

THE DESIGN OF HOME ENERGY-MANAGEMENT INTERFACES:
EFFECTS OF DISPLAY TYPE ON THERMOSTAT TEMPERATURE SELECTION

by

Joshua Stein

May 2013

A thesis submitted in conformity with the requirements
for the degree of Masters of Applied Science,
Graduate Department of Mechanical and Industrial Engineering,
University of Toronto.

The Design of Home Energy-Management Interfaces:
Effects of Display Type on Thermostat Temperature Selection

Joshua Stein

Masters of Applied Science,

2013

Graduate Department of Mechanical and Industrial Engineering,

University of Toronto.

Abstract

This thesis explores home energy management (HEM), an emerging field for interface design and sustainability. Section 1 introduces HEM's broader context. In Section 2, I review the literature surrounding HEM. Section 3 outlines the usability study on the ecobee Smart Thermostat, to evaluate the technology's ease-of-use, and better understand users' experience with current HEM technology. Section 4 describes a "Critical Making" workshop, where participants investigated HEM through material interaction and discussion. Section 5 describes and evaluates the potential design spaces gleaned from previous sections. In Section 6, I return to the literature to investigate key concepts underlying the design intervention for the chosen design space. Section 7 describes my design intervention and experimental evaluation. In Section 8, I present the study results, which suggest enhanced display labelling had a significant and directional effect on user-selected temperatures. In Section 9, I discuss these results, study limitations, and make conclusions and recommendations.

Acknowledgements

I want to thank my supervisor, Prof. Greg Jamieson, for his continued support and patience through my many ups and downs in my research, for providing me with just enough guidance to find my way, but not too much that I would not learn anything on my own. Thanks to Prof. Matt Ratto in the Faculty of Information for connecting me with an incredible opportunity to study HEM in depth with ecobee Inc. Thanks to fellow Master's student Mike Borg for putting together the components of the FEDDEV project with me and for helping me to run the Usability Study. Thanks to ecobee Inc for their enthusiastic participation in this project. Thanks to Kevin Trihn and Antony Hilliard carrying the sustainability flame at CEL and for their input and encouragement. Thanks to Antony as well for eagerly reviewing early versions of my thesis and other important documents along the way.

Table of Contents

1.0	Introduction.....	1
2.0	Review of the Initial Literature	3
2.1	The Problems with Residential Energy Consumption	3
2.1.1	The Disconnect Between Attitudes and Behaviours	3
2.1.2	Poor Mental Models	3
2.2	Existing Approaches to Residential Energy Conservation	6
2.2.1	Curtailment-Strategies Overview	6
2.2.2	Feedback.....	9
2.2.3	Habits	11
2.3	Home Energy Management.....	12
2.3.1	Electrical HEM.....	13
2.3.2	Programmable (and Manual) Thermostats.....	14
2.4	Moving Forward.....	17
3.0	Observational Review of a Current HEM Product.....	20
3.1	Purpose.....	20
3.2	Methods	20
3.3	Participant Demographics.....	24
3.4	Results.....	25
3.4.1	Quantitative Results by Task	25
3.4.2	Qualitative Results by Issue	27
3.5	Discussion.....	34
3.5.1	Favoured Attributes.....	34
3.5.2	Least Favoured Attributes	34
3.5.3	Participant Ideas and Desired Functionality	35
3.5.4	Other User Comments.....	35
3.5.5	Expert Workarounds.....	36
3.6	Recommendations	36
3.7	Moving Forward.....	36
4.0	The Critical Making Workshop.....	38
4.1	Critical Making.....	38
4.2	Development of Our Session	38
4.2.1	Details of our Critical-Making Workshop Idea.....	39
4.3	Session Overview and Highlights	39
4.3.1	Qualitative Session Highlights.....	40
4.3.2	Design Spaces Highlighted from the Critical Making Session.....	41
4.4	Moving Forward.....	42
5.0	Design Spaces and Selection.....	43
5.1	Top Five Design Spaces.....	43
5.1.1	Ambiguous understanding of Furnace Operation	43
5.1.2	Lack of Device Context → Activity-Based Energy Feedback.....	43
5.1.3	Accommodating Multiple (sometimes conflicting) desires for Heating	43
5.1.4	Cost versus Social and Environmental Motivations	44
5.1.5	Temperature Selection Ambiguity	44

5.2	Selection Criteria	44
5.2.1	Automation Pitfalls (Lee, 2006)	45
5.2.2	Automation Strategies (Lee, 2006)	46
5.3	Evaluation Based on Selection Criteria.....	47
5.3.1	Pitfalls and Strategies Applied to Temperature Selection Ambiguity.....	49
5.4	Choice of Design Space.....	51
6.0	Comfort, Control Labelling, and The Anchoring Effect	52
6.1	Comfort and the Design of H&C Systems	52
6.1.1	Adaptive Comfort Standard for Naturally-Ventilated Buildings	55
6.2	Control Labelling	57
6.2.1	Human Factors Principles and Guidelines for Controls and Displays	57
6.2.2	Warnings Literature	59
6.2.3	Enhanced Control Labelling.....	60
6.3	The Anchoring Effect	62
6.3.1	Anchoring and Social Norms.....	63
6.4	Moving Forward.....	64
7.0	Prototype Development and Experimental Evaluation.....	67
7.1	Research Questions	67
7.2	Prototype Details and Research Method.....	68
7.3	Experimental Design.....	70
7.4	Hypotheses.....	74
7.5	Participants.....	74
7.6	Materials.....	74
7.7	Procedure	75
7.8	Moving Forward.....	77
8.0	Experimental Results.....	78
8.1	Normality and Homogeneity of Variance.....	78
8.2	Descriptive Statistics for the Independent Variables	79
8.3	Main Effects.....	79
8.4	Confidence Scores	81
8.5	Supplementary Data Collected	82
8.6	Questionnaire Data	83
8.6.1	Mental Model Checks	84
8.7	Moving Forward.....	87
9.0	Analysis and Discussion	88
9.1	Mental Models.....	90
9.2	Limitations	92
9.3	Conclusions and Recommendations	94
10.0	References	96
	Appendix A – User Recruitment Letter	106
	Appendix B – Letter of Consent for Usability Study	108
	Appendix C – User Aides.....	110
	Appendix D – List of Tasks.....	111

Appendix E – Usability Study Participant Questionnaire.....	112
Appendix F – Quantitative Data.....	113
Appendix G – Subjective Usability Metrics.....	114
Appendix H – Full List of Qualitative Usability Issues.....	115
Appendix I – Critical Making Character Questionnaire.....	121
Appendix J – Informed Consent.....	123
Appendix K – Thermostat Paper-Prototypes.....	125
Appendix L – Participant Questionnaire	130
Appendix M – Debrief Form.....	131
Appendix N – Normality Plots.....	132
Appendix O – Study Calculations and Statistical Tests.....	142

List of Tables

Table 3-1 – Simplified list of usability tasks	23
Table 3-2 – Participant Details	25
Table 3-3 – Average participant task-time and errors across tasks 4, 5a, 5b, 6a, 7, and 10	27
Table 7-1 – Participants’ Academic Programs	74
Table 7-2 – Number and percent of participants recruited at each location.....	75
Table 8-1 – Descriptive statistics for each condition.....	79
Table 8-2 – Number and percent of participants who rated at different levels of confidence in their selected temperatures.....	81
Table 8-3 – Participants’ answer to the question “How comfortable are you right now?” ..	83
Table D-10-1 – List of Tasks in the Usability Study	111
Table F-10-2– Errors and task time, by task, for each participant – dashes indicate tasks not recorded/performed	113
Table H-10-3 – Qualitative Usability Issues – Category: Confusion	115
Table H-10-4 – Qualitative Usability Issues – Category: System Errors.....	116
Table H-10-5 – Qualitative Usability Issues – Category: User Errors	116
Table H-10-6 – Qualitative Usability Issues – Category: Action Difficulty.....	117
Table H-10-7 – Qualitative Usability Issues – Category: System Quirks and Pitfalls	118
Table H-10-8 – Qualitative Usability Issues – Category: Accessory-Specific Issues (not mutually exclusive from above categories).....	119
Table O-10-9– Tests of Skewness.....	142
Table O-10-10– Tests of Kurtosis	142
Table O-10-11– SPSS Output for Tests of Normality.....	143
Table O-10-12 – SPSS Output for Levene’s Test	143
Table O-10-13 – SPSS Output for the mixed-ANOVA within-subjects effects, for instruction type	144
Table O-10-14 – SPSS Output for the mixed-ANOVA between-subjects effects, for display type	144
Table O-10-15— SPSS Output for between-subjects planned contrasts, for display type ..	145
Table O-10-16 — SPSS Output for the Kruskal-Wallis Test of confidence scores by display type	146
Table O-10-17 – SPSS Output of the test statistics for the K-W test.....	146

Table O-10-18 – SPSS Output from the Wilcoxon Signed-Rank test of confidence scores by instruction type	147
Table O-10-19 – SPSS Output of the test statistics for the WS-R test	147

List of Figures

Figure 2.1 – Household savings broken down by feedback type (Ehrhardt-Martinez et al., 2010)	10
Figure 3.1 – Time to complete each task.....	26
Figure 3.2 – Number of user-errors, by task.....	26
Figure 6.1 – The adaptive standard for naturally ventilated buildings (Brager & De Dear, 2000).....	56
Figure 8.1 – User-selected temperature for each condition. Error bars represent the 95% confidence intervals.....	80
Figure 8.2 – Number of participants who rated at different levels of confidence in their selected temperatures	82
Figure 8.3 – User-selected temperature for participants with and without control over their H&C system. Error bars represent the 95% confidence intervals.	85
Figure 8.4 – User-selected temperature for mental-model question #2. Error bars represent the 95% confidence intervals.	86
Figure 8.5 – User-selected temperature for mental-model question #3. Error bars represent the 95% confidence intervals.	87
Figure B.10.1 – Using the Touch Screen	110
Figure B.10.2 – Feature Buttons.....	110

1.0 Introduction

Home energy management (HEM) is an emerging field for both interface design and energy conservation. In Canada, households use 16% of total energy (Government of Canada, 2013b) and produce 14% of CO₂ emissions (Government of Canada, 2013c); space-heating accounts for 63% of residential energy use and 62% of CO₂ emissions, compared to 20% energy used by electrical appliances and lighting and 21% of CO₂ emissions (Government of Canada, 2013a).

Of this, at least 25% of the variance in home energy use can be attributed to behaviour (Sauer, Franke, & Ruettinger, 2008; Verhallen & Van Raaij, 1981). Parker et al (1996) found that, in low-income centrally-ventilated apartments, every 1°C below a setting of 28°C accounted for an increase of 23% air conditioning (AC) energy consumption. They also found that, on a peak-use day, an indoor temperature difference of 6.5° C between homes coincided with a three-times difference in energy use (between 14 and 45 kWh/day).

Opportunities for savings also highlight the cost of inaction. In summertime, increasing room temperatures by just 1°C may lower AC energy consumption by ~two to four percent in large commercial buildings (Lovins, 1992). In another summertime study, an experimental group cut their electricity demand by 20% by increasing their indoor temperature by about 1.5 °C compared with control, with minimal change in comfort or clothing levels (Lovins, 1992).

Sometimes incidental differences in ecology can have a great effect on energy consumption. For example residents of master-metered buildings use 35% more energy on average than those in separately-metered units (Lutzenhiser, 1993).

A central problem addressed in this thesis is the failure of demand-side energy technology to effectively address the human factor. As we will see below, users generally find modern home-energy management devices and thermostats difficult to use and understand. For

instance, despite the great savings opportunities, 90% of programmable thermostats never get programmed. (Meier, Aragon, Peffer, Perry, & Pritoni, 2011).

This work addresses HEM technology, and does so from a number of perspectives. Section 2 reviews the literature, defining the challenges in HEM and how they have been addressed thus far. Section 3 outlines a usability study on a current HEM device. Section 4 describes a participant “Critical Making” workshop, where we investigated HEM through material interaction and discussion. Section 5 describes and evaluates the potential design spaces, identifying one problem to address through design. Section 6 returns to the literature to investigate key concepts underlying the design intervention. Section 7 describes the basis for my design intervention and details the experimental evaluation. In Section 8, I present the results of this evaluation. I finish with Section 9, an analysis and discussion of these results, followed by conclusions and recommendations.

This thesis takes a narrative approach to describing the contents of my research, as each phase of my activities were dependent on the previous phase. For instance, the research questions were only developed after a process that illuminated several design spaces from my previous investigations.

2.0 Review of the Initial Literature

This section first outlines some of the challenges associated with residential energy consumption, follows with a review of existing strategies to address the behavioural challenges, and ends with a review of home-energy management technology.

2.1 The Problems with Residential Energy Consumption

The following subsections explore some of the challenges of residential energy management through looking the human attitude-behaviour split and the issue of poor mental models.

2.1.1 The Disconnect Between Attitudes and Behaviours

It often takes more than pro-environmental attitudes for people to reduce their energy consumption. A few studies on the use of domestic appliances found no overall association between environmental concern and energy use (Sauer & Rüttinger, 2007; Sauer, Wiese, & Rüttinger, 2004). For example, investigating energy use of vacuum cleaners, Sauer et al (2002) found high environmental concern correlated with *longer* cleaning times (thus consuming more energy). This becomes relevant for interventions that attempt to change attitudes without addressing the needed behaviours.

