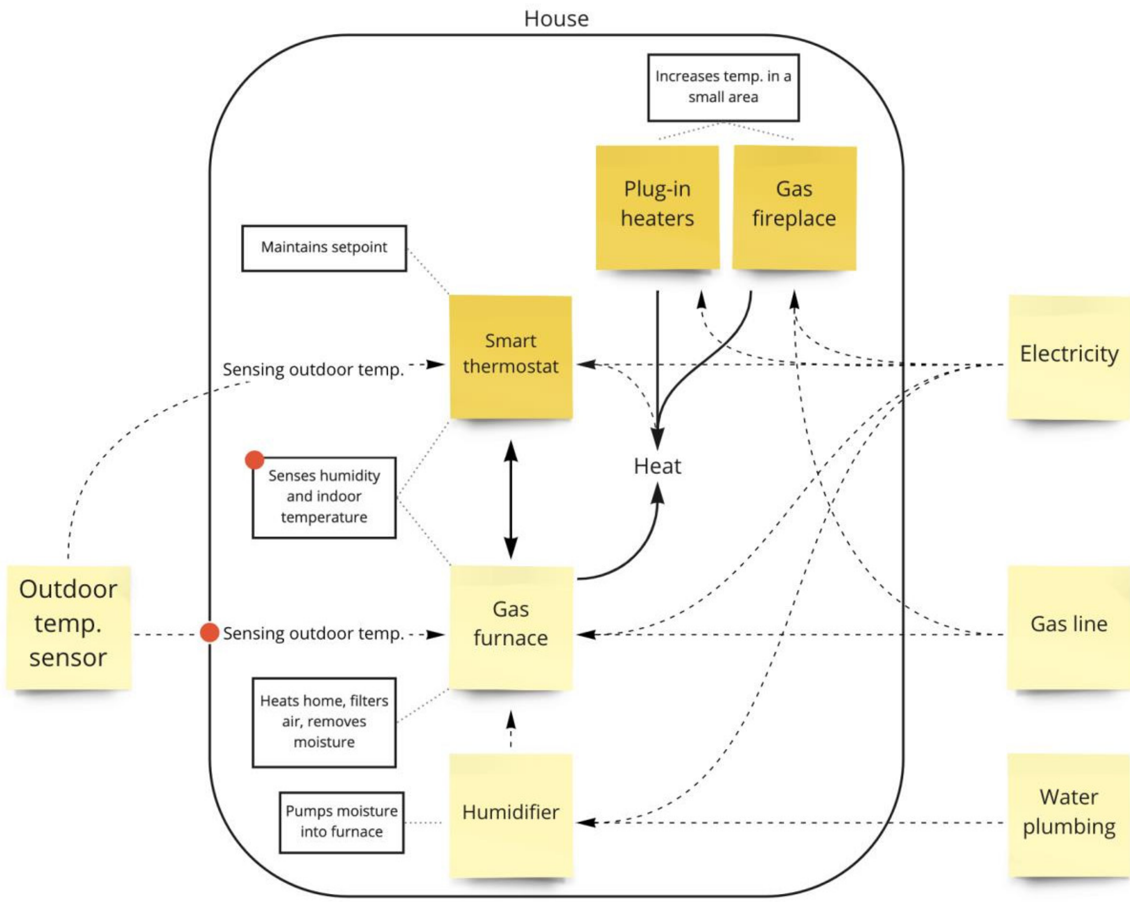


Figure 3.5: UMM description of home heating for participant C, redrawn for clarity.

Like participants A and B, participant C appears to adhere to feedback theory (highlighted in Table 3.4). The UMM (Figure 3.5) shows a loop between where the heat is measured (smart thermostat), produced (gas furnace) and delivered (heat). Participant C included a large scope of components in her UMM, with supplementary heating devices (plug-in heaters, gas fireplace) and external sources (gas, electricity, and plumbing). Uncertainty was expressed for where the temperatures (indoor and outdoor) and humidity are measured (i.e., if temperatures were measured at the thermostat or gas furnace).

Although participant C knew that she could operate her thermostat from her phone, she found it undesirable, and did not include it in her UMM. She reported being regularly frustrated when attempting to use her smart thermostat interface, which related to her low motivation for pursing remote control. She didn't consider remote interaction a worthwhile feature, noting that it would only be useful to monitor the status of the heating system (e.g., checking the heat indoor temperature to ensure the pipes don't freeze).

Participant C's priority during the heating season was to keep warm. She was very conscious of the humidity and air mixing across the three floors in her home and adjusted her system with both priorities in mind. Because there was a large temperature gradient between the levels in her home, she kept the furnace fan on high to promote air mixing. However, the air movement often made her uncomfortably cool, and she reported regularly placing cushions over floor registers.

Both participants A and C operated their thermostat without setbacks. However, unlike participant A, participant C's home was sub-metered. She noted that there was no need to program a schedule because the house was always occupied. Like participant B, she believed a consistent setpoint produced the highest energy savings. However, unlike participant B (who thought increasing interaction increased energy consumption), her ideology stemmed from previous experience and personal patterns of adjustment. Participant C used to employ manual overnight setbacks but often found herself feeling uncomfortably cold in the morning. Because of this, she would reactively increase the setpoint, overshooting her desired setpoint temperature and then forget about it, instead of turning it back down once comfortable. She noted this as wasteful and has since opted to have a consistent setpoint. Participant C's UMM may represent a more traditional instance of feedback theory, reflecting Kempton's findings for manual thermostats [29] where users knew the thermostat regulated the temperature and were less likely to employ overnight setbacks. This is especially applicable as Participant C had only used manual thermostats before adopting her smart thermostat.

Since choosing to maintain a consistent setpoint, the only reason Participant C adjusts the thermostat is for humidity control. For example, if the basement feels damp, she increases the setpoint to "dry the air." When asked if using a smart thermostat influenced her energy consumption, participant C responded:

> ***PARTICIPANT C:*** I think because it controls the temperature a bit better, I am less likely to adjust the temperature. Because the dial-style thermostat shows a range, you're not exactly sure where you're at within that range so I think there's a little bit more fluctuation there (using a manual thermostat) than the digital style smart thermostat.
>
> ***RESEARCHER:*** So, you would need to interact with it more if it was manual?
>
> ***PARTICIPANT C:*** We did interact with it more. Yeah.
>
> ***RESEARCHER:*** Would you ever want to receive feedback about your energy usage?
>
> ***PARTICIPANT C:*** That's not relevant to me because I can afford the bills at this point and time. So that wouldn't necessarily be relevant and because I don't change the dial, the temperature doesn't fluctuate. And I know my main purpose is to have it at a particular temperature. I'm not as concerned about energy savings or energy consumption. That's sort of beyond my control so the information would be irrelevant to me.

What's most notable about participant C is that although employing a smart thermostat system, she still prefers to operate it as if it were a manual thermostat, without programming. She noted repeatedly that the smart thermostat interface was challenging to use, which may have contributed to her style of operation. She was not interested in the financial savings accrued through temperature setbacks. She demonstrated a lack of engagement with energy saving features, such as the use of a thermostat schedule or energy-related feedback. Her frustrating experiences attempting to navigate the device might have deterred her from engaging further. Besides the small areas of uncertainty (i.e., if the temperature is sensed at the thermostat and/or the furnace), she has a relatively accurate understanding of the heating system components and recognizes that the thermostat regulates the temperature, adhering to the feedback model.