2.1.2 Poor Mental Models

A number of researchers have presented evidence suggesting ordinary people lack accurate mental models of the use of energy at home, and thus rely on folk theories (Karjalainen & Vastamaeki, 2007; Kempton, Feuermann, & McGarity, 1992; Kempton & Montgomery, 1982; Kempton, 1986). Kempton (1986) describes a 'folk' theory as a (usually) inaccurate model of reality, made in an attempt to make sense of the world through everyday experience and social interaction. Even though folk theories are usually inaccurate, they subjectively appear to be confirmed by people's normal experience (Lutzenhiser, 1993; Norman, 2002).

Kempton and Montgomery (1982) found, for instance, that homeowners tracked their home's energy efficiency (and the effectiveness of behavioural and structural changes) by the cost of their utilities rather than the actual amount of energy consumed. People experimented by changing a single factor and looking at the difference in their monthly energy bill, often to be discouraged by the small savings (or increased cost) unknowingly due to other factors (like seasonal changes).

Kempton (1986) describes most homeowners as having one of two folk theories of home heat control. Users holding the 'Valve' model (between 25-50% of Americans) believe that a thermostat works like a valve, in that a higher temperature setting will deliver heat at a faster rate than a lower setting. The valve model is inaccurate because real heating systems (and ACs) produce (or remove) heat at a constant rate, and can only be turned on or off by the thermostat. Conversely, those holding the 'Feedback' model believe that thermostats sense and compare current and set temperatures, controlling the heater (or air conditioner) to add (or remove) heat *at a constant rate* until both current and desired temperatures meet; in 'Feedback' theory, the amount of temperature increase depends on length of time. While the 'Feedback' theory is closer to the physical functionality of a heating and cooling system (H&C system), it fails to acknowledge the effect of heat-loss (or gain) from inside-outside temperature differences affecting both the rate of temperature change and energy consumption^{1,2}. Paradoxically, the more incorrect 'Valve' model motivates homeowners to lower temperatures to save energy in the winter, which they falsely attribute to a reduced instantaneous rate of heat output. 'Feedback'-theory holders would actually avoid setting back temperatures, because they believe the energy saved from a setback would be used up again when the temperature was restored. Rathouse and Young (2004) drew similar conclusions, where survey participants debated the efficiency of controlling their heat continuously (consistent with Feedback theory) versus intermittently (consistent with Valve theory).

¹ The rate of heat transfer between two adjacent zones is proportional to the temperature difference between those two zones. Thus in a building, the rate of heat loss (or gain) between inside and outside, proportional to the temperature difference, is also approximately equal to the power required from a heat source (or AC) to replace that heat loss (or gain) (Kempton, 1986; McQuiston, Parker, & Spitler, 2004).

² Thermostatic radiator valves also control heat through thermostatic feedback. Standard radiator valves simply turn radiators on or off ("Thermostatic Radiator Valve and standard radiator valves," 2012).

People also show variable and generally limited comprehension of their heating controls, buildings, and heating systems, with some not at all interested in how they worked (Lutzenhiser, 1993; Rathouse & Young, 2004). Residents surveyed in the UK showed flawed mental models similar to those found by Kempton, including the understanding that a thermostat is an on-off switch (Rathouse & Young, 2004). Some did not understand how lower temperature settings were more energy efficient, that heating had any connection to energy use, nor the connection between opening windows and energy waste. Sources of misinformation included friends, family, advertisements, 'official' information, and rumours.

Meier et al. (2011) found that some users failed to comprehend the importance of how energy-saving technologies are used: "We had the impression, based on the responses to survey questions and supplementary information, that some consumers believed that purchasing and installing an 'energy-saving programmable' thermostat would automatically result in lower energy use." (p. 1897).

Another energy-misperception is that people perceived that more salient devices consumed more energy (Kempton and Montgomery, 1982). Kempton and Montgomery found lights were falsely thought to consume more energy than a water heater.

People also seem to have a flawed understanding of appropriate temperature selection. Karjalainen & Vastamaeki (2007) found that 41% of surveyed occupants of Finish homes believed (falsely) that comfortable indoor temperatures should be warmer in the winter than in the summer. This goes against well established research that humans find warmer temperatures more comfortable in the summer than in the winter, because of clothing levels and psychophysical adaptation (Brager & De Dear, 1998, 2000; De Dear & Brager, 1998; Karjalainen & Vastamaeki, 2007; Lovins, 1992). Only 15% correctly thought the reverse; that comfortable indoor temperatures are lower in the winter than in the summer. When asked for specific temperatures, participants reported, on average, they would prefer 21.2°C in the winter and 20.5°C in the summer. The topic of temperature selection is a primary theme in this thesis.

2.2 Existing Approaches to Residential Energy Conservation

Abrahamse, Steg, Vlek, & Rothengatter (2005) outline two broad categories of existing approaches to residential energy conservation. “Efficiency” approaches are one-time investments to increase the energy efficiency of the home and its appliances; these will not be discussed further. “Curtailment” strategies aim at changing people’s daily energy-use behaviours; these are outlined below. One particular curtailment strategy, providing feedback, deserves separate mention below. This subsection ends with a review of the importance of considering habits in behavioural interventions.

2.2.1 Curtailment-Strategies Overview

This section overviews some of the more prominent and thesis-relevant curtailment strategies used to change energy-consumption behaviours.

Providing Information

Most campaigns aimed at curtailment focus on providing people with information in an attempt to change their attitudes or highlight economic benefits (McKenzie-Mohr, 2011). Consistent with the attitude-behaviour split (Section 2.1.1), providing information may change attitudes, but does not necessarily change related behaviours (McKenzie-Mohr, 2011; Verplanken & Wood, 2006) or lower energy consumption (Abrahamse et al., 2005). For example, Sauer, Wiese, and Ruttinger (2003) found that participants’ knowledge of environmental impacts did not predict their energy consumption in their use of consumer appliances.

However, providing information is still necessary and can be effective if well implemented. McKenzie-Mohr (2011) posits that for a message to be effective and influential, it should (to name a few): capture the reader’s attention; be vivid and captivating; be tailored to the attitudes and beliefs of the intended audience, and their perceived barriers and benefits to taking action; cite a credible source; frame the message to highlight a potential loss; provide actionable solutions when highlighting something that may threaten the reader; keep instructions clear, specific, and easy to remember; and be combined with other approaches (like those mentioned in this section).

Social Norms

Social norms are people's perceptions of what is "normal", with respect to behaviours and attitudes. Decades of research has established that norms effectively, directly, and meaningfully, trigger and guide behaviours (Schultz, Nolan, Cialdini, Goldstein, & Griskevicius, 2007). People generally rate normative interventions as least likely to affect their behaviour, but they are nonetheless most affected by them compared to other appeals (Cialdini, 2007).

There are (generally) two categories of social norms, which are thought to separately influence behaviour: injunctive and descriptive (Bailey, Nofsinger, & O'Neill, 2004; Cialdini, 2007; Eyssel, Böhner, & Siebler, 2006; Schultz et al., 2007). Injunctive norms are people's perceptions of what others approve or disapprove of, and motivate behaviour based on avoidance of informal social sanctions or seeking social approval. They can be effective even if they are not from personally-significant others (i.e., friends and family). Descriptive norms are people's perceptions of what others actually do. They give people information about what is appropriate and adaptive conduct in a particular situation.

Shultz et al. (2007) theorized and then showed that descriptive norms motivate people to bring their behaviour closer to the norm, regardless if they are above or below it. By showing people the average energy consumption in their neighbourhood, households above the norm conserved more the following week; however, those below the norm actually *increased* their use; a phenomena Shultz et al. (2007) labelled the "Boomerang Effect". However, combining an injunctive norm (e.g., using smiley and sad faces), expressing approval for being below the norm and disapproval for being above, cancelled the effect so that households below the norm continued conserving energy.

There is mixed evidence about the independent effectiveness of injunctive and descriptive norm approaches (Croy, Gerrans, & Spielman, 2010). Bailey, Nofsinger, and O'Neill (2004) studied effect of injunctive and descriptive norms used together, but found no interaction effects. Cialdini (2007) recommends use of descriptive norms for

behaviours that are already prevalent, and injunctive norms for those that occur less than 50%.

While evidence is mixed about the effect of the perceived social proximity of the “other” on norm effectiveness, it generally reveals that people’s behaviour is more influenced by the groups they associate more strongly with (Berkowitz, 2004; Goldstein, Griskevicius, & Cialdini, 2007; Lewis & Neighbors, 2006; Neighbors et al., 2007). For example, Neighbors et al (2007) found gambling frequency among college students was most highly correlated with the perceived approval of friends, less so by family, and slightly negatively correlated with approval of other students. In Social Comparison Theory, people tend to follow the norms of those who are more similar to themselves, both in personal characteristics and in having a shared context in which a decision is made (Goldstein et al., 2007). In a study by Goldstein et al. (2007) on message cards encouraging hotel guests to reuse their towels, the researchers found that a context-specific descriptive norm (stating that 75% of guests who stayed in their room reused their towels) led to reuse rates of 49%, compared to 44% for a more global message (stating that 75% of guests who stayed at the hotel reused their towels), and 37% for non-normative environmental appeal.

Rewards/Incentives

Incentives, rewards, and disincentives provide people with extrinsic motivation to perform existing, or learn new, behaviours that they would otherwise be indifferent or resistant to (Abrahamse et al., 2005; McKenzie-Mohr, 2011). Implemented correctly, incentives are very effective at fostering sustainable behaviours. For example, introducing bottle deposits in Oregon, Vermont, and Michigan, saw decreases in litter of 68%, 76%, and 82%, respectively (Syrek & Michigan Legislature, 1980). A program in California that charged residents for the amount of waste they put out on the curb, saw a 46% reduction in landfill-bound waste and a 158% increase in recycling (A *Municipal Guide on Economic Instruments to Support Municipal Waste Management Programs*, 1996). However, Abrahamse et al (2005) found that while rewards produce large effects, these effects quickly diminish once the reward is discontinued.

Prompts

Prompts are simple reminders (delivered at the right time and place) to spur action on specific behaviours (McKenzie-Mohr, 2011). Their purpose is not to affect motivation, but to remind people to do something they are already predisposed to do. Good prompts are specific and actionable (McKenzie-Mohr, 2011). In one Australian study, prompts displaying the environmental impacts of water use, placed around taps and showers, decreased water consumption by 23% (Kurz, Donaghue, & Walker, 2005). Another project in the Netherlands saw a 50% reduction in litter through the placement of prompts over waste bins (Kort, McCalley, & Midden, 2008).

Convenience

This strategy stems from the notion that inconvenient (i.e., unpleasant or time consuming) behaviours have low participation levels. McKenzie-Mohr (2011) asserts that the strategies mentioned in this thesis will be ineffective if the intended behaviour is inconvenient. He compares backyard composting, adopted by approximately 30% of Ontario residents, to curbside recycling, adopted by approximately 80% of residents, as one example of the power of convenience (McKenzie-Mohr, 2011). Compare too, in 2011 Toronto residents (as a whole) diverted 100,663 tonnes of waste through the curb-side Green Bin program, compared to only 18,970 tonnes of waste through backyard composting (City of Toronto, 2011). To address convenience, one must address the real and perceived barriers to performing the specific behaviour (McKenzie-Mohr, 2011).

2.2.2 Feedback

Providing users with feedback of their electricity use has been shown to reduce consumption by approximately 5 to 15%, depending on various factors (Ehrhardt-Martinez, Donnelly, & Laitner, 2010; Hargreaves, Nye, & Burgess, 2010). The American Council for an Energy-Efficient Economy (ACEEE) published a report reviewing residential feedback research (Ehrhardt-Martinez et al., 2010). The ACEEE and Abrahamse et al (2005) agree that feedback becomes more effective to the degree it is given more frequently. Figure 2.1 shows a breakdown of electricity savings by the type of feedback employed.

A chart of average annual percent household savings broken down by granularity of feedback (Ehrhardt-Martinez et al., 2010, p. iii). This material has been removed due to copyright restrictions. It can be found here:

<http://sedc-coalition.eu/wp-content/uploads/2011/06/ACEEE-08-06-01-Energy-Information-Feedback-Studies1.pdf>

Figure 2.1 – Household savings broken down by feedback type (Ehrhardt-Martinez et al., 2010)

Hargreaves et al. (2010) argue that feedback is effective, in part, because energy lacks saliency in two important ways. First, electricity and heat are invisible and abstract forces entering a space through hidden wires (and ducts). Second, most energy-use behaviours are part of routines and habits, and thus not easily noticed.

Acknowledging 30 years of research into residential energy feedback—reporting savings between 0 and 20%—Karlin, Au, Goneva, and Zinger (2011) were interested in what feedback factors led to effective energy savings. They conducted a meta-analysis of the 42 home feedback studies meeting their criteria (out of 104 empirical studies conducted between 1976 and 2010), focusing on effect sizes. Feedback in general was found to have a highly significant ($p < .001$) effect on energy savings, accounting for 4.8% of the variance in energy use.

Karlin's team ran a moderator analysis, looking at the effect of particular presentation variables (both between and within studies). They found feedback presentation variables (display medium, duration, frequency, granularity, comparison, and combination) each had a *significant* effect on energy savings (Karlin et al., 2011).

In another study, summer electricity use (in a group of townhouses) reduced 10.5% through daily consumption feedback; more frequent feedback (combined with goal-setting) achieved 13.0% savings (Lovins, 1992). In this same study, using a simple signal light prompted people to reduce energy use by 15.7%, by indicating for people to open their windows instead of using the air conditioning to cool their home.