3.3.2. RQ2: What attitudes are held towards smart thermostat devices & how does the adoption of smart thermostats relate to energy consciousness/consumption?

In relation to RQ2, participants were asked about their attitudes, preference, and usage. Five different smart thermostat brands are represented in the participant sample. Users noted how and why they chose their smart thermostat, and the positive and/or negative aspects of their system.

For some, adopting a smart thermostat occurred at a natural threshold – the device was more compatible with their newly installed heating system or was discounted and promoted by their energy supplier. Others acquired their devices through suggestion from friends or family. Remote adjustment was a major attraction. Participants noted that some activities, such as programming a schedule, favoured the remote touchscreen interface (e.g., smartphone application). Additional features, including notifications related to usage and increased environmental data, motivated the adoption of smart thermostats in place of traditional, programmable devices.

3.3.2.1. **Attitudes and preference**

Smart thermostats present new features, many with the goal of promoting energy-savings. However, their influence remains somewhat undetermined. All participants in this study received feedback related to energy use (e.g., heating system run times) and could compare their profile with other smart thermostat users. However, most (9/10) users noted this did not influence their consumption but was "good to know" or "nice to see." While these features received positive reception, their intended impact may not be fully realized.

To further disseminate user attitudes and preferences, Table 3.5 provides a comprehensive list of the reported advantages and disadvantages related to smart thermostat operation.

Table 3.5: Participant reported advantages/disadvantages for smart thermostat usage

<table>
<tr><th>Advantages</th><th>Disadvantages</th></tr>
<tr><td>Interface</td><td rowspan="6">- Steep learning curve
- Relies on Wi-Fi connection
- Battery-dependent temperature sensors (users forget to change)
- Breadth of features is intimidating
- Requires user to remember and recall procedures (e.g., how to schedule)
- Difficult to physically install</td></tr>
<tr><td>- Displays weather forecast
- Displays indoor and outdoor temperatures
- Displays humidity
- Ability to select which sensor dictates the setpoint
- Aesthetically pleasing
- Dial-control adjustment
- Motion-activated display
- Provides schedule flexibility beyond Energy Star's recommended periods (wake, leave, return, sleep) [88]</td></tr>
<tr><td>Energy-related</td></tr>
<tr><td>- Ability to compare energy consumption with other users
- Provides information about energy use and potential savings</td></tr>
<tr><td>Other</td></tr>
<tr><td>- Ability to employ multiple temperature sensors
- Ability to interact remotely (e.g., from smartphone)
- Ability to interact verbally
- Compatibility with smart home system
- Provides reminders (e.g., change furnace filter)</td></tr>
</table>

The majority (7/10) of participants expressed a desire to further automate temperature setbacks based on occupancy. Given a choice between motion detection or geo-fencing, they opted for the latter. One respondent wanted the ability to set specific setpoint preferences to each occupant (e.g., temperatures for adults vs. children) and have the heating system automatically adjust based on who occupied the home.

3.3.2.2. Usability and usage

Revell and Stanton [80] previously identified common usability issues for programmable thermostat interfaces and argued that smart thermostat systems introduce more features and complexity, rather than designing for intuitive or usable interfaces. The sample of smart thermostats was reviewed to determine if the usability issues persisted. Some issues appear to remain, detailed in Table 3.6.

Table 3.6: Usability issues for programmable thermostat interfaces (identified by Revell and Stanton [80]) and observation for smart thermostat sample.

Issue	Observation
Programmable thermostat interfaces do not effectively communicate the feedback function of the device.	Previous work has identified various shared theories among manual and programmable thermostat users (Table 3.2). This study, conducted with smart thermostat users, produced ten UMMs, each adhering to feedback theory. Thus, the participants in this study demonstrate that smart thermostats have effectively communicated the feedback function of the smart thermostat device.
Programmable thermostat interfaces do not communicate where the temperature is sensed.	Thermostat users have demonstrated confusion or uncertainty related to where the temperature is measured in the home. This issue, first identified in manual and programmable thermostat users, was observed within the evaluated sample of smart thermostat users. Two of three case studies (participants B and C) demonstrated uncertainty in their UMMs related to where the temperature was measured by their respective heating systems.

Dial control triggers a valve schema [81], resulting in a UMM that reflects valve theory.	Among the smart thermostats evaluated, 3/10 employed dial control at the interface. Users that operated a dial-controlled interface did not adhere to valve theory. All UMMs reflect a feedback theory understanding of the heating system.
The programming procedure and interface is confusing and discourages engagement	During interviews, participants noted that their thermostat interfaces were confusing, and they had been discouraged when attempting to program their device. Some noted a steep learning curve required to effectively use programming functions, others had abandoned the activity altogether.
The program schedule is hidden.	Previous work noted that programmable thermostat users cannot easily view their thermostat program schedules - users were often required to cycle through individual time periods to see setpoint changes. Smart thermostats appear to remedy this issue, specifically with the use of smartphone applications. During the interview, participants viewed their weekly setpoint changes without difficulty. One participant shared a screenshot from their smartphone, showing the weekly heating schedule.
The link between thermostat and furnace operation is not clear; users assume that scheduled periods equate to active heating periods	When asked what activates the heating system, participants noted that heat delivery is dependent on the margin between setpoint and current temperature(s). Within the sample, there was no indication that users believed the heating system was constantly producing heat (reflecting valve theory) and instead understood it to be regulated by the thermostat (reflecting feedback theory).

While some usability issues remain, it appears that smart thermostat users are still able to form effective UMMs that resemble institutional theory. Therefore, these usability issues may not contribute to misunderstanding in users, relative to how their HVAC system functions and relates to the thermostat. However, usability issues still influence usage. Multiple users had trouble when attempting to program their thermostat, resulting in reduced engagement with energy saving features.

3.3.2.3. **Thermostat schedules**

While some case study participants (A and C) did not employ overnight setbacks, this was not reflective of the whole sample. During the interview, participants relayed their thermostat schedules verbally, often while viewing it from their phone. From this, a profile for weekly setpoints/durations was produced, demonstrating the use of temperature setbacks for unoccupied and sleeping periods. Figure 3.6 and Figure 3.7 show individual and average setpoint schedules, respectively.

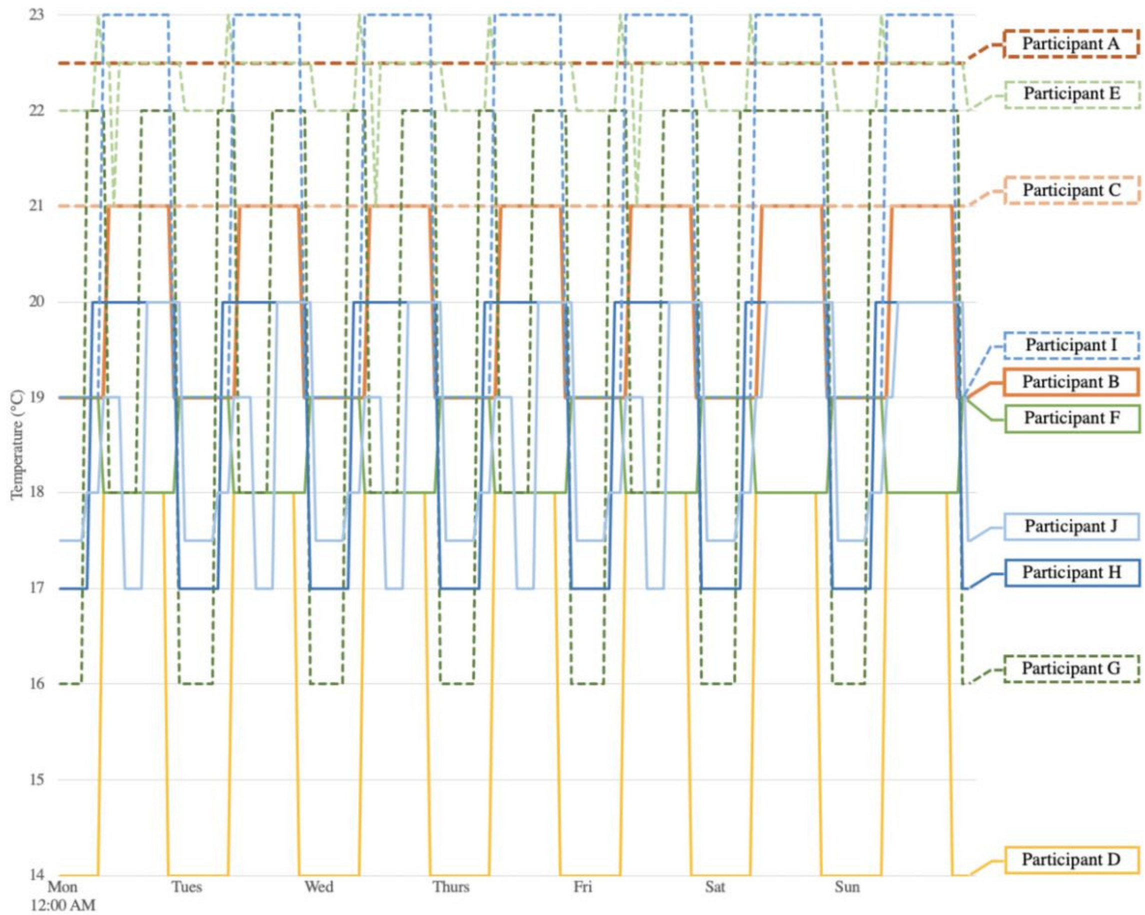


Figure 3.6: Self-reported setpoint schedule for ten smart thermostat users.

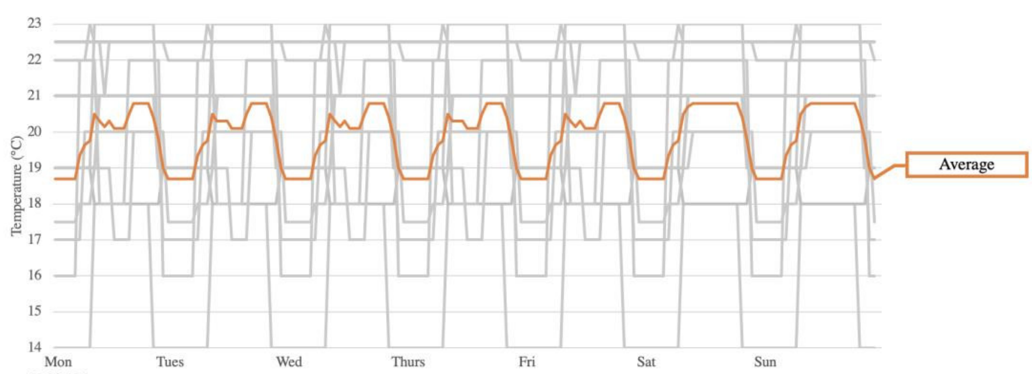


Figure 3.7: Self-reported setpoint schedule for ten smart thermostat users with averaged trendline.

It is significant that most participants (8/10) employ temperature setbacks. While programmable thermostats have been frequently cited as too complicated for use [3], [12]–[17], [32]–[34], the smart thermostat sample appears to support efficient behaviour. Participants noted that more complex tasks, such as programming a schedule, were easier to achieve using the smartphone application. The ease of programming a schedule on the smartphone interface may have contributed to the widespread use of temperature setback. The two participants who did not use temperature setbacks (participant A and C) were included in the case study analysis and noted their respective reasons for not employing a schedule.

Participant A lived in a bulk-metered apartment and indicated that his thermostat was 'scheduled', meaning that it was scheduled to hold the same temperature. He noted that the lack of financial incentive influenced his usage behaviour, but that the thermostat was easy to operate. Participant C lived in a sub-metered apartment and expressed difficulty operating her thermostat. She found the interface confusing and did not attempt to control it remotely. She also noted that a lack of financial incentive influenced her thermostat usage. From the case studies, Participant B's thermostat schedule reflects the lowest energy consumption (he was the only one to employ consistent setbacks). Participant A (bulk-metered) and C (sub-metered) consumed 12% and 5% more energy than participant B, respectively. For all case study participants, temperature adjustments were motivated by comfort, rather than energy or financial incentives. Thus, while smart thermostats offer significantly more methods of feedback and usage, results suggest that occupants still require further motivation to alter their behaviour and engage with energy saving features.

3.3.2.4. **Relative heating degree days**

The participants' self-reported setpoint schedules were assessed against Ottawa's typical weather (typical meteorological year) to estimate the relative heating degree days (HDD) associated with each schedule (Table 3.7). Based on Touchie and Siegel's research [89], which found a typical balance temperature of 14°C, we assumed that the homes would start requiring heat if the outdoor temperature is 7°C less than the current setpoint. A sensitivity analysis revealed that the results are not very sensitive to this assumption.

Table 3.7: Heating degree days for each participant's thermostat schedule

Participant	**Average setpoint temperature (°C)**		**HDDs** (°C-day)	**Relative HDDs,** compared to 21°C case (%)
	Daytime (8:00 AM – 8:00 PM)	**Nighttime** (8:00 PM – 8:00 AM)		
A	23		3474	112
B	21	19	2822	91
C	21		3088	100
D	18	14	1952	63
E	22		3423	111
F	18	19	2498	81
G	21	18	2740	89
H	20	18	2592	84
I	23	19	3092	100
J	19	18	2522	82
		Average	2793	90
		21 °C	3088	100

3.3.3. RQ3: How effective is the use of online software & remote strategies in the investigation of UMMs?

The use of online software to capture UMMs largely met the research objectives, with some considerations and limitations. While Miro enabled remote diagramming, the method was influenced by participants' lack of experience using the software. Often, the researcher would support the creation of diagrams, aiding the participant in the generation of post-it notes, the mapping, and the placement of text. Conscious that this could produce bias or inaccurately represent the internal constructs of thermostat users, extensive checks for clarification and confirmation were taken at each step to ensure the diagram accurately represented the participants' understanding and intentions. After completing the exercise, each heating system component and link was confirmed individually for correct interpretation. Additionally, participants were given the opportunity to indicate a degree of confidence for each element of the heating system diagram; red markers denoted areas of low confidence. While this proved effective, a UMM constructed by the participant themselves may present subtleties lost in the analyst supported model. The analyst supported UMM is inherently more susceptible to bias, represented in the placement of post-it notes and annotation.