Further, Abrahamse et al. (2005) and Hargreaves, Nye, and Burgess (2010) both found evidence that homeowners actually changed their energy-use habits as a result of being exposed to feedback over a certain period, both persisting after feedback ceased.

2.2.3 Habits

When addressing technologies and interventions that promote behaviour change, an understanding of habits is essential because most residential energy is consumed through homeowners' routines and habits (Carlsson-Kanyama & Lindén, 2007; Hargreaves et al., 2010; Lutzenhiser, 1993; Sanders & McCormick, 1993; Sauer et al., 2003). Many routine energy-consuming actions go unnoticed or forgotten (Lutzenhiser, 1993); diary-studies reveal that homeowners are often surprised by how frequently these behaviours occur. Sauer et al. (2003) attribute low energy savings in the use of domestic appliances (vacuum cleaners and kettles) to habits and low motivation.

Verplanken and Wood (2006) define habits as “a form of automaticity in responding that develops as people repeat actions in stable circumstances” (p. 2). Habits are formed through the repetition of initially-planned actions, undertaken to achieve some goal. Over time, people form memory links between the action and the reoccurring situational features that said action is situated in. These features (prior responses in a chain of actions, environmental cues, internal states, or the presence of a people associated with the behaviour) eventually become automatic cues to trigger the action without any planning or intention.

Verplanken and Wood (2006) suggest that information campaigns, especially those that ignore a behaviour's context, have trouble changing behaviours when they do not address previously existing habits. They give two reasons for this:

- Habits create expectations that reduce a person's sensitivity to changes in their context, reduce their need to look for new (and disconfirming) information, and reduce their need to deliberate about the action. Some habits (like addictions) are also instantly rewarding.

- Environmental cues are very effective at triggering well-practiced behaviours, despite a person's intention to do otherwise, and these habits do not require much attention, conscious intention, or control to be done efficiently.

Verplanken and Wood (2006) describe two effective habit-change intervention strategies: 1) downstream-plus-context-change (D+CC) interventions, and 2) upstream interventions. D+CC interventions involve putting (or finding) people in a new context (where old environmental cues are not salient), and then motivating the individual to change the behaviour through education, counselling, or any number of other strategies, like those listed above. An example would be giving free bus passes to new city residents to encourage public transit over driving. Upstream interventions address the behaviour *indirectly* by addressing larger contextual influences on behaviour. These may include economic incentives, legislation, structural changes to environment (read: affordances & constraints(Norman, 2002)), and education (to those—like children—not yet with a strong habit).

With respect to creating new habits, successful interventions must: change the contextual cues triggering old the habit, motivate the person to perform the new action (through incentives and establishing intentions), and ensure the action is repeated in stable circumstances so that new environmental cues can form (Verplanken & Wood, 2006).

2.3 Home Energy Management

One further approach to residential energy conservation is to utilize automation, in the form of home energy management (HEM) technology. HEM technology encapsulates devices designed to monitor and sometimes control residential energy use in the home. Electrical HEM (eHEM) technology refers to those devices and monitors focused on electrical energy, while Programmable Thermostats (PTs) focus on thermal energy.

2.3.1 Electrical HEM

eHEM monitors provide users with energy-consumption feedback on the whole-home or appliance level of granularity (Karlin et al., 2011). Recent work by Hargreaves et al (2010) looked qualitatively at the impact of three kinds of eHEM devices on households in the UK. They were interested in how applying feedback leads to energy savings.

Hargreaves et al (2010) found two primary ways that the devices changed behaviour in most participants, reducing their energy use. First and most commonly, participants would “use it hot,” reacting in the moment to energy feedback by turning appliances off when consumption went above a perceived baseline value (i.e., energy used when “you’re doing nothing”) (Hargreaves et al., 2010). The monitors made energy use salient and thus prompted behavioural responses. Second, people used the monitors to identify ‘greedy’ devices, and then made “considered choices” to use them differently or replace them (Hargreaves et al., 2010).

Hargreaves et al (2010) also identified four ways in which behavioural changes were limited. First, many people saw their ‘greedy’ devices as essential and simply felt they had no control over these devices’ consumption. The visibility of this consumption often caused anxiety, exacerbating the problem. Second, people justified their high consumption as reasonable in service of providing a warm, comfortable, and well-lit home. As one participant reported: “life is for the living” (p 6117). Third, participants reported numerous instances of tensions and disputes between household members. Examples include: Children not complying with parent’s requests to turn devices off; lights turned off by one member and back on by the other; people using the monitors to spy on family members’ behaviours and use that information to feed their arguments; etc. Finally, in response the feedback, many people complained of a lack of external resources or support in terms of recommended behaviour changes or purchasing decisions.

Participants reported that absolute measures of consumption (those aggregating use across a household) were unhelpful because amounts were too low to motivate action

on short time scales, and too high to feel realistic when extrapolated to large time scales (Hargreaves et al., 2010). They found units like kWh and CO₂ too abstract to be useful. However participants were very interested in comparisons between specific devices' consumption, which made sense in a way that scientific units could not. Participants were also interested in graphical or symbolic feedback (ie. fuel-tank and cross or check-mark with respect to the daily target).

Hargreaves et al (2010) presented a few other findings of interest. A novelty effect was seen, that after initially high interest for a few weeks, use would drop significantly. Feedback was also a source of stress, for example if the person felt conflicted between warmth and not wasting money from their heater. The relationship between aesthetics & location in the home seemed more vital to their use than each monitor's level of provided feedback.

2.3.2 Programmable (and Manual) Thermostats

Programmable thermostats (PTs) are those that allow users to set an automated schedule of different temperature settings over different times of the day and through a standard week (McCalley & Midden, 2004; Peffer, Pritoni, Meier, Aragon, & Perry, 2011). By programming temperatures that are lower in winter (and higher in summer) while residents are away or sleeping, household energy consumption can be reduced in comparison to maintaining a constant temperature. Unfortunately, most PTs are too difficult for users to operate and understand to be effective at reducing energy consumption (Combe, Harrison, Craig, & Young, 2012; Meier et al., 2011; Rathouse & Young, 2004). The following review looks at this issue in depth.

Usability Studies

Combe, Harrison, Craig, and Young (2012) tested the usability of three PTs with a groups of younger (24-44) and older (62-75) people, finding discouraging results; not one of the older users could finish the experimental tasks, and only 35% of the younger users did so with the most usable PT, whose mean successful task time was more than 7 minutes.

Meier, Aragon, Pepper, Perry, and Pritoni (2011) conducted a series of five investigations—both qualitative and quantitative—into the usability of residential PTs. Of particular interest, they ran a usability study on five state-of-the-art PTs, including the ecobee unit addressed in Section 3.0 of this thesis. In one task, asking users to turn the unit from Off to Heat, 26% of all participants were unable to complete the task, and that number rose to 50% for the lowest performing PT; time to complete this task for one particular unit varied from 20s to 260s. Some participants were unable to open (or even recognize) the hinged covers of some of the units, which concealed important controls.

On Controls:

Rathouse and Young (2004) found most residents they surveyed were unsatisfied with their PTs' controls. While some were satisfied, many found the buttons too small and controls too complicated. Similarly, Combe et al. (2012) reported their users were frustrated by PT interfaces that had too few buttons, and were distracted or intimidated by interfaces with too many. The PTs' physical position (too high, too low, difficult to reach, etc.) also impacted their use (Rathouse & Young, 2004).

Patterns of Use:

Rathouse and Young (2004) suggest people have difficulty achieving comfortable temperatures because of difficult-to-use controls, variable heat distribution in their homes, and conflicting preferences of householders. Other factors affecting how people use heating controls include the inconvenience of programming their PTs (versus the convenience of manually setting), the perception that hotter or colder temperatures are healthier, a personal norm for conservation, and the “rumbling” sound of their furnace (Black, Stern, & Elworth, 1985; Rathouse & Young, 2004). People surveyed by Karjalainen (2007) often admitted to not knowing how to use a thermostat to achieve a desired affect. As a result of this and for a fast effect (holding “Valve” theory), they typically chose minimum or maximum settings.

Many people surveyed by Rathouse and Young (2004) avoided programming their thermostats because they perceive it to be tricky or to be more efficient to run their

homes at a constant temperature. Building occupants tend to use PTs manually by adjusting them often or using them as an on-off switch (Lutzenhiser, 1993; Parker et al., 1996). Meier et al. (2011) found most occupants (90%) neglected the programming features and used their PTs manually. In a survey where participants photographed their PTs, 20% of them displayed the wrong time³, while above 50% were set to “hold” (manual override).

Rathouse and Young (2004) found some people did not know what temperature their PT was set to. Meier et al. (2011) found many people kept their PTs set to a constant temperature; this suggests people may not recognize a need for different temperatures in different situations .

Similar to eHEM, many investigations found frequent disagreements within households on temperature selection (Carlsson-Kanyama & Lindén, 2007; Parker et al., 1996; Peffer et al., 2011; Rathouse & Young, 2004). In a study of first-time PT users McCalley and Midden (2004) found that when one group was reminded to consider the preferences of other household members, they would make less-energy-efficient settings on their PTs than those not given the reminder.

Rathouse and Young (2004) found people are more likely to over-heat than under-heat their homes in the winter because manually intervening on their PT is more readily triggered by low-temperature discomfort; people are more likely to open windows when over-heating than to adjust the controls, and then leave them open too long either due to forgetting or to get ‘fresh’ air with the heat.

On motivations and attitudes:

Rathouse and Young (2004) found comfort was the main consideration determining people’s PT use (more so than cost). Some people did *not* consider cost because they prioritized heating as essential, they saw it as affordable, or they did not know how to cut their bills. Some people were motivated to conserve for fear that “the supply will run out,” while others (especially older participants) simply preferred to avoid waste.

³ the correct time is required for a PT program to work correctly

The researchers found little-to-no environmental concern among participants.

On Instructions:

Many people liked having instructions on their PTs, but were also confused by the meaning of symbols (Rathouse & Young, 2004). People also find written instruction manuals are too complicated and do not provide them with enough support (Combe et al., 2012).

2.4 Moving Forward

In this section, I reviewed the literature surrounding the management of energy consumption in the home. I began by outlining some of the challenges faced in this field like ineffective systems, the disconnection between attitudes and behaviours, and inaccurate mental models. I then reviewed some existing curtailment strategies, with particular focus on energy feedback and habit-change interventions. I finished by focusing on the research on electrical and thermal HEM technology, with emphasis on PTs.

From this review, it seems people are generally unable to effectively manage their own energy use, and many existing strategies and technologies fall short of adequately helping home occupants reduce their energy consumption. These challenges and other findings of this review present opportunities for designers and researchers.

Most striking are the accounts of poor usability and user-comprehension of HEM devices (especially PTs). For these technologies to be used properly and to be effective at reducing energy consumption, designers must begin to consider how their devices may be used in a residential context (by users who are unlikely to read instruction manuals). For instance, human-factors principles and design techniques can be implemented to improve the usability of these devices. One particular unexplored opportunity to design for use by heterogeneous groups (i.e. families) rather than individuals.

I also found that people have variable and inaccurate understandings of how to effectively manage their energy use. In particular, I saw a poor conception of how to choose energy-efficient or comfortable temperatures; this topic is a primary theme in this thesis. There

might be opportunities to correct these inaccuracies through design. Research into the effect of different mental models on the choices of individual users is also warranted.

Designers may be able to integrate the aforementioned curtailment strategies to influence users to make more energy-efficient HEM choices. Social norms could be integrated to HEM displays, perhaps in a manner similar to those used by Shultz et al. (2007). Prompts may be ideal for these interfaces as well since they could be presented in close spatial and temporal proximity to energy use choices. Designers may also choose to consider convenience as strategy to increase usability and thus performance. Considering habit-change, a novel-enough interface may present a new behaviour context in which to promote the formation of habits that are more energy-efficient.

From Karlin et al.'s (2011) moderator analysis, there are ample opportunities to study electrical feedback presentation variables (display medium, duration, frequency, granularity, comparison, and combination) independently. I have also found no mention of thermal feedback in the PT literature; this presents a chance to investigate if such information could effect residents energy consumption or the even the use of their PTs.

Some noteworthy gaps and weaknesses in the literature deserve mention and are in need of further research:

- There is still a lack of scientific consensus on whether injunctive or descriptive social norms are more effective at motivating behaviour, leaving practitioners without guidance on which to employ.
- As will be discussed in Section 6.3, there is evidence to suggest that, at least in some cases, the success of social norms can be attributed to the cognitive anchoring effect and not to their social meaningfulness (Eyssel et al., 2006; Lombardi & Choplin, 2010). Researchers should include checks for this possibility when investigating social norm interventions.
- While addressing convenience is a compelling strategy, there exists little empirical research on how convenience impacts behaviour change.
- Karlin et al.'s (2011) moderator analysis could not completely identify the independent effectiveness of all presentation variables. For instance computer-

based display medium gave the highest effect size of display mediums, however, only two studies used computer presentation, and both were at appliance-level granularity.

This literature was the foundation for the remaining investigations, documented herein. The following section outlines a usability study of a current HEM device.

3.0 Observational Review of a Current HEM Product

After gathering this knowledge from the above literature, I decided to carry out an investigation on a piece of current HEM technology. I designed and performed a usability study of one of the current state-of-the-art HEM devices (at the time).

The following is an abridged copy⁴ of the test-plan and report (Stein, Borg, Ratto, & Jamieson, 2011; Stein, 2011) for a usability study conducted as part of a FedDev project between ecobee Inc. and Prof. Matt Ratto of the ThingTank Lab at the University of Toronto (UofT). The usability study was conducted by myself, with assistance from Master-of-Information student Mike Borg (Faculty of Information) in late October and early November 2011.