The software was perhaps even more effective for capturing the UMMs compared to previous methods [2], [29], [31], [70], where UMM diagrams were hand-written and had to be re-interpreted when digitized by the researcher. The use of Miro allowed the participant to validate the exact digital representation of their finalized UMM diagram and reduced opportunities for misinterpretation or misrepresentation. Furthermore, the video recordings, opposed to voice recordings [28], allowed the analyst to review the interview with additional context (e.g., expressions, body language, etc.).

The remote inquiry allowed respondents to participate from their home environment, which may have provided useful context as they attempted to describe their home heating system. Participants frequently opted to have the analyst manipulate the Miro interface while they provided feedback. The same dynamic was present in previous work [28]. If adopting the same remote methodology in the future, it would add value if the analyst shared their screen during the call while the UMM was constructed, especially considering that the analyst's and participants' cursors can be viewed simultaneously on the Miro interface. This would allow dialogue and gestures to be captured parallel to UMM development, providing rich insight and context for future review.

3.3.4. **Limitations**

While extensive measures were taken to reduce bias and accurately represent participants, the research methodology has inevitable limitations. Before the UMM diagramming exercise, participants were interviewed about their behaviour and responded to different usage scenarios. The discussion before UMM construction aided in identification of heating system components and user-initiated terminology (e.g., participant B mentions the advanced smoke detector in response to an interview question; analyst makes a post-it note labelled "advanced smoke detector" to later be mapped within the heating system). However, interview questions asked participants for specific details related to their thermostat, increasing the likelihood of the term "thermostat" being included in the UMM. This contrasts a previous UMM investigation [28], where the omission of the word "thermostat" represented a key finding.

A previous UMM investigation, conducted by Revell and Stanton [28], considered six participants with the same (or relatively similar) home layouts, thermostat interfaces, and experience with home heating systems. The same participant controls were not available in this investigation. While living in a similar geographic area and operating the same type of heating system (forced-air gas furnace), participants differed in experience, dwelling type, and thermostat interface(s). Because the respondents had less in common, their UMMs cannot be directly compared. The configuration of this study produces a broader, more generalized sample of UMMs, attitudes, preferences, and usage for smart thermostat systems – without definitive statements or conclusions.

Previous work has demonstrated that small participant samples can be effective in identifying usability issues [52]. The sample used in this study is larger than that previously employed [28] to evaluate UMMs for programmable thermostat users. Furthermore, the nature of qualitative research and UMM investigation favours detailed analysis for a small sample of participants. However, while this work holds value, providing anecdotes and insight into user experience, as well as areas of repeated difficulty or misconception, a sample of ten participants has limited application for quantitative results.

3.4. Closing remarks

An investigation was performed to identify UMMs of home heating systems for smart thermostat users using a remote-interview methodology. Data was collected from ten participants in Canadian homes during the heating season using semi-structured interviews and diagramming activities, reflecting previous literature [28], [29]. Five different smart thermostat brands/devices were included in the study.

All UMMs for the sample of participants appeared adhere to feedback theory. Foundations of cognitive frameworks were consistent across the sample; however, the level of detail, certainty, and scope of the UMM diagrams varied between participants. Results indicate that thermostat interface designers may be able to assume or produce specific UMMs by providing key information and use this foundation to inform energy-saving strategies.

Three case studies were analysed. Participant A lived in a bulk-metered apartment and employed a high degree of smart home technology (relative to other participants), including the use of two smart home devices, an additional temperature sensor and a preference for remote adjustment from his smartphone or verbally. He had the highest energy consumption, based on his thermostat settings. Participant B lived in a sub-metered apartment. His thermostat schedule consumed the least amount of energy. He was the only case study participant to employ overnight setbacks and opted to control his thermostat manually after difficulty using the scheduling interface. Participant C lived in a sub-metered, single-detached house and operated her smart thermostat manually, citing confusion and frustration when using the interface beyond setpoint adjustment.

While two of three case studies highlighted participants that struggled to use their thermostat interface, overall, respondents expressed positive attitudes. They noted significantly more advantages than disadvantages for smart thermostat usage. However, some usability issues seen in programmable thermostat interfaces persisted in the smart thermostat sample; participants were confused where the temperature was sensed and experienced frustration when attempting to program a schedule, resulting in user disengagement.

The use of online software, including video-call and online diagramming, was effective in capturing and verifying UMMs. Future work should consider the use of virtual exercises for evaluating participant attitudes and cognitive frameworks.

Chapter 4: Conclusions

This research employed participant interviews to evaluate smart thermostat interfaces on the basis of usability and user comprehension. Quantitative and qualitative methods were employed to produce an in-depth analysis and comparison of smart thermostats, addressing a gap in the literature. The summary of conclusions and contributes for each chapter is available in Sections 4.1 and 4.2.

4.1. Summary & contribution

4.1.1. A quantitative usability evaluation for manual, programmable, and smart thermostat interfaces

In this chapter, the usability of thermostat interfaces was evaluated using a mixed-methods approach. A sample of 51 manual, programmable, and smart thermostat users in Ottawa, Ontario were interviewed remotely. Each participant attempted six usability tasks on their respective thermostat interface. Video-recordings were used to determine usability metrics and produce a framework for evaluating a variety thermostat interfaces in terms of time and interactions taken per task, success/failure of the task, and participant feedback. The key findings of this investigation are listed:

- Significantly more participants of the smart thermostat sample (78%) successfully completed the most complex task (setting a schedule) compared to programmable thermostat users (52%).
- Smart thermostat interface users performed better than programmable thermostat users in usability trials, reflected by participant-assigned usability ratings.
- When asked how participants knew their heating system was working, auditory feedback was the most reported (indicated by 82% of participants), and visual feedback, the least (28%).

- Participants noted a strong desire for increased visual feedback from their thermostat interface, related to projected energy consumption and cost.
- Half (51%) of the participants indicated they either desired or enjoyed thermostat control through a smartphone application.

4.1.2. A qualitative evaluation of smart thermostat users

In this chapter, ten smart thermostat users were evaluated using qualitative methods to produce ten unique user mental model (UMM) diagrams which represent how the user imagines their home heating system to operate. UMMs detailed the interactions of various heating system components and allowed participants' cognitive frameworks to be analysed and compared to previously established shared theories, with implications for effective and efficient thermostat operation. From the UMM inquiry, key findings were derived:

- All participant UMMs adhered to feedback theory and the majority (8/10) employed temperature setbacks, suggesting that smart thermostat interfaces communicate accurate representations of the home heating system and enable efficient behaviour.
- Some usability issues identified in programmable thermostat interfaces continued to persist in the smart thermostat sample, negatively influencing user engagement and the related energy consumption.
- Methods of thermostat control lend themselves to different tasks (e.g., remote thermostat interfaces, such as a smartphone application, were favored when programming a schedule).
- Verbal interaction with the smart thermostat introduced new constraints for adjustment, with implications for setpoint selection and behaviour. Participants were more likely to make

subjective temperature adjustments (e.g., warmer, cooler) instead of selecting a specific setpoint.

- Energy use feedback was desired and well received among smart thermostat users, however, most (9/10) noted it had no impact on their behaviour.

4.2. **Research contributions**

4.2.1. **A quantitative usability evaluation for manual, programmable, and smart thermostat interfaces**

This chapter contributes to the body of literature focused on thermostat operation, usability, and development. Unlike previous work, this research included a large participant sample with strong representation of smart thermostat users. It provided a comprehensive comparison of manual, programmable, and smart thermostats, and contextualized findings with previous usability studies. Usability was evaluated based on time and interactions required to complete usability tasks, success/failure in task attempts, and feedback from participants. From usability testing, programmable thermostat users appeared to require significantly more time and interactions to complete tasks - especially energy-saving operations, such as programming a schedule. This finding was further supported by participant feedback, which assigned programmable devices the lowest usability rating, compared to smart and manual thermostats. Given that manual thermostats have limited functionality and do not support energy-saving measures, such as scheduled temperature setbacks, smart thermostats appear to be the most desirable interface in terms of usability and functionality. Results from usability testing were used to generate a time and success (TS) metric, which varies based on the tasks considered (e.g., tasks one to six). To allow comparison, the TS metric was calculated multiple times such that it aligned with previous usability tests and could be directly related to findings in literature. Independent of the tasks considered, smart thermostats consistently demonstrated the highest TS metric, indicating a comparatively high degree of usability. Users indicated they desire further feedback related to their usage (e.g., energy consumption) and inputs (e.g., confirmation prompts on the thermostat interface). Users also demonstrated a strong desire for or willingness to adjust their thermostats remotely (e.g., from a smartphone application). This research employed a novel methodology, where all usability testing was conducted remotely, and serves as an example of adapting traditional research methods to increase accessibility and efficiency.

4.2.2. **A qualitative evaluation of smart thermostat users**

This chapter contributes to the body of literature focused on human-computer interaction and interactive design, specifically related to residential systems, as well as qualitative research methods for design. The chapter provided the first user mental model (UMM) analysis performed in relation to smart thermostat systems and followed a similar methodology previously employed with manual and programmable thermostat users, such that results can be broadly compared. Notably, all smart thermostat users appeared to employ feedback theory when describing their heating system and in their UMM. Feedback theory is recognized as a functional theory, in that it provides thermostat users with an accurate depiction of their heating system, enabling effective and efficient usage. The case studies suggested that some usability issues persist in the smart thermostat sample, however, many respondents expressed positive attitudes towards their smart thermostat device and its functionality. Moreover, participants demonstrated a willingness to employ further energy saving measures, such as automated temperature setbacks through geo-fencing. The use of online software to conduct interviews and record UMMs represents a novel research method that was highly effective in capturing, evaluating, and verifying feedback from participants.

4.3. **Recommendations for future work**

This research attempts to address gaps in the literature related to smart thermostat systems, including their usability and usage compared to other devices, as well as user attitudes, preferences and understanding. From the investigation, areas warranting further research have been identified, and are presented as follows:

- Throughout the usability study (Chapter 2), participants, especially those without smart thermostats, indicated that they desired more feedback from their thermostat system, related to their energy consumption for specified periods (e.g., monthly). In the UMM study (Chapter 3), participants received extensive usage feedback from their smart thermostat systems; however, the majority (9/10) noted that this information did not influence their behaviour or adjustments. Further research should investigation the effectiveness of feedback in motivating efficient behaviour from users, in relation to thermostat operation.
- This research has demonstrated that key usability issues exist that limit smart thermostat users from engaging with energy saving features, including temperature setbacks and the incorporation of energy-use feedback. Thermostat design may be improved by considering usability issues, including interface complexity and communication. Research showed that when successful, the use of remote control through smartphone applications provides and intuitive interface for complex operations, such as programming. However, users are still frequently confused or frustrated with the thermostat interface, reducing engagement. Future

interface designs should provide more contextual information, helping users to map where and how to engage with energy saving features. Furthermore, different avenues for delivering energy consumption and usage feedback should be explored, as current models appear to receive low engagement. Future work should follow a similar approach as Revell and Stanton [79], using the available literature to prototype an ideal smart thermostat interface, informed by UMMs, usability studies and human-computer interface research (e.g., Norman's design recommendations for everyday devices [70]).

- The development of usable smart thermostats appears likely to promote the engagement of users and yield energy savings. Specific interest should be given to opportunities to suggest or nudge behaviour and investigate the use of automated setbacks (e.g., geo-fencing).
- The growing range of interaction (smartphone application, verbal communication, etc.) is likely to have implications on usability, frequency of interaction and related energy consumption. Thus, traditional and emerging interaction opportunities should be evaluated and compared for their influence on user control.
- The thermostat usability study observed traditional, wall-mounted thermostats. The methodology limited the types of interfaces that could be observed, as participants performed interviews on video-call using their smart phones. Thus, users were unable to video-record their interactions with alternative interfaces, such as a smartphone application. As devices evolve, future work should consider including additional methods of thermostat control, such as remote (e.g., web or phone applications) or voice-control interfaces (e.g., Amazon Alexa voice control).

- UMM investigation requires a significant amount of time conducting interviews, constructing diagrams, confirming interpretation, analysing video recordings and participant response, etc. To produce both qualitative and quantitative results, future work could consider a two-part interview: the first conducted with a large sample of participants, producing quantifiable results. From this this sample, a smaller group of users could be selected to participate in the second leg of the interview, focused on UMM investigation and other qualitative inquiry.
- The link between UMMs and household energy consumption represents a gap in the literature. Future work should consider relating specific energy outcomes with user conceptualization (e.g., shared theories) to better understand what is required in a UMM to elicit efficient usage.

The purpose of this research was to investigate smart thermostats and improve the body of knowledge, with a focus on both quantitative (usability testing) and qualitative (user mental model investigation) data. The results are limited to the devices studied and will need to be repeated as new methods of control and features emerge.

Appendices

Appendix A: Additional Tables

Table A.1: Thermostat sample

Type	Brand	Model	Count
Manual	Bryant	TSTATBBNQ001	1
	Centurion	Unknown	1
	DC Dimplex	Unknown	1
	Honeywell	RLV3150A1004	1
		T87N1000	1
		TH1110DV1009 PRO 1000 Non-Programmable Vertical	1
		T6575C1001	1
	Stelpro	R32W2	1
	White-Rodgers	1F37-407	1
		1A65	1
		1F78-144	1
Programmable	Honeywell	Magic Stat 97-4730	4
		RTH221B1047	3
		PRO 2000 Vertical	3
	White-Rodgers	1F86-0471	2
		1F80-361	2
	Bryant	SYSTXBBECC01-B	1
		T6-PRH01-B	1
	Carrier	SYSTXCCUID01-V Infinity control	1
	Emerson	1F78-151	1
		1F95-1277	1
	Honeywell	WIFI 7-day programmable touchscreen thermostat	1
		RTH7460D1018	1
		TH6220D1028	1

		RTH2300B1012	1
		RTH7500D	1
		RTH6400	1
		Focus Pro 6000	1
		RTH8500D	1
	Noma	THM301M	1
	Schluter systems	Ditra heat touch screen	1
	Vive comfort	T-P-705	1
	Robertshaw	9600	1
	White-Rodgers	1F80-71	1
		1E78-151	1
Smart	Carrier	Infinity touch	2
	Ecobee	3 Lite	2
		3	2
		4	1
	Goodman	CTK03 ComfortNet	1
	Honeywell	WIFI Color touchscreen thermostat	1