3.1 Purpose

The purpose of this study was twofold. First, we sought to evaluate the ease-of-use of a state-of-the-art in home energy-feedback technology, namely ecobee's Smart Thermostat and accessories. The evaluation was formative in nature, aimed at finding the interface's positive and negative aspects and identifying improvements. Second, we wanted to better understand users' thoughts, strategies, and needs when interacting with this kind of system.

3.2 Methods

The methods were based on recommendations from Jakob Nielsen's *Usability Engineering* (Nielsen, 1994, Chapter 6). Each participant was put in front of the ecobee Smart-Thermostat (ST), and asked to perform tasks, while being video recorded.

The study was conducted at the basement level of 376 Bathurst St, in proximity to the ThingTank Lab. Each session lasted approximately one hour per participant, including intake (10min), task observation (40 min), and debrief (10 min). One experimenter

⁴ I excluded redundant sections and added contextual content to fit this into my thesis.

interacted with the participant while the other managed recording of video and performance measures and kept time.

The ST was running “ecobee Smart” version 2.4, which was ecobee’s most current software version for the unit; it gave users the ability to control and program their central heating and air conditioning system, set preferences, monitor various sensors, and control the use of smart plugs. Before each test session, *Preferences and Program* was reset, but *Registration* was not altered. The smart plugs remained installed a) to maintain a use history and b) because they are difficult to reset. The system initially was set to *Off* mode (as opposed to *Heat*, *Cool*, or *Auto*) and the user started from the home screen.

Participants

We recruited four participants through the ThinkTank Consortium, via email invitation (see Appendix A). They were selected to be representative of current ecobee users. ecobee’s residential demographics data were analyzed to determine an acceptable range, often representing about two thirds of their current users in each demographic category. Some allowances were made in our criteria because a) we were equally interested in potential users of this technology as in current ones, and b) excessive constraint would limit our ability to recruit enough participants. To be representative of potential users, participants must:

- currently be responsible for their home electrical and heating bills
- have control over their home’s thermostat
- be familiar with touch screen technology
- *not* be HVAC professionals or usability experts

Further, to be representative of current ecobee users, ideal participants should:

- be between 30-55 years old
- live in a household with 2-4 people (themselves included)
- have a total household income exceeding \$100,000
- typically use the internet at least 2 hours each day

Ideally, all users would have never seen or interacted with an ecobee Thermostat. However, since insights from current users would be valuable in exposing expert strategies and workarounds, we chose to recruit an ecobee employee to serve as an 'expert' user, for comparison.

Participant Intake

Before beginning the tasks, each participant was briefed on the test and asked to sign an informed consent (Appendix B)⁵.

We then explained how the test would run, including the instruction to think aloud while performing tasks. Participants were encouraged to ask questions to highlight what they found unclear about the interface; however we usually did not answer these questions because we wanted to know if the system could be used without help. Finally, participants had a chance to ask any questions before they started.

Because the system was novel, sections of ecobee's "Smart Thermostat: Quick Start Guide" ("Smart Thermostat Quick Start Guide," 2008) were made available, namely the "Using the touch screen" and "Feature Buttons" panes (see Appendix C). No other aides were made available.

Tasks

Each task was read from a script and presented one at a time, to ensure everyone got the same instructions. The list of tasks and their exact wording is found below in a simplified format (Table 3-1) and in full (as communicated to participants) in Appendix D.

⁵ The study protocol was reviewed by UofT ethics.

Table 3-1 – Simplified list of usability tasks

Task	Task Description
1	Report indoor temperature.
2	Report outdoor temperature
2a	Return to home screen
3	Turn furnace on.
4	Create new program, using typical workday, choosing preferred temperatures when prompted.
5a	Edit the program to arrive home late on two days.
5b	Edit the program to add a lunch at home on two other days.
6	Alter current setpoint to achieve comfort when it is too cold.
6a	Return to preprogrammed setpoint
7	Enter Quick-Save mode
7b	Interpret
8	Install 2nd temperature sensor
9	Check 2nd temperature sensor feedback
10x	Find Smart Plug Dashboard
10	Turn smart plug off
10a	Turn smart plug back on
11	Set smart-plug program so it is on only when you are at home
12	Compare and interpret usage of two plugs
13	Find hourly electrical use data
13a	Interpret

The experimenter refrained from helping participants, unless a) they were clearly stuck and getting frustrated, or b) the problem was clearly not novel. We also refrained from giving positive or negative feedback

For the purpose of timing, each task was deemed finished when the participant completed the task and returned to the home screen, unless otherwise indicated by the instructions. However, a task could have been ended prematurely if the participant was having excessive difficulty completing it or if it had become too unpleasant.

Debrief

Once they had either finished all tasks or run out of the allotted time for the study, each participant was debriefed. First, they filled out a demographics questionnaire (Appendix E). Then they were asked for comments on the system and suggestions for system improvement.

Data Collection

The screen was video-recorded to capture exactly how the participant interacted with the interface and what they said as they did so.

To comprehend what the participants were thinking, they were asked to *think aloud*: vocalize what they were thinking as they performed each task. Time to complete each task, the number of errors, and specific qualitative issues were recorded post-hoc through video analysis.

3.3 Participant Demographics

Table 3.1 outlines participant demographics. Income was omitted as most participants chose not to share. Note that participant 2 is not directly responsible for his energy bills because his wife manages these items. No participants are usability or HVAC experts.

Table 3-2 – Participant Details

Test User	Age	Job	Household Size	Hours Online	Pays Heating Bills?	Control over T-stat?	Background with ecobee ST?	Describe Current T-Stat
1	31	Student	2	10	Yes	Yes	No	digital, non-program, simple display with current & set temp
2	50	Software Designer	2	5	No	Yes	No	digital, programable, straight-line dial
3	37	Partner/ Design Firm	2	2	Yes	Yes	Heard of it but never seen.	digital, programable, numerical interface
4 (expert)	46	VP Marketing	4	5	Yes	Yes	Yes	ecobee ST

3.4 Results

3.4.1 Quantitative Results by Task

We recorded time-to-complete and the number of user errors for each task⁶. See Table 3-1 for the list of tasks, and Appendix F for a complete table of results. Figures 3.1 and 3.2 summarize and compare the results of the novice average against the expert user. Negative red values are tasks not recorded for the expert. Error bars represent minimum and maximum for novices. Some tasks were only recorded or performed by one novice (no error bars – tasks 2a, 6, 11, & 13).

⁶ Due to video-recording errors and testing time constraints, not all tasks were recorded for each participant, limiting the amount of quantitative data.

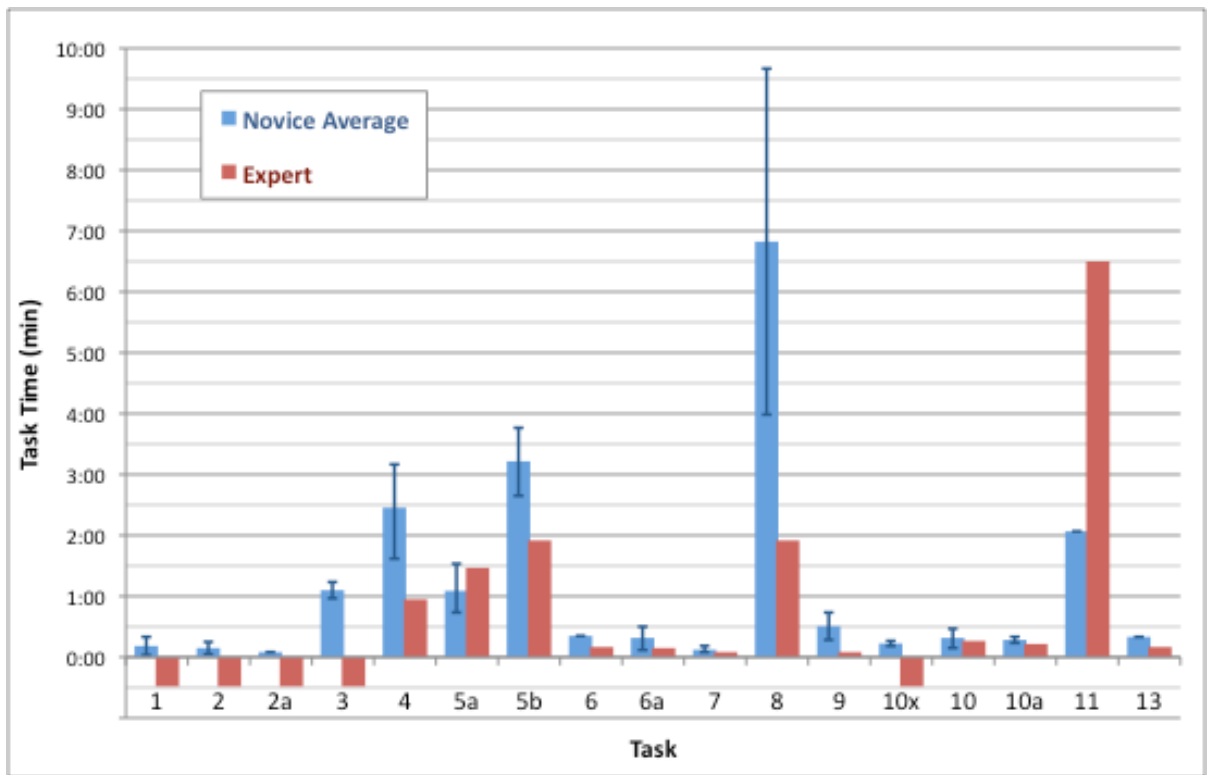


Figure 3.1 – Time to complete each task

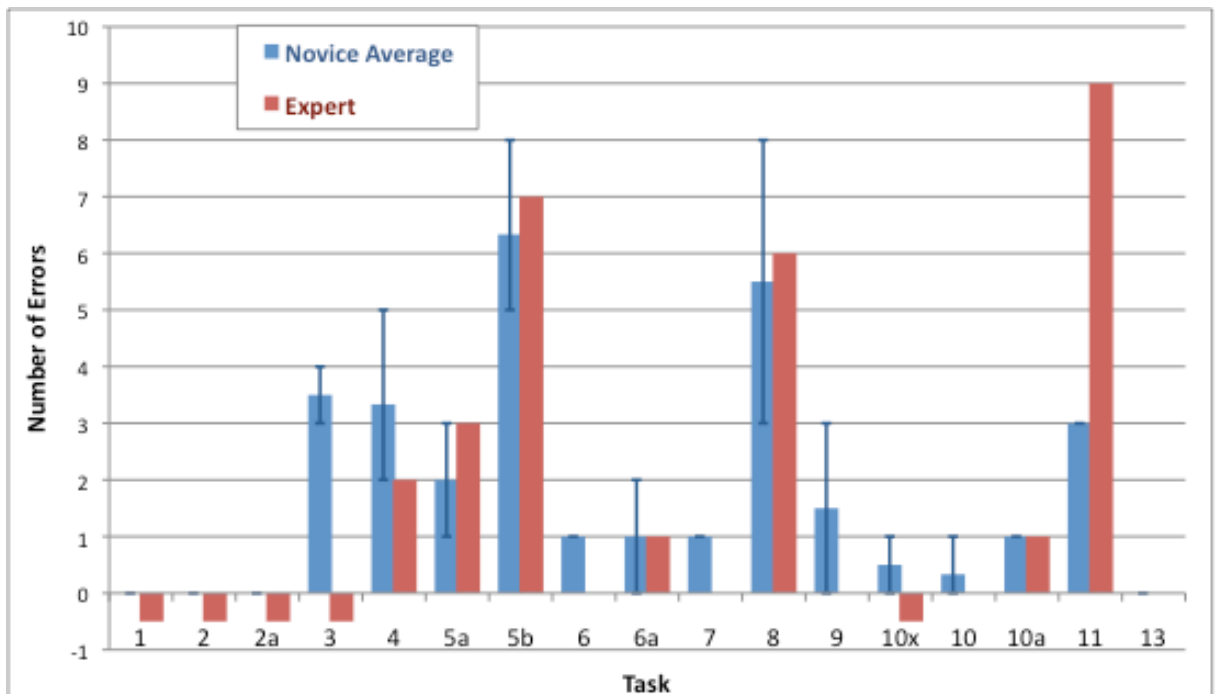


Figure 3.2 – Number of user-errors, by task

These data suggest a few things about how participants performed on particular tasks. Certain tasks were performed with high proficiency. Tasks 1, 2, 2a, 6, 7, 10x, 10, 10a, and 13 were all done quickly and with one or fewer errors. Second, note the expert was able to do most tasks faster than the novices, and performed tasks 3, 4, 6, 7, 9, and 10 with fewer errors, suggesting some learnability on those tasks (see further detail below).

These figures also highlight—based on long task time and/or high number of errors—that novices seemed to find tasks 4 (program Wizard), 5b (edit program with lunch item), 8 (install temperature sensor), and 11 (set smart-plug program) especially difficult, suggesting room for improvement in these areas.

Participant Comparison

To compare both between the novice participants and against the expert, scores were averaged across the tasks that all participants completed⁷ (Table 3-2). Most notable, is that while expert's task time is lower than the novices, he still made about the same number of errors.

Table 3-3 – Average participant task-time and errors across tasks 4, 5a, 5b, 6a, 7, and 10

Test User	Average Task Time	Average Errors
1	1:23	2.5
2	1:11	2.0
3	1:10	2.5
4 (expert)	0:48	2.2

3.4.2 Qualitative Results by Issue

Each usability issue was recorded and scored for severity (1: minor, 4: critical), expected frequency (1: rare, 4: frequent), and occurrence (fraction of participants effected); these were integrated into what I am calling a “priority score” (to denote the level of priority that developers should assign to addressing each issue) for each issue (see Appendix G). We found 69 distinct issues and rated each with a priority score

⁷ Task 8 was omitted for one participant because time was running out, in favour of smart-plug tasks

(product of the three factors). The highest possible score is 16, and the lowest is 0.3⁸. Each issue listed below is given a priority score in parenthesis. See Appendix H for a complete list of all issues and their scores.

3.4.2.1 Overall

Top Issue: Unresponsive Screen (12)

All participants had difficulty due to a lack of screen responsiveness, which was most apparent in making selections near the screen edges and on scrollable lists. This caused much frustration in all participants. In response, users tended to press harder, which may eventually lead to screen damage. The issue was not confined to the test unit.

According to a conversation with interface-design experts on the 3rd floor of the building⁹ (who had the ST installed as part of our project), this may be due to: 1) a sensitivity issue (the system may interpret a touch gesture to be a scroll, due to small finger movements), 2) icons too small for the current screen resolution.

Solution: Adjust the software's touch sensitivity to better distinguish touch from scroll gestures. Increase the size of buttons and move them further from the edge.

2nd Top Issue: "Hold" Ambiguity (12)

None of the novice participants understood what "Hold" means, when they moved the temperature slider. One thought it might mean his desired action is "on hold" until some unknown time in the future.

Solution: Find a way to explicitly indicate what "hold" means, or consider different wording to describe this function.

⁸ Note the exponential nature of the priority score. If each factor is halfway (2/4 severity, 2/4 frequency, and 50% occurrence), the priority score is 2.

⁹ Of Normative Design - <http://normativedesign.com>

3rd Top Issue: “Ok” in Quick-Save Mode (10.5)

Novice participants *all* pressed “Ok” after entering Quick-Save (QS) mode, unknowingly disabling the feature. This effectively renders QS mode useless.

Solution: Instead of “Ok”, the button could say something like “Disable Quick-Save”. Alternatively, find a way to maintain QS mode, without remaining in the QS screen, like an on/off function.

3.4.2.2 By Category

I sorted the issues by affiliation (i.e. into groups of similar issues) into five categories that I generated: Confusion, System Errors, User Errors, Action Difficulty, and System Quirks & Pitfalls.

Category: Confusion

We found 21 issues distinguished by confusion or a lack of understanding on the user’s part.

Top Issue: “Hold” Ambiguity (12)

- **see 3.4.2.1**

2nd Top Issue: Quick-Save Confusion (9)

Novice users were confused when interpreting the Quick-Save screen. User comments included: *“I’m not sure where 2.2° came from”; “I don’t know what ‘set point’ is, and I don’t know what ‘set-back’ means”; “If the set point is 15°, and the current temperature is 22°, then it should be set back by 7°, not 2.2°”; “I would not be using that feature”,*

Solution: A more effective dialogue could more clearly explain how QS works and why it is important, maybe using imagery, or animating the slider moving down. Also, for simplicity, 2.2° (4° F) could be rounded down to 2°.

3rd Top Issue: Current vs Set Temperature (8)

Two out of three novices confused the current temperature with the set point. This impacts general understanding and inhibits the user's ability to use the system effectively.

Solution: Show that set point is the desired future setting, which the current temperature is moving towards.

Category: System Errors

We found 8 items distinguished as system faults or bugs.

Top Issue: Unresponsive Screen (12)

- *see 3.4.2.1*

2nd Top Issue: "Add New Item" Edits List Element (6)

On the Edit screen, pressing "Add New Item" accidentally activated the bottom-most list item (usually "*I go to sleep at ...*") to be edited. One user unknowingly changed his program to sleep at noon as a result.

Solution: Have developers investigate and fix this software bug

3rd Top Issue: Erratic Temperature Slider (5.3)

Two of three novices experienced problems with the temperature slider erratically jumping around while being adjusted, making precise selection difficult.

Solution: This may be related to screen responsiveness. Fix this bug, or make the tap-adjust function more visible and intuitive.

Category: User Errors

We found 10 items distinguished as mistakes on the user's part.

Top Issue: “Ok” in Quick-Save Mode (10.5)

- *see 3.4.2.1*

2nd Top Issue: Not Clear to Press “Resume” (8)

When asked to return the temperature back to its original setting, two out of three novices moved the slider back, instead of pressing resume.

Solution: Make clearer what “hold” does (see 3.4.2.1).

3rd Top Issue: Pressed Physical Button → Lost Changes (7.9)

Three participants (expert included) pressed the physical button to complete their actions, unknowingly losing their changes. Sometimes happened after the “Done” button failed to respond (screen-responsiveness issue).

Solution: Have an alert asking “Are you sure you want to exit? Doing so will cancel your changes.” Also, fix the responsiveness issue with “Done” (see 3.4.2.1)

Category: Action Difficulty

We found 16 items distinguished by difficulty in the user’s attempt to achieve some goal.

Top Issue: Too Many Steps to Turn Plug On/Off (6)

In the Smart-Plug Dashboard, both novices in this situation complained that turning a plug on or off took too many steps to complete (ie. press the plug icon, select on or off, set a hold duration, press done). One participant commented: *“It’s a two-step, you have to go “turn device on,” then “done”, when in fact that should be a toggle.... Why do you have two steps?... that’s kind of weird.”*

Solution: Number of steps could be reduced by one or two. Consider a simple toggle on/off functionality from the Smart-Plug Dashboard.

2nd Top Issue: Attempted Interacting with Text (4)

All participants (expert included) tried to interact with plain text at some point during the test, thinking it would produce more information.

Solution: Either make more text interactive, or make appear less interactive. Anything that is interactive should look different than plain text.

3rd Top Issue: Sensor Screen: “Configure” Button Says To Go Elsewhere (3)

To install the temperature sensor, after mistakenly going to the sensor screen, both participants in this situation hit “Configure” only for it to tell them to go to the installation settings. It is also not clear that this function is meant to designate a control sensor.

Solution: Distinguish between configuring a new sensor and configuring a control sensor from the Sensor screen. Provide a direct shortcut instead of telling them to go somewhere.

Category: System Quirks and Pitfalls

We found 14 items that did not fit into the above categories, but seemed to be distinguished as potential system risks, shortcomings, or annoyances.

Top Issue: No Reference/Comparison Point for Plug Historical Graphs (9)

The historical energy use graphs for the smart-plugs were difficult to interpret due to a lack of point of reference or comparison. One participant commented “this information wouldn’t be all that useful to me; maybe if I knew in total what my kW/h [sic] should be ... [or] a point of comparison.”

Solution: Put the data into context: Include the ability to visually compare devices side-by-side in a chart. Include a reference or a prompt to show efficient or average use of the device.

2nd Top Issue: “Consumption” has no units, in Smart-Plug Dashboard (9)

In the smart-plug dashboard, both novices in this situation were not sure what “Consumption” meant, and noticed the number had not units.

Solution: Add kWh to the number and give context (is it hourly, daily, etc?)

3rd Top Issue: Time-Wheel Defaults to 12 AM (6)

In the Edit screen, when adding a new item, the time-wheel starts at 12 am. When noticed, it caused annoyance (to have to scroll through more numbers). When it was not noticed, it led to errors in the program.

Solution: Set the default to 12 pm.

3.4.2.3 Specific to Functionality of “Accessory” Units

Of the 69 issues found, 26 items were directly related to accessories units (temperature sensor and smart plugs), specifically the temperature sensor and the smart plugs.

Top Issue: No Reference/Comparison Point for Plug Historical Graphs (9)

- *see 3.4.2.2*

2nd Top Issue: “Consumption” has no units, in Smart-Plug Dashboard (9)

- *see 3.4.2.2*

3rd Top Issue: Resume Avoided to Turn Plug Off (7.5)

Neither novice realized that they could turn a plug off by pressing “Resume”. Instead, both chose indefinite hold, essentially disabling the program. Related to the top issue in *Action Difficulty*.

Solution: If a device has been turned on, putting the program in hold, then there should be only one way to turn it off; do not prompt “for how long?”.

3.5 Discussion

3.5.1 Favoured Attributes

Participants liked having a touch screen interface. They all especially liked the temperature-control slider, commenting that it was very intuitive. Also, in the post-test questionnaire, all participants indicated that (ignoring cost) they all preferred the ecobee Smart Thermostat to their current programmable thermostats.

Participant Two liked the Edit Screen because it showed all his program settings in one place. Participant three thought it was “fun” to be able to turn things on and off remotely. Participant four (expert user) especially appreciated the ability to “go in and set my preferences”.

3.5.2 Least Favoured Attributes

All participants were disappointed with the poor screen responsiveness. Two of them commented they would probably use the web-portal or the mobile app instead.

One participant said the system is inconsistent with itself and going through the menus was confusing. Another commented the screen does not look crisp and that it seems small. One said that the fire symbol was unsettling “Is my fire alarm going off?” Another commented that the lack of brand reinforcement (i.e. not having an ecobee logo on the screen) made the software look generic.

3.5.3 Participant Ideas and Desired Functionality

Some participants commented they would like more system information from the home screen (like power use etc). One participant wanted to see his temperature settings from the Program screen diagram. Another wanted more clarity that the furnace was on or off, and easier control. After getting lost in the menu structure, one user suggested a breadcrumb or tab structure. Finding the sensor installation difficult, he suggested a “6-step Wizard” to guide him through the technical choices.

Some participants were interested in home-automation, while others wanted more control: one suggested a light sensor that would turn off lights when the sun shines, and a motion sensor that would turn off lights when nobody is home. Another participant said: *“a good interface design allows you to more easily control things, not ... take control away from you.”*

3.5.4 Other User Comments

One user said: *“It looks like it was designed by an engineer,”* referring to a lack of user-centred design. Participant two commented that confirmation is key, but *“thermostats, in general, don’t indicate if an action is having an effect.”* Our expert user said that personalization is very important to him.

Some participants wanted to save money, while others wanted to learn about their energy use. When asked why they chose a certain temperature setting, participants were usually unable to justify their choice beyond *“that’s what my home is set to.”* This indicates users may be able to be influenced to settings that are more energy-efficient¹⁰. One participant demonstrated a moral licensing effect: he thinks his energy-efficient home justifies not worrying about heating.

¹⁰ This issue came up again in the critical making session and served as the chosen design space for my thesis.

3.5.5 Expert Workarounds

The expert demonstrated three workarounds, which indicate opportunity for better system functionality. First, he used his fingernail when the screen was unresponsive (1st System-Errors issue). Second, he used the tap-adjust gesture to alter the temperature set point, maybe pointing to the erratic slider movement (3rd System-Errors issue). Third, he went into his preferences and set the hold duration to “Ask Me,” suggesting the current hold default is not ideal.

3.6 Recommendations

As shown here, there are many opportunities to improve the ST’s usability. We recommended to ecobee to consider the suggested solutions for the each of the issues listed in Section 3.4.2, as well as to review the complete list of issues in Appendix H. More generally, a ubiquitous Help feature would be immensely useful, for when users do not understand what something means or what to do.

3.7 Moving Forward

In this section, I described the usability study that I conducted on the ecobee Smart Thermostat in the fall of 2011. The purpose was to 1) evaluate the technology’s ease-of-use, and 2) better understand users’ experience with state-of-the-art HEM technology. I described and justified the methods, then described the participants. This followed with an overview and interpretation of the quantitative results of task time and number of errors. I then described the top qualitative results, both overall, and in each of five categories (Confusion, System Errors, User Errors, Action Difficulty, and System Quirks and Pitfalls). I finished with the notable participant comments, three expert workarounds, and some recommendations to ecobee.

This study demonstrates first-hand many of the challenges related to HEM and in particular the design of PTs. It highlight’s people’s difficulties and misunderstandings with such tasks and technology. Consistent with the findings of Meier et al (2011) on this same device (Section 2.3.2), this study suggests that even the most state-of-the-art in HEM technology has substantial usability issues, and is thus still insufficient in terms of the user

experience. There is reason to believe that if HEM devices are not usable, then homeowners will simply neglect their energy-saving features; fortunately for ecobee, users expressed interest in interacting with the system through alternative mediums like the web portal. Additionally, the issue of users' ambiguous understanding of comfortable and efficient temperature selection arose in their inability to justify chosen settings; this topic will be explored more in depth later. These findings were also used to inspire the design of the 'Critical Making' session, described in the following section.

4.0 The Critical Making Workshop

The usability study inspired the content of the following stage of this investigation. This section briefly describes the development and execution of a “Critical Making” session, a central component of the FedDev project already described. It was conducted by Prof. Matt Ratto, Mike Borg, and myself on April 21, 2012. It allowed us to explore HEM firsthand through discussions with workshop participants, and use that to scope out potential design spaces.

4.1 Critical Making

Critical making (CM) is an interactive mode of engagement that utilizes and combines physical making with critical thinking (*Ignite Toronto 2*, 2009; Ratto, 2011). Unlike participatory design (Nielsen, 1994), the goal of CM is not to produce a usable prototype, but rather to highlight viscerally-implicit concepts and stimulate theoretical discussion. By physically engaging with some material elements related to a particular theoretical domain, CM participants can uncover and verbalize their tacit understanding and reaction to said domain, adding to the depth of critical thinking available in group discussion.