Table A.2: "Dream thermostat" features

Features	Count
Feedback	
More energy use feedback (generally)	9
Gas used per day	3
Indication of monetary savings when adjusting setpoint (dynamic)	2
Electricity used per day	1
Projections of monetary spending per month/year/season based on setpoint (static)	1
Projections of energy savings for different usage profiles	1
Hourly breakdown of costs	1
Wants interface to include...	
Touchscreen display	7
Intuitive scheduling	5
Large screen for readability	3
Display setpoint and current temperature (simultaneously)	3
Display outdoor temperature	2
Display weather	2
Bright screen for readability	2
Suggestions for energy savings	2
Use of natural language	2
More simplicity (there are too many features)	2
Voice control	2
Display outdoor windspeed	1
Display bus schedule	1
Display furnace filter status	1
Display furnace errors	1
Prominent clock	1
Large text for readability	1
Save button	1

Undo button	1
Suggestions for vent adjustment based on remote sensors	1
Display humidity	1
Ability to switch between Fahrenheit and Celsius	1
Preview schedule	1
System features, users want…	
Remote (e.g., app) control	26
Zone control heat distribution	13
Ability to program a schedule	4
Smart thermostat	3
Remote sensors	3
AI learning	3
Humidity control	2
Manual thermostat control	2
Vacation mode	2
Reminders to setback temperature (text from your thermostat)	1
Program flexibility (constrained to 4 time periods)	1
Occupancy sensing and automatic adjustment	1

Appendix B: Interview questions

Part 1: Establishing participant and household characteristics

1. How old is your home?
2. How old are you?
3. Are you the primary user of the thermostat?
4. Do you own or rent your current home?
5. Approximately many homes have you lived in or owned?
6. What kind of heating system do you have?
    a. Can you please take me to your heating system and show it to me with your camera?
7. Where is the thermostat located in your home?
    a. Can you show me where the thermostat is located in your home with your camera?
    b. Is it susceptible to temperature fluctuations?
    c. Are there any obstructions to using the thermostat interface?
    d. Do you have any remote sensors connected to the thermostat in your home?

Part 2: Think-aloud exercise

Instruction: using your camera, show me the thermostat interface as you complete the following 6 tasks. The purpose of this exercise is for me to see what you're doing and know what you're thinking. Please explain your logic for every action and what you're thinking while you complete the tasks.

For Smart or Programmable Thermostat	
Task	Action
1	Turn the thermostat from ‘HEAT to ‘OFF
2	Change the thermostat’s time to 3:00pm
3	Identify the temperature the device is set to reach “Show me the setpoint temperature” “Show me the current temperature”
4	Identify the temperature the thermostat is set to reach on Thursday at 9pm
5	Put the thermostat in vacation mode to keep the temperature the same while you're gone
6	Program your thermostat to be 24 degrees Celsius, 8am-10pm and 22 degrees Celsius, 10pm-8am

For Manual Thermostat	
Task	Action
1	Turn the thermostat from ‘HEAT to ‘OFF
2	Identify the temperature the device is set to reach “Show me the setpoint temperature” “Show me the current temperature”

Part 3: Thermostat investigation

8. How may thermostats do you have in your home?
9. What is your thermostat(s) brand/model number?
10. Do you consider your thermostat aesthetically pleasing?
    a. Do you care if it is aesthetically pleasing?

11. Did you purchase your thermostat yourself or was it already installed when you moved in?
    a. Yes: Why did you choose this thermostat & what was your previous thermostat lacking?
    b. No: What do you like/dislike about your thermostat?
12. Are you considering replacing your thermostat? If yes, why?
13. What is a comfortable setpoint temperature for your home?
14. Do you feel that the temperature displayed on your thermostat interface is accurate? Why do you think so/not?
15. Are there uncomfortable areas in your home?
    a. Does the thermostat improve comfort in those places?
16. How often do you adjust your thermostat manually?
    a. Are there challenges when adjusting your thermostat?
17. Have you programmed your thermostat to a schedule?
    b. If yes, was it challenging?
    c. If no, why not?
18. What functions does your thermostat have (e.g., sleep, hold, vacation)?
    a. Which of these functions do you use and when?
19. Do you feel confident adjusting the temperature on your thermostat?
    c. Why/why not?
20. How does your thermostat work in relation to your heating/cooling system?
    a. At what temperature does the heating/cooling come on?
    b. How does your adjustment of the thermostat relate to the heating system?

c. Does setting the thermostat to a more extreme value speed up the heating/cooling?

d. Does having a lower temperature at night save energy?

21. When you adjust your thermostat, how do you know the HVAC system is working?

    a. Do you hear anything?

    b. Do you see anything?

    c. Do you feel anything?

22. How long does it take for the system to respond after you adjust the thermostat?

23. Do you try to save energy with your thermostat? If you wanted to, do you know how?

    a. Are you concerned with saving energy with your thermostat usage?

    b. Do you know that you can use your thermostat to impact your energy consumption?

    c. Are you conscious of energy consumption when you adjust your thermostat?

    d. Have you taken any steps to try and save energy with your thermostat?

        i. What have you done?

    e. Do you find it difficult to save energy using your thermostat?

24. If you make an error, how do you fix it?

    a. Does your thermostat have an undo or reverse change function?

25. Are there any features you wish your thermostat did/didn't have?

26. If you were to rate your thermostat's usability, where would it lie?

References

[1] Natural Resources Canada, "Heating equipment for residential use," *NRCAN*, 2022. [Online]. Available: https://www.nrcan.gc.ca/energy-efficiency/products/product-information/heating-equipment-for-residential-use/13740.

[2] T. Peffer, M. Pritoni, A. Meier, C. Aragon, and D. Perry, "How people use thermostats in homes: A review," *Build. Environ.*, vol. 46, no. 12, pp. 2529–2541, 2011.

[3] M. Nevius and S. Pigg, "Programmable thermostats that go berserk? Taking a social perspective on space heating in Wisconsin," *Energy Cent. Wisconsin Dep. Sociol. Univ. Wisconsin*, pp. 233–244, 2000.

[4] A. Meier, C. Aragon, and B. Hurwitz, "How People Actually Use Thermostats," *ACEEE Summer study Energy Effic. Build.*, 2010.

[5] M. Shipworth, S. K. Firth, M. I. Gentry, A. J. Wright, D. T. Shipworth, and K. J. Lomas, "Central heating thermostat settings and timing: building demographics," *Build. Res. Inf.*, vol. 38, no. 1, pp. 50–69, 2010.

[6] J. K. Day and D. E. Gunderson, "Understanding high performance buildings: The link between occupant knowledge of passive design systems, corresponding behaviors, occupant comfort and environmental satisfaction," *Build. Environ.*, 2015.

[7] T. Peffer, D. Perry, M. Pritoni, C. Aragon, and A. Meier, "Facilitating energy savings with programmable thermostats: Evaluation and guidelines for the thermostat user interface," *Ergonomics*, vol. 56, no. 3, pp. 463–479, 2013.

[8] P. Foran, "Thousands of Ontario homeowners are eligible for a free smart thermostat and here's how to get one," *CTV News*, Toronto, 19-Oct-2021.

[9] Enbridge Gas, "Home Winterproofing Program." [Online]. Available: https://www.enbridgegas.com/residential/rebates-energy-conservation/home-winterproofing-program. [Accessed: 13-Jun-2022].