In a CM workshop, which resembles a focus group, participants are first introduced to the necessary background theory on the session’s chosen topic and given a tutorial on the technical ‘making’ aspects of the workshop. Discussions are facilitated before, during, and after ‘making’ around the workshop’s topic. Participants and experimenters leave with an enriched understanding of that specific field of interest.

This particular workshop was the first attempt at using CM in engaging and supporting an industrial client (ecobee Inc.).

4.2 Development of Our Session

We integrated the findings from the technical literature (Section 2.0), the usability study (Section 3.0), and some social-sciences research on human approaches to decision making

(not covered here) to inform the development a critical-making workshop on the topic of “Energy Monitoring and Data Visualization”.

4.2.1 Details of our Critical-Making Workshop Idea

We designed our CM session to engage users on the topics of HEM and home automation. We prepared cardboard ‘dollhouses’ equipped with miniature appliances (furnace, oven, washer/drier, two lights, entertainment system, and a dishwasher) outfitted with LEDs. The appliances were connected to the wireless network through an Arduino microcontroller. A software interface was developed to communicate with the appliances and allow users to create a schedule of appliance-use during a simulated 24-hour period.

We grouped participants into ‘family’ units (with adults and children) where each person assumed a persona of a family member. The members of these family units were asked to cooperate to decide on and program schedules for the use of their major appliances and heating system. We then provided them with different scenarios for a typical day in the their characters’ lives. Each scenario would occur over a single day for typical April weather in Toronto.

4.3 Session Overview and Highlights

We had 11 male participants, recruited mostly from the ThingTank consortium, including those from academic and the business domains.

We started by giving them a basic introduction to important technical knowledge that would help them understand the domain, including basic concepts in understanding the difference between energy and power, and how it related to the use of heat and consumer appliances in the home. This was followed by a tutorial on how to use the Arduino and software interface.

Participants broke into four family-units, of 2-3 members each, and each participant fabricated and then assumed a role of one of those family members. By filling out a questionnaire (see Appendix I) each had to specify their character’s preferences for

system-automation versus human-agency (by marking an X on a line between the two extremes), their character's preferred decision-making style, between rational and socially influenced, and a few metrics about their views and preferences on comfort and energy efficiency.

We administered two scenarios (S1: a typical weekday, and S5: large family dinner with guests), wherein each family unit collectively programmed their dollhouse by setting a schedule for the use of their home's major appliances and heating system. For each scenario, we recorded energy consumption for each household over the virtual day, and facilitated a group discussion over the results from each household.

In the second part of the session, each family unit was asked to mock up paper prototypes of what a new interface for the home's energy-management system could look like.

4.3.1 Qualitative Session Highlights

Some of the qualitative results from participant discussions include the following:

- Many participants seemed to lack understanding of their how furnaces operated. Even when appearing to have the more-accurate Feedback model of home heat control (see Section 2.1.2), some participants falsely believed that the energy savings from lowering temperatures at night or while away would be lost when the temperature is set back to normal levels.
- We saw some examples of family interactions that could affect energy management. For instance, there was a lot of negotiation and compromise between household members on decisions. One group commented "grandpa sets the temperature because he pays the bills." Another participant commented, "he sets it down and I set it back up".
- In the family dinner scenario, some participants were hesitant to set temperatures too low, out of a motivation to accommodate houseguests. Another commented that they do not want to be managing their energy use while they have company.

- In their design ideation, one group commented that people do not use appliances in isolation, but as part of activities, and the interface should reflect this.
- In terms of temperature settings, many participants remarked that there was a perceived conflict between saving energy and being comfortable.
- Again, participants did not seem to understand how to set appropriate temperatures for comfort or efficiency. A related issue was that nobody set their temperatures below 17°C for away-from-home times out of fear that “the pipes would burst”; however they seemed to not understand what temperatures would cause this.

4.3.2 Design Spaces Highlighted from the Critical Making Session

The discussion results were the basis for further identification and development of problem spaces to address through design. Based on my notes and the audio-recorded discussions from the CM session, I identified thirteen potential design spaces that were highlighted. In no particular order:

- Temperature Selection Ambiguity
- Ambiguous understanding of furnace operation
- Activity-based energy feedback
- Self-diagnosing systems
- Accommodating multiple (sometimes conflicting) desires for heating
- Cost vs Social/Environmental Motivations
- Confusing Controls for Refrigerators and Freezers
- Automated fail-safes
- Altered behaviour as a result of ineffective automation
- Time-of-use pricing confusion
- Accommodating the comfort of house guests
- Simple physical controls versus complex online controls
- Customizable interfaces.

4.4 Moving Forward

In this section, I described the Critical Making participant session conducted in the spring of 2012. I described CM as a process in which participants generate and develop their comprehension of a particular subject through physically engaging with related material elements and then discussing their insights. I then described how our particular workshop on HEM was developed, through integrating findings from the literature review and usability study, and how it was implemented. I finished by describing the problem spaces that this investigation highlighted.

This CM investigation gave me another useful perspective on how people understand and relate to their energy use in the home. One useful experience was seeing how heterogeneous groups might negotiate (or fail to) to manage their energy use. I also was intrigued by comments about how houseguests factor into people's decisions. More generally, I found the CM session allowed me to better understand how people relate to the broad range of issues related to HEM.

Similar to the usability study and the literature, I again saw evidence for an ambiguous understanding of temperature selection and home heat control in general.

I narrowed down the list of design space to the five I was most interested in pursuing (see Section 5.1); the next section outlines how I then evaluated each for its readiness for design intervention and chose one to address further.

5.0 Design Spaces and Selection

This section describes the process of taking the design spaces discovered from the CM workshop, identifying and applying selection criteria to them, and choosing one area (Temperature Selection Ambiguity) to explore through design.

5.1 Top Five Design Spaces

From the literature review, usability study, and CM workshop, I identified a list of unexplored design spaces with the potential for exploration. From those, I identified the five I saw as most interesting to explore. In no particular order:

5.1.1 Ambiguous understanding of Furnace Operation

Problem: People do not understand how their H&C systems use energy, and how different situations affect that use (Lutzenhiser, 1993; Rathouse & Young, 2004). This was also seen through the usability study and CM workshop. Incorrect mental models could lead to inefficient settings/behaviours.

5.1.2 Lack of Device Context → Activity-Based Energy Feedback

Problem: One comment from the CM workshop was the idea that people do not use appliances in isolation, but as part of activities. However, existing HEM interventions target mostly appliance-level or whole-home energy consumption (Karlin et al., 2011). Activities could provide a more meaningful context for feedback than isolated appliances.

5.1.3 Accommodating Multiple (sometimes conflicting) desires for Heating

Problem: Thermostats and HEM devices are not designed to handle the variable demands associated with families (or other heterogeneous groups) who often have different individual desires for comfort and convenience, along with hierarchies of decision making (Carlsson-Kanyama & Lindén, 2007; Hargreaves et al., 2010; Parker et al., 1996; Peffer et al., 2011; Rathouse & Young, 2004). I saw examples of household disagreements and negotiation in the CM workshop. For example, recall that having to consider other householders caused first-time PT users to set less energy efficient

temperatures (McCalley & Midden, 2004). Different individuals may also be responsive to different interface modes (i.e. wall device vs web-portal vs mobile-app).

5.1.4 Cost versus Social and Environmental Motivations

Problem: The common perspective on H&C control seems to focus on cost, which seems to not be very effective in promoting energy conservation, especially with low-consumption appliances (Hargreaves et al., 2010) (e.g. an ecobee employee discovered their laptop cost only \$3/year to charge, which was insufficient to motivate change). Current metrics for representing energy (e.g. kWh, BTU, J, tonnes CO₂) also lack meaning (Hargreaves et al., 2010). As we saw from Rathouse and Young (2004) in section 2.3.2, cost is not the primary consideration in home heating control.

5.1.5 Temperature Selection Ambiguity

Problem: As seen in the literature, the usability study and the CM workshop, people do not seem to understand why and how to select an appropriate temperature for comfort or efficiency. For instance, recall that Karjalainen & Vastamaeki (2007) found only 15% of surveyed occupants of Finish homes correctly believed comfortable indoor temperatures should be cooler in the winter than in the summer, and vice versa. Kempton's Valve model (Kempton, 1986) suggests people regard temperature settings as a degree of heat output, rather than static conditions. Recall in Section 3.5.4 that one Usability study participant chose 22° only because that is what his wife set it to at home. In the CM workshop, I noticed nobody set their temperatures below 17°C for an away-from-home scenario, partly for fear that their pipes would freeze. These seemingly arbitrary choices of temperature may be a missed opportunity for people to choose the most optimal conditions to both be comfortable AND save energy.

5.2 Selection Criteria

Choosing a design space to address required criteria against which to evaluate and compare each space. To this end, I chose the work of Lee (2006), on automation design. He

describes a number of “Automation Pitfalls”, followed by a list of strategies that address these pitfalls. The contents of the following two sections are extracted from this paper¹¹.

5.2.1 Automation Pitfalls (Lee, 2006)

1. **Out-of-the-Loop Unfamiliarity:** “the diminished ability of people to detect automation failures and to resume manual control” (Lee, 2006, p. 1571)
2. **Clumsy Automation:** “when automation makes easy tasks easier and hard tasks harder” (Lee, 2006, p. 1572)
3. **Automation-Induced Errors:** when the introduction of automation brings about new forms of human error, including:
 - a. **Mode errors:** “arise when operators fail to detect the mode or recognize the consequence of mode transitions in complex automation.” (Lee, 2006, p. 1572)
 - b. **Configuration Errors:** mistakes made during “complex configurations or setups” (Lee, 2006, p. 1572)
4. **Inappropriate trust:**
 - a. **Misuse:** failures occurring “when people inadvertently violate critical assumptions and rely on automation inappropriately” (Lee, 2006, p. 1573)
 - b. **Disuse:** (distrust) failures occurring when people reject the automation’s capabilities
 - c. **Complacency:** (overtrust) overreliance on automation “result[ing] from trusting [it] more than is warranted” (Lee, 2006, p. 1573)
5. **Behavioural Adaptation:** “the tendency of operators to adapt to the automation’s new capabilities, particularly to change behaviour so the technology’s potential safety benefits are not realized.” (Lee, 2006, p. 1575)
6. **Inadequate Training and Skill Loss:** when the “introduction of automation leaves the operator without the appropriate skills to accommodate the demands of the job.” (Lee, 2006, p. 1576)

¹¹ Only those pitfalls and strategies interacting with the design spaces are mentioned here.

5.2.2 Automation Strategies (Lee, 2006)

1. **Fitts's List and Function Allocation:** "assess each function to determine if a human or automation would perform it better." (Lee, 2006, p. 1580). Fitts' list is an overview of human versus automation strengths at different information-processing stages (Lee, 2006).
2. **Dynamic Function Allocation:** "addresses need to adjust degree and type of automation according to individual differences, the operator's state, and the system state." (Lee, 2006, p. 1581)
 - **Adaptable Automation:** "the operator can engage or disengage [automation] as needed" (Lee, 2006, p. 1581)
 - **Adaptive Automation:** "automatically adjusts automation level based on operator's performance, the operator's state, or the task situation." (Lee, 2006, p. 1581)
3. **Representation Aiding:** "capitalizes on the power of visual perception to convey complex dynamic relationships." (Lee, 2006, p. 1583)
 - "Integrating automation-related information with traditional displays and combining low-level data into meaningful information are two important ways to enhance feedback without overwhelming the operator." (Lee, 2006, p. 1583)
4. **Multimodal Feedback:** "provides operators with information through haptic, tactile, auditory, and peripheral vision to avoid overwhelming the operator" (Lee, 2006, p. 1583)
5. **Matching Automation to Mental Models:** "Automation designed to perform in a manner consistent with operators' preferences and expectations can make it easier for operators to recognize failures and intervene." (Lee, 2006, p. 1585)

5.3 Evaluation Based on Selection Criteria

I used the automation pitfalls and strategies to evaluate and compare each of the five design spaces for each space's potential to be utilized with a design intervention, in service of selecting a candidate for design and evaluation. The five potential design spaces were crossed with the pitfalls and strategies, to reveal interactions (Table 5.1). In the table, "x" denotes a likely interaction, while "?" denotes a possible one.

Table 5.1 – Automation Pitfalls and Strategies applied to the Design Spaces

Design Spaces	1. Poor Understanding	2. Poor Context	3. Conflict	4. Motivations	5. Temperature
Pitfalls					
1. Out-of-the-Loop Unfamiliarity	x	?	?	x	x
2. Clumsy Automation		x	x	x	?
3. Automation-Induced Errors	x		x		x
3a. Mode Errors	x				x
3b. Configuration Errors	x		x		x
4. Inappropriate Trust	x	x	x	x	x
4a. Misuse	x	?	x		x
4b. Disuse	x	x	x	x	x
4c. Complacency	x	x	x	x	
5. Behavioural Adaptation	x	?	x	x	x
6. Inadequate Training & Skill Loss	x	?	x	x	x
7. Interaction Between Problems	x	x	x	x	x
Strategies					
1. Fitt's List & Function Allocation	x	x			x
2. Dynamic Function Allocation	x		x		x
2a. Adaptable Automation	x		x		x
2b. Adaptive Automation			x		x
4. Representation Aiding	x	x		?	x
5. Multimodal Feedback	?	?	?	x	
6. Matching Automation to Mental Models	?	x	x	x	x

**Design spaces:*

1. Ambiguous Understanding of Furnace Operation
2. Lack of Device Context -> Activity-Based Energy Feedback
3. Accommodating Multiple (sometimes conflicting) desires for Heating
4. Cost vs Social/Environmental Motivations
5. Temperature Setting Ambiguity

Section 5.3.1 outlines how these criteria applied to the chosen design space of Temperature Selection Ambiguity; this space was chosen, in part, because it had the highest number of pitfalls and strategies.

5.3.1 Pitfalls and Strategies Applied to Temperature Selection Ambiguity

Applicable Pitfalls:

- **Out-of-the-loop unfamiliarity:** Because heating and cooling processes are mostly invisible to residents, most of them simply do not know what environmental conditions are most comfortable or energy efficient (Karjalainen, 2007; Rathouse & Young, 2004).
- **Clumsy Automation:** It is easy to set a temperature in the moment, as well as (if programmed properly) to have the H&C system run by a set program. However, it is difficult to choose an optimal temperature for different circumstances or needs.
- **Automation Induced – Mode Errors:** Occur in different modes of a thermostat program, like when a user cannot understand why their temperature has decreased for seemingly no apparent reason (night-time setback mode).
- **Automation Induced – Configuration Errors:** When asked to choose temperatures during PT setup, users are left only to their intuition to choose (potentially inappropriate) set points, thus leaving them at the mercy of a non-optimal program.
- **Inappropriate Trust – Misuse:** Any non-optimal temperature selection action could be considered misuse, as well as overriding the program with a 'hold'. For example, when turning the temperature up to the maximum because the user holds the 'Valve' model of thermostat operation.
- **Inappropriate Trust – Disuse:** Many programmable thermostats are left in 'hold' mode or simply turned off, instead of relying on the program (Meier et al., 2011). While turning the PT off may actually be more energy efficient in some cases, it represents a failure of the automation in that it requires more effort on the user's part.

- **Inappropriate Trust – Complacency:** Some people (including one of my Usability participants) think just buying a PT will save them energy, and thus do not go through the process of actually programming it or setting the correct time (Meier et al., 2011). Similarly, they may avoid setting different temperatures in a program (ie. choosing 22° for all times) because they think the PT will automatically save energy.
- **Behavioural Adaptation:** Users may show conservative behaviour if they think the system will not do what they desire. They may leave the temperature high while away so their pipes do not freeze. Users may overuse the 'hold' feature (overriding the energy-saving program) to set their own temperatures that they may see as no different than the energy-saving temperatures.
- **Inadequate Training & Skill Loss:** Residents receive no formal training in thermostat operation nor are they given much information about what temperatures are appropriate, comfortable, and/or efficient in different scenarios.

Applicable Strategies:

- **Fitts's List & Function Allocation:**
 - Allocate temperature selection to the automation, or have it recommend a temperature
 - Make temperature selection less salient; instead, have the user choose a pre-set profile¹²
- **Dynamic Function Allocation:**
 - **Adaptable:**
 - Customize to each user how much control the system has over temperature selection.
 - **Adaptive:**
 - The system could automatically turn down temperature when users are away from home
 - Prompt users with warnings or other feedback when they change the set-point, to encourage energy saving¹³

¹² Inspired by a method used by EnergyHub Inc

- **Representation Aiding:**
 - Make relationships between temperature selection and energy use salient.
 - Guide user choices with recommendations and visual constraints. For instance, ASHRAE¹⁴ standards show what 90% of people find comfortable at different seasons or outdoor temperatures (Brager & Dear, 2000).
 - Recommend (or have default setting for) temperature while home is unoccupied.
- **Matching Automation to Mental Models:**
 - There may be different or unexplored ways to represent temperature or energy feedback. For example, people are better at sensing temperature *differences* than static set points (“I want it warmer/colder” versus “what temperature do you want?”).

5.4 Choice of Design Space

I chose Temperature Selection Ambiguity for three reasons: it tied with “Ambiguous Understanding of Furnace Operation” for most number of pitfalls; more strategies were applicable to this design space than to any other; and I observed interaction difficulties, user comments, and published references related to this design space in all previous stages of this design investigation.

Already, a simple solution became apparent from considering representation aiding, and was desirable to explore: to provide users with thermal comfort guidance, directly on the control interface. Investigating this design space and addressing a possible solution required a second exploration of the literature, seen in the following section.

¹³ like Nest Labs does with the leaf icon on their thermostat (“Next Learning Thermostat,” 2013)

¹⁴ American Society of Heating, Refrigerating and Air-Conditioning Engineers

6.0 Comfort, Control Labelling, and The Anchoring Effect

Having chosen the temperature-selection design space, my design objective was to provide thermostat users with comfort-range information to help them choose more-appropriate temperatures. I was interested to determine if different kinds of message framings—like social norms versus non-social information—would help motivate these users to accept and act on this information. Further, since I would be presenting users with numbers, I wanted to determine if the cognitive anchoring effect was a factor. As such, a further literature investigation was required into thermal comfort control and perception, control labelling, and the cognitive anchoring effect.

6.1 Comfort and the Design of H&C Systems

Recall from Section 2.3.2 that Rathouse and Young (2004) found comfort was the main consideration determining people's PT use. As such, focusing on comfort may prove more persuasive than focusing on efficiency. It is necessary to review the literature on human thermal comfort and the thermal-comfort standards used in the design of H&C systems.

The engineering paradigm of human comfort states that all people experience comfort in a uniform and narrow range of temperatures, humidity, airspeed, clothing, and metabolic rate, and thus need mechanized H&C systems to keep those conditions as stable as possible (Kuijter & De Jong, 2011; Lovins, 1992). This view has dominated widely-adopted thermal-comfort design standards¹⁵ in the design of buildings and their mechanized H&C systems, approaches that define comfort conditions based on laboratory climate-chamber experiments.

However, several researchers argue that a more accurate description is people experience comfort differently, based on the combination of complex physiological, psychological, and cultural factors, highly influenced by recent thermal experiences (Brager & De Dear, 1998; Kuijter & de Jong, 2011; Lovins, 1992). According to Pineau (1982) "Comfort is not a

¹⁵ i.e. ASHRAE standards

universal concept. It does not have the same meaning for all individuals, who use different frames of reference to evaluate it, from which comfort needs arise."(p 282).

Psychological Influences of Thermal Perceptions

Many subjective factors influence a person's perception of their comfort. For example, in laboratory studies of thermal comfort, participants placed in a meat-locker reported a perceived temperature significantly lower than those placed in a standard temperature chamber of the same temperature (Lovins, 1992). However, when the meat-locker was decorated to look like a comfortable room, that perceived difference disappeared.

A sense of control over one's thermal environment, even if never exercised, can lead to people feeling more comfortable and less stressed (Hedge, Khalifa, & Zhang, 2009). Perceptions of thermal comfort are also affected by recent meal, alcohol, and nicotine consumption (Lovins, 1992).

Control of Temperature

This subjectivity of temperature perception is also seen in the control of thermal conditions. In one wintertime study, some experimental groups reduced indoor temperatures by up to 2.7°C compared with the non-treatment group, with no significant difference in perceived comfort (Lovins, 1992). In a related summer study, residents cut AC use by an average of 34% "with no changes in perceived comfort or clothing worn, and with minimal temperature change (1.5 °C) in the home" (Lovins, 1992, p. 12).

Some research suggests that comfort is not even the main determining factor of how much heating or cooling people want and use in their own homes (Lovins, 1992). It suggests degree to which H&C systems are used depends strongly on complex relationships between comfort and various expectations, values, and preference. In a study involving controlling ACs to run 50% of their normal runtime on the hottest days, load savings were uncorrelated with perceived discomfort (Kempton, Reynolds, Fels, & Hull, 1992). Another study found 55% of electricity variance in AC use between identical townhouses was explained by attitudinal variables alone (Seligman, Darley, & Becker, 1978).

Variations

When choosing their comfort temperatures in everyday life, people show a much wider range between individuals than in carefully controlled climate-chamber experiments (Lovins, 1992). In their research, Parker et al (1996) found indoor temperature varied from 21.4 to 27.8°C between houses on the hottest day in Miami. Likewise, Kempton et al. (1992) showed a 22.2 to 28.4°C range in comfortable summer temperature in NJ. In wintertime, Shipworth et al (2010) also found a large variation in residential thermostat settings (SD = 2.5-3 °C), with a mean/median at 21°C, but 30% set below 20°C and 40% set above 22°C. All these findings stand in contrast to the $\pm 1.2^\circ$ range suggested by laboratory studies (Kempton, Feuermann, et al., 1992).

Some of these variations can be explained by physiological differences. Humans exhibit a large interindividual variability in skin temperature and evaporative heat loss (12.5% to 75% of total body heat loss) at similar activity levels, and thus will also have profoundly different sensitivity to humidity & air movement (Lovins, 1992).

Some gender differences in comfort also deserve mention. In one UK study, women seemed to experience lower temperature more negatively than men (Carlsson-Kanyama & Lindén, 2007). Addressing clothing differences in office buildings, Lovins (1992) points out that men dressed in full suits tended to sit side by side with women in light dresses.

Variations also exist across cultures. Temperatures in households, in countries with similar energy prices and incomes, keep winter indoor temperatures that vary from 10°C in New Zealand, to 14°C in Japan, to 17°C in Norway, to 21°C in Sweden (Lovins, 1992).

Human Adaptability

Humans can adapt to different thermal environments physiologically (acclimatization), behaviourally (through conscious and unconscious actions), and psychologically (unconsciously shifting perceptions and expectations) (Brager & De Dear, 2000). There is a well documented strong positive correlation between people's comfort temperature and the prevailing indoor and outdoor temperatures (Brager & De Dear, 2000).

Mechanized H&C and Natural Ventilation

One result of applying the engineering paradigm is that—due to human adaptability—it ironically reinforces itself. Mechanized H&C actually gives building occupants an expectation and a perceived need for constant thermal conditions (Brager & De Dear, 1998; Kuijer & de Jong, 2011; Lovins, 1992). To these occupants, non air-conditioned buildings seem hotter. These centrally-conditioned occupants are also twice as sensitive to thermal changes as those used to natural ventilation, and tend to become more critical when their expectations are not met (Brager & De Dear, 2000).

In comparison to centrally conditioned buildings, indoor temperatures in naturally ventilated buildings more closely match the daily and seasonal variations outdoors (Brager & De Dear, 1998). Occupants of these buildings have more relaxed expectations, become more tolerant of dynamic and non-uniform indoor conditions, and even prefer having a closer connection with daily and seasonal changes (Brager & De Dear, 1998, 2000). While comfort ranges for occupants of both building types are related to seasonal temperatures, those for naturally-ventilated occupants vary twice the amount with the seasons.

6.1.1 Adaptive Comfort Standard for Naturally-Ventilated Buildings

Brager and De Dear (Brager & De Dear, 2000) argue that reliance on engineering-paradigm comfort standards based on laboratory studies “has allowed important cultural, social, and contextual factors to be ignored, leading to an exaggeration of the “need” for AC.” (p. 2). Further, giving people more thermal control, and allowing temperatures to more closely follow outdoor patterns, could both improve occupant satisfaction and reduce energy consumption (Brager & De Dear, 2000).

The ASHRAE Adaptive Comfort Standard, shown in Figure 6.1, reflects research from naturalistic field studies where people in naturally-ventilated buildings (working in at different indoor temperatures from each other) were asked to rate their current comfort level (Brager & De Dear, 1998, 2000). The standard was developed to link indoor temperatures to a building's climatic context, accounting for occupants' past thermal experiences and expectations. Compared to other design standards, this

standard simplifies the determination of acceptable temperatures by relating it only to mean-monthly outdoor temperature¹⁶ and leaving out factors like humidity and air speed. “Acceptability limits” refer to the percent of occupants finding conditions acceptable.

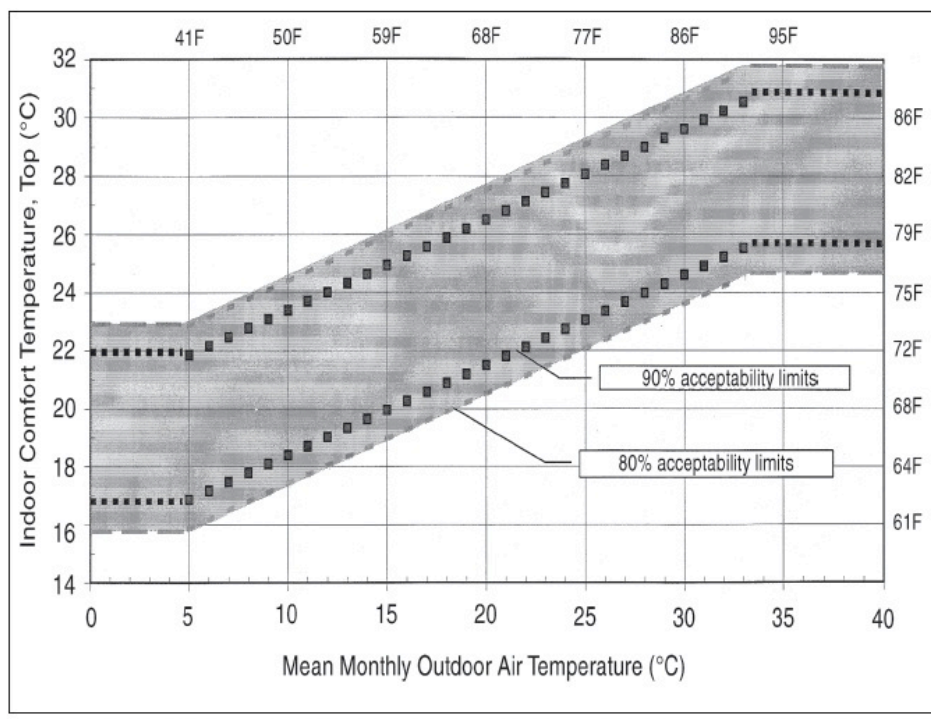


Figure 6.1 – The adaptive standard for naturally ventilated buildings (Brager & De Dear, 2000). © ASHRAE Journal, 2000. This use is by permission of the copyright holder.