[10] City of Ottawa, "Heritage conservation: Overview," 2022. [Online]. Available: https://ottawa.ca/en/planning-development-and-construction/heritage-conservation#:~:text=The City has over 300,and management of heritage resources. [Accessed: 13-Jun-2022].

[11] City of Ottawa, "Heritage conservation districts," 2022. .

[12] P. J. Boait and R. M. Rylatt, "A method for fully automatic operation of domestic heating," *Energy Build.*, vol. 42, no. 1, pp. 11–16, 2010.

[13] R. Critchley, J. Gilbertson, M. Grimsley, G. Green, and W. F. S. Group, "Living in cold homes after heating improvements: evidence from Warm-Front, England's Home Energy Efficiency Scheme," *Appl. Energy*, vol. 84, no. 2, pp. 147–158, 2007.

[14] R. C. Diamond, "Comfort and Control: Energy and Housing for the Elderly," *Environ. Des. Res. Assoc.*, vol. 15, 1984.

[15] A. Freudenthal and H. J. Mook, "The evaluation of an innovative intelligent thermostat interface: universal usability and age differences," *Cogn. Technol. Work*, vol. 5, no. 1, pp. 55–66, 2003.

[16] S. Karjalainen, "The characteristics of usable room temperature control," 2007.

[17] A.-L. Lindén, A. Carlsson-Kanyama, and B. Eriksson, "Efficient and inefficient aspects of residential energy behaviour: What are the policy instruments for change?," *Energy Policy*, vol. 34, no. 14, pp. 1918–1927, 2006.

[18] D. Cross and D. Judd, "Automatic setback thermostats: Measure persistence and customer behavior," in *Proceedings of the international energy program evaluation conference*, 1997.

[19] C. Haiad, J. Peterson, P. Reeves, and J. Hirsch, "Programmable thermostats installed into residential buildings: Predicting energy savings using occupant behavior & simulation," *South. Calif. Edison*, 2004.

[20] G. Dewis, "Thermostat use in Canadian homes," *EnviroStats*, vol. 2, no. 3, pp. 3–22, 2008.

[21] R. L. W. Analytics, "Validating the impact of programmable thermostats," *Middletown, CT*, 2007.

[22] N. Michaud, L. Megdal, P. Baillargeon, and C. Acocella, "Billing Analysis & Environment that 'Re-Sets' Savings for Programmable Thermostats in New Homes," in *Energy Program Evaluation Conference, Portland*, 2009, pp. 597–606.

[23] D. Tachibana, "Residential customer characteristics survey 2009," *Seattle City Light*, 2010.

[24] J. K. Day *et al.*, "A review of select human-building interfaces and their relationship to human behavior, energy use and occupant comfort," *Build. Environ.*, vol. 178, no. May, p. 106920, 2020.

[25] B. Huchuk, W. O'Brien, and S. Sanner, "A longitudinal study of thermostat behaviors based on climate, seasonal, and energy price considerations using connected thermostat data," *Build. Environ.*, 2018.

[26] S. Karjalainen, "Usability guidelines for room temperature controls," *Intell. Build. Int.*, vol. 2, no. 2, pp. 85–97, 2010.

[27] K. Herter and Y. Okuneva, "2015 Communicating Thermostat Usability Study," Sacramento, 2015.

[28] K. M. A. Revell and N. A. Stanton, "Case studies of mental models in home heat control:

Searching for feedback, valve, timer and switch theories," *Appl. Ergon.*, vol. 45, no. 3, pp. 363–378, 2014.

[29] W. Kempton, "Two theories of home heat control," *Cogn. Sci.*, vol. 10, no. 1, pp. 75–90, 1986.

[30] D. Norman, *The Design of Everyday Things*. 2016.

[31] K. M. A. Revell and N. A. Stanton, "Mental model interface design: putting users in control of home heating," *Build. Res. Inf.*, vol. 46, no. 3, pp. 251–271, 2018.

[32] R. Vastamäki, I. Sinkkonen, and C. Leinonen, "A behavioural model of temperature controller usage and energy saving," *Pers. Ubiquitous Comput.*, vol. 9, no. 4, pp. 250–259, 2005.

[33] H. Fujii and L. Lutzenhiser, "Japanese residential air-conditioning: natural cooling and intelligent systems," *Energy Build.*, vol. 18, no. 3–4, pp. 221–233, 1992.

[34] J. Sauer, D. G. Wastell, and C. Schmeink, "Designing for the home: A comparative study of support aids for central heating systems," *Appl. Ergon.*, vol. 40, no. 2, pp. 165–174, 2009.

[35] C. Preville, "Ushering in the modern age of heating," *Air Cond. Heat. Refrig. News*, vol. 214, no. 11, pp. 27–31, 2001.

[36] D. Perry, C. Aragon, A. Meier, T. Peffer, and M. Pritoni, "Making Energy Savings Easier: Usability Metrics for Thermostats," *J. Usability Stud.*, vol. 6, no. 4, pp. 226–244, 2011.

[37] A. Meier, C. Aragon, T. Peffer, D. Perry, and M. Pritoni, "Usability of residential thermostats: Preliminary investigations," *Build. Environ.*, vol. 46, no. 10, pp. 1891–1898, 2011.

[38] T. Malinick *et al.*, "Destined to disappoint: programmable thermostat savings are only as good as the assumptions about their operating characteristics," in *ACEEE Summer Study on Energy Efficiency in Buildings*, 2012, vol. 12.

[39] M. Pritoni, A. K. Meier, C. Aragon, D. Perry, and T. Peffer, "Energy efficiency and the misuse

of programmable thermostats: The effectiveness of crowdsourcing for understanding household behavior," *Energy Res. Soc. Sci.*, vol. 8, pp. 190–197, 2015.

[40] K. Macchiarelli, "Pioneered to perfection: Introducing the new ecobee SmartThermostat with voice control," *Ecobee Newsroom*, 2019.

[41] Nest Labs, "Energy Savings from the Nest Learning Thermostat : Energy Bill Analysis Results," 2015.

[42] S. Karjalainen and O. Koistinen, "User problems with individual temperature control in offices," *Build. Environ.*, 2007.

[43] B. Bordass, A. Leaman, and R. Bunn, "Controls for end users: a guide for good design and implementation," *Build. Control. Ind. Assoc.*, pp. 1–26, 2007.

[44] M. Lee, C. M. Carswell, W. Seidelman, and M. Sublette, "Green expectations: The story of a customizable lighting control panel designed to reduce energy use," *Proc. Hum. Factors Ergon. Soc.*, pp. 1353–1357, 2013.

[45] H. C. A. Dale and C. M. Crawshaw, "Ergonomic aspects of heating controls," *Build. Serv. Eng. Res. Technol.*, vol. 4, no. 1, pp. 22–25, 1983.

[46] S. Karjalainen, "Thermal comfort and use of thermostats in Finnish homes and offices," *Build. Environ.*, vol. 44, no. 6, pp. 1237–1245, Jun. 2009.

[47] N. Combe, D. Harrison, S. Craig, and M. S. Young, "An investigation into usability and exclusivity issues of digital programmable thermostats," *J. Eng. Des.*, vol. 23, no. 5, pp. 401–417, 2012.

[48] S. Jung *et al.*, "Effect of Touch Button Interface on In-Vehicle Information Systems Usability," *Int. J. Human–Computer Interact.*, pp. 1–19, 2021.

[49] R. Li, Y. V. Chen, C. Sha, and Z. Lu, "Effects of interface layout on the usability of in-vehicle information systems and driving safety," *Displays*, vol. 49, pp. 124–132, 2017.

[50] B. Huchuk, W. O'brien, and S. Sanner, "Exploring smart thermostat users' schedule override behaviors and the energy consequences," *Sci. Technol. Built Environ.*, vol. 27, no. 2, pp. 195–210, 2020.

[51] C. Brackley, W. O'Brien, and C. Trudel, "Evaluation of a Modern Commercial HVAC Control System in the Context of Occupants' Perceived Control.," *ASHRAE Trans.*, vol. 126, no. 2, 2020.

[52] R. A. Virzi, "Refining the test phase of usability evaluation: How many subjects is enough?," *Hum. Factors*, vol. 34, no. 4, pp. 457–468, 1992.

[53] B. Martin and B. Hanington, *Universal methods of design: 100 ways to research complex problems, develop innovative ideas and design effective solutions*. 2012.

[54] K. A. Ericsson and H. A. Simon, "Verbal reports as data.," *Psychol. Rev.*, vol. 87, no. 3, p. 215, 1980.

[55] A. Freudenthal, "The design of home appliances for young and old consumers," 1999.

[56] M. J. Rooden, "Design for anticipating future usage." Delft, 2001.

[57] National Renewable Energy Laboratory, "NREL: A Year in Clean Energy Innovations - A Review of NREL's 2011 Feature Stories," 2011.

[58] A. Fanara, "ENERGY STAR Programmable Thermostats Proposal Industry," 2003.

[59] K. Rathouse and B. Young, "RPDH15: Use of domestic heating controls," *Defra's Mark. Transform. Program.*, 2004.

[60] R. C. Diamond, "Energy use among the low-income elderly: a closer look," 1984.

[61] T. G. Moore and A. Dartnall, "Human factors of a microelectronic product: the central heating timer/programmer," *Appl. Ergon.*, vol. 13, no. 1, pp. 15–23, 1982.

[62] L. Lutzenhiser, "A question of control: alternative patterns of room air-conditioner use," *Energy Build.*, vol. 18, no. 3–4, pp. 193–200, 1992.

[63] B. Park, A. Korbach, and R. Brünken, "Does thinking-aloud affect learning, visual information processing and cognitive load when learning with seductive details as expected from self-regulation perspective?," *Comput. Human Behav.*, vol. 111, p. 106411, 2020.

[64] L. Bourland, "The History of Thermostats and Their Evolution to Smart," *Mysa*, 2020. .

[65] D. Melanson, "Ecobee debuts energy-saving, WiFi-connected Smart Thermostat," *Engadget*, 2008. .

[66] Mordor Intelligence, "Smart Thermostat Market - Growth, Trends, and Forecast (2020-2025)," 2020.

[67] M. Bäckman, "Smart Homes and Home Automation," *IOT Res. Ser.*, 2021.

[68] J. W. Moon and S.-H. Han, "Thermostat strategies impact on energy consumption in residential buildings," *Energy Build.*, vol. 43, no. 2–3, pp. 338–346, 2011.

[69] K. M. A. Revell and N. A. Stanton, "When energy saving advice leads to more, rather than less, consumption," *Int. J. Sustain. Energy*, vol. 36, no. 1, pp. 1–19, 2017.

[70] D. Norman, *The design of everyday things: Revised and expanded edition*. Basic books, 2013.

[71] L. A. Rafferty, N. A. Stanton, and G. H. Walker, "The famous five factors in teamwork: a case study of fratricide," *Ergonomics*, vol. 53, no. 10, pp. 1187–1204, Oct. 2010.

[72] J. M. Carroll and J. R. Olson, "Mental models in human-computer interaction," *Handb. human-computer Interact.*, pp. 45–65, 1988.

[73] Z. Mack and S. Sharples, "The importance of usability in product choice: A mobile phone case study," *Ergonomics*, vol. 52, no. 12, pp. 1514–1528, 2009.

[74] R. J. Branaghan, C. M. Covas-Smith, K. D. Jackson, and C. Eidman, "Using knowledge structures to redesign an instructor–operator station," *Appl. Ergon.*, vol. 42, no. 6, pp. 934–940, 2011.

[75] D. Lockton, D. Harrison, and N. A. Stanton, "The Design with Intent Method: A design tool for influencing user behaviour," *Appl. Ergon.*, vol. 41, no. 3, pp. 382–392, 2010.

[76] K. M. A. Revell and N. A. Stanton, "Mind the gap–Deriving a compatible user mental model of the home heating system to encourage sustainable behaviour," *Appl. Ergon.*, vol. 57, pp. 48–61, 2016.

[77] K. M. A. Revell and N. A. Stanton, "The Quick Association Check (QuACk): a resource-light, 'bias robust' method for exploring the relationship between mental models and behaviour patterns with home heating systems," *Theor. Issues Ergon. Sci.*, vol. 17, no. 5–6, pp. 554–587, Sep. 2016.

[78] I. S. Walker and A. K. Meier, "Residential thermostats: comfort controls in California Homes," 2008.

[79] K. M. A. Revell and N. A. Stanton, *Mental models: Design of user interaction and interfaces for domestic energy systems*. CRC Press, 2017.

[80] K. M. A. Revell and N. A. Stanton, "Using Interface Design to Promote a Compatible User Mental Model of Home Heating and Pilot of Experiment to Test the Resulting Design," 1st ed., CRC Press, 2017, pp. 183–200.

[81] K. Manktelow and J. Jones, "Priniciples from the psychology of thinking and mental models,"

in *Applying cognitive psychology to user-interface design*, 1987, pp. 83–117.

[82] R. Tamas, W. O'Brien, and M. S. Quintero, "Residential thermostat usability: Comparing manual, programmable, and smart devices," *Build. Environ.*, vol. 203, p. 108104, 2021.

[83] Nautral Resources Canada, "2015 Survey of Household Energy Use (SHEU-2015)," 2015.

[84] Government of Canada, "Monthly Climate Summaries," 2022. [Online]. Available: https://climate.weather.gc.ca/prods_servs/cdn_climate_summary_e.html.

[85] D. Dewees and T. Tombe, "The impact of sub-metering on condominium electricity demand," *Can. Public Policy*, vol. 37, no. 4, pp. 435–457, 2011.

[86] J. Teres-Zubiaga, E. Pérez-Iribarren, I. González-Pino, and J. M. Sala, "Effects of individual metering and charging of heating and domestic hot water on energy consumption of buildings in temperate climates," *Energy Convers. Manag.*, vol. 171, pp. 491–506, 2018.

[87] S. Siggelsten and B. Hansson, "Incentives for individual metering and charging," *J. Facil. Manag.*, 2010.

[88] U.S. Environmental Protection Agency (EPA), "Definitions for Energy Star Programmable Thermostat Proposal." pp. 1–2, 2003.

[89] M. F. Touchie and J. A. Siegel, "Residential HVAC runtime from smart thermostats: characterization, comparison, and impacts," *Indoor Air*, vol. 28, no. 6, pp. 905–915, 2018.