In comparison, the more widely used ASHRAE Variable Temperature Standard for centrally conditioned buildings provides a narrow range of acceptable temperatures for summer and winter conditions, based on humidity, wet-bulb temperature, clothing levels, and activity level (DE DEAR & BRAGER, 1998). At the 90% acceptability level and standard conditions, it recommends an optimum temperature of 22.5 ± 1.2 °C in winter and 23.5 ± 1.2 °C for summer. Not only is the difference between summer and winter noticeably more narrow than the Adaptive Standard (just 1°C, compared to as much as 6°C for Toronto weather¹⁷), the acceptable range is also half as large (2.4°C compared to 5°C).

¹⁶ the average of daily highs and lows over a month

¹⁷ Typical Toronto summers reach MMOAT of 25 °C and winters drop below 5 °C (“Climate Data Online,” n.d.)

I used the Adaptive Standard to produce the comfort range presented in the final design prototype.

6.2 Control Labelling

I investigated the literature on effective labelling of controls and displays. Control labelling (CL) is referred to minimally in the classical human-factors literature and I only found a single set of researchers looking at CL used for influencing or aiding the use of controls. Consequently, I briefly explored research on the effective use of warning labels to compensate for this gap in the literature.

6.2.1 Human Factors Principles and Guidelines for Controls and Displays

Many classical human factors texts and standards recommend criteria for the design of controls and displays. (Boff & Lincoln, 1988; “Military Standard 1472D,” 1989; O’Hara, Brown, Lewis, & Persensky, 2002; Sanders & McCormick, 1993; Woodson, Tillman, & Tillman, 1992). However, most of these texts mention CL only in reference to coding and panel labelling, used to identify and distinguish only between controls.

Guidelines that may apply to my work include the following:

- Labels should be placed above the controls so they are not blocked by the user’s hand when being used (“Military Standard 1472D,” 1989; Sanders & McCormick, 1993; Woodson et al., 1992).
- Scale units should increase in magnitude from left to right (or bottom to top) (Boff & Lincoln, 1988; O’Hara et al., 2002; Woodson et al., 1992). There are some other specific guidelines about scale-markings on dials and gauges, like width and proximity to each other and the dial, etc. (Boff & Lincoln, 1988; Woodson et al., 1992).
- Distance and proximity between labels and controls should be balanced so the label can be seen and still be associated with the control (“Military Standard 1472D,” 1989; Woodson et al., 1992). A slider’s length should be long enough to read and position precisely, but short enough to avoid long response times (O’Hara et al., 2002).

- To be read quickly and easily, labels should be oriented horizontally, unless they are not safety-critical (“Military Standard 1472D,” 1989; O’Hara et al., 2002).
- Label wording should be familiar to the operator; a primary consideration over brevity (Boff & Lincoln, 1988; “Military Standard 1472D,” 1989). At the same time, they must convey verbal meaning as directly as possible, using abbreviations only when familiar to the user (“Military Standard 1472D,” 1989; O’Hara et al., 2002).
- Controls and displays should be labelled with information the user requires “for proper identification, utilization, actuation, or manipulation of the element”(“Military Standard 1472D,” 1989).
- Control labels should “indicate the functional result of the control movement (e.g. increase, ON, OFF)”(“Military Standard 1472D,” 1989).
- Units of measurement should always be displayed (“Military Standard 1472D,” 1989).
- Labels should describe what is being displayed, not just label it (Boff & Lincoln, 1988; “Military Standard 1472D,” 1989).
- Sliders should be used with discrete controls—those in which users can select one of a limited number of settings (Sanders & McCormick, 1993)—with more than two levels “to allow easy recognition of relative switch settings” (Boff & Lincoln, 1988). Boff and Lincoln recommend rotary dials for continuous controls—those allowing selection of any value on a continuum (Sanders & McCormick, 1993)—where precise adjustment is required.
- Critical limits or different operating ranges should be coded graphically into a slider facilitate recognition (O’Hara et al., 2002). Similarly (and relevant to my purposes), reference values should be provided to help users “judge the appropriateness of values”.

Some texts briefly mention the use of instructions on controls. For instance, on-product instructions (and any labels) should be as brief as possible without obscuring their meaning (“Military Standard 1472D,” 1989; Woodson et al., 1992). Instructions should be all capital letters unless they are more than one line (“Military Standard 1472D,”

1989). Boff and Lincoln (1988) say instructions should “stand out” in person-computer displays.

Boff and Lincoln (1988) share suggestions for systems with infrequent or casual users, including: training users in principles, not details; making available choices explicit and constrained; and using natural language. They also describe a relevant concept termed “disruption of psychological set” related to problem-solving aids: “such an aid is intended to disrupt any bias or “sets” that the user may employ, and thereby stimulate more creative, or novel, problem-solving attempts”. An enhanced control label may serve this function.

6.2.2 Warnings Literature

Research on providing visual safety warnings may help to identify characteristics of effective CLs. Laughery and Wogalter (in press) posit warnings—in line with the intention for enhanced control labels—serve three purposes: to provide information, to influence behaviour, and to remind or cue users of what they already know. Effective warnings do this in three necessary stages: attracting the user’s attention, eliciting knowledge and comprehension, and enabling or motivating compliance behaviour. Without the message first being attended to or providing necessary information, a person cannot make an informed decision to comply (Laughery & Wogalter, In Press; Sauer et al., 2008).

In general, message brevity is associated with a greater likelihood of being read (Laughery & Wogalter, In Press). Also, the more familiar or experienced someone is with a product or affiliated behaviour, the less likely they are to seek or read a warning on that product, and less likely to comply with said warning. Time pressure or multi-tasking¹⁸ can also reduce warning compliance. In general, Laughery and Wogalter (in press) suggest that warning compliance is lowest when the cost-of-compliance is high, and higher when hazard is perceived to be greater.

¹⁸ Factors relevant to those programming and interacting with PTs

6.2.3 Enhanced Control Labelling

This section investigates the use of CL enhancements in the residential domain. Sauer, Wiese, and Ruttinger (2002) suggest domestic users differ from employed workers in five key ways: domestic users are heterogeneous; have no chance for formal training; cannot be selected for competence; define their own tasks; and have no performance supervision or feedback from other users. Referring to low-complexity electrical consumer products (ECPs), Sauer et al. (2003) suggested there are fewer possibilities to alter behaviour to reduce energy use because 1) habits develop quickly and 2) instruction manuals are used even less than with complex ECPs.

Task performance is closely related to the quality of the mental model (Wickens & Hollands, 1999). Sauer et al (2002) suggest that, in the absence of formal training in the domestic domain, the mental model may be improved by product information and design features (visibility, feedback, controls design, etc.). However instruction-manuals are generally not read by users (Sanders & McCormick, 1993; Wiese, Sauer, & Rüttinger, 2004). Recall from section 2.3.2 that many people liked having instructions on their PTs, but were also confused by the meaning of symbols (Rathouse & Young, 2004). People also found written instruction manuals were too complicated and did not provide them with enough support (Combe et al., 2012).

The proximity-compatibility principle states that items of high processing proximity should be placed near each other on an interface (Wickens & Hollands, 1999). It should follow that task-relevant information or instructions should be placed near the relevant controls. On product information (OPI) and display-control labelling (DCL) aim to accomplish this.

OPI Theory

On product information (OPI) may be effective in modifying user behaviour. OPI provides a knowledge-conveying function, informing the user about how to best operate an appliance (Sauer et al., 2003). Sauer et al (2008) argue OPI is better than instruction manuals, since it is permanently visible. Similar to warnings, OPI's efficacy is influenced by a number of factors: location proximity, procedural explicitness, print

design, and amount of information presented. Limiting to 1 or 2 messages may also increase a label's effectiveness (Wiese et al., 2004).

Display Control Labelling (DCL) Theory

Many ECPs lack information on appropriate control settings (Sauer et al., 2008).

Labelling is often limited to quantitative information, like physical units, which are likely unhelpful to domestic users. Without qualitative labelling, novice users are more likely to use maximum settings, as they may connect maximum power with maximum effectiveness (Sauer et al., 2004).

Display control labels (DCLs¹⁹), or OPI integrated into displays and controls, makes “a stronger link between behavioural advice given and the control device needed to implement the advice” (Sauer et al., 2008, p. 72). If the user's mental model is lacking, enhanced DCLs can provide information about optimum control setting under specific operational situations (Sauer et al., 2002). In the energy conservation domain, enhanced DCLs can provide the user with information about the most energy-efficient settings on controls, conveying knowledge and prompting behaviour change (Sauer et al., 2004).

A limitation of DCLs is the lack of information quantity and complexity that can be conveyed to the user (Sauer et al., 2002). There is also mixed evidence for effectiveness of enhanced DCLs to influence behaviour (Sauer et al., 2002, 2003).

Study Results

Sauer et al. (2008, 2002, 2003, 2004) performed a number of investigations into OPI and DCL in the domestic domain. Sauer et al. (2003) found that providing OPI on kettles prompting the user to “only boil as much water as required”, significantly reduced energy use .

Sauer et al., (2002, 2004) tested enhanced versus standard DCLs on vacuum cleaners. In their first study (Sauer et al., 2004), the enhanced label conveyed a degree of

¹⁹ Sauer's term. I will use DCL and CL interchangeably in this thesis.

dirtiness using colour-coded LEDs. There was no effect of label on any measure, likely because the control was in the distal position (away from the user's view), emphasizing the need for label-saliency. In their follow-up study, Sauer et al. (2002) tested enhanced (showing the most-efficient setting) versus standard DCLs with controls in the distal (on the vacuum body) versus proximal (on the hand-grip) positions. They found the enhanced label led to significantly reduced suction levels (and thus energy use), but only in the proximal position.

Sauer et al. (2008) also studied the effectiveness of enhanced DCLs on the temperature control of a pressure washer (providing recommended ranges for different objects). In one experiment involving photo-prototypes, they found the enhanced label led to significantly lower temperature settings, compared with standard labelling (temperatures only). In a second experiment, using enhanced labels on the device itself, they found users of the enhanced labels chose significantly different settings between use scenarios (range 51°C), while standard-label users tended towards a narrow range of settings (range 3°C). A significant interaction between the control-label settings and scenario indicated most users complied with the enhanced DCL recommendations. The researchers speculated that, with the standard label, users would choose a medium setting, observe adequate cleaning performance, and thus settle at that setting.

6.3 The Anchoring Effect

In my design intervention, I wished to see if different message framings of the comfort range information, especially social norms, would influence users to more readily use the presented information. If the intervention were to be effective, I wanted to know whether it was due in part to a cognitive anchoring effect on the presented numerical comfort range or the message's meaningfulness. In this section, I explore anchoring, and how it interacts with social norms.

The anchoring effect is a robust cognitive heuristic, shown to be effective across many domains involving judgement performed under uncertainty (Furnham & Boo, 2011; Tversky & Kahneman, 1974). The dominant viewpoint suggests that when a numerical

anchor is presented, people tend to bias their answers towards the anchor. People seem to make their estimates by starting at the anchor and then insufficiently adjusting their response towards the correct answer. Many underlying mechanisms have been proposed, but none can fully account for the effect (Furnham & Boo, 2011). Anchoring is effective whether the anchor is provided or self-generated, and regardless of how meaningful or relevant it is (even if it is known to be wrong) (Furnham & Boo, 2011; Switzer & Sniezek, 1991). In one study, for example, estimations of spending in a restaurant were affected by the name of the restaurant (“Studio 17” or “Studio 97”) (Furnham & Boo, 2011).

Conditions that tend to increase the anchoring effect include: high ambiguity, low familiarity or personal connection to a problem, more trustworthy sources, and anchors that are closer to the range of plausible answers (Furnham & Boo, 2011). Knowledge or certainty about the target domain has been shown to reduce the effect (however even experts are still influenced by anchors) (Furnham & Boo, 2011; Lombardi & Choplin, 2010; Tversky & Kahneman, 1974); the more people know about the value they intend to estimate, the less they will rely on anchors. Most anchoring studies typically provide only a single anchor. The little existing research suggests providing multiple anchors tends to reduce the effect of any one anchor, especially when they contradict each other (Switzer & Sniezek, 1991; Whyte & Sebenius, 1997).

6.3.1 Anchoring and Social Norms

Research suggests that social norms (people’s perceptions of what socially-relevant others accept or do) may be confounded or even explained by the anchoring effect (Eyssel et al., 2006; Lombardi & Choplin, 2010). Interpreting the success of norms as being due to anchoring does not require the group of others be perceived socially relevant, nor that the perceiver be motivated to belong, nor even that the information about others’ beliefs be perceived as true (Eyssel et al., 2006). The following outlines specific investigations into this hypothesis.

Switzer and Sniezek (1991) conducted two experiments wherein they provided anchors of different degrees of relevance, including descriptive social normative, first at two different anchor levels, then with multiple anchors. In both cases they found a

significant effect of the anchors numerical level, but no significant effect of anchor's relevance (i.e. no difference between social and non social anchors of the same level). Subjects used the anchoring information to make their judgments regardless of their objective usefulness.

Testing the efficacy of social-normative alcohol-cessation interventions, Lombardi and Choplin (2010) suspected that educating students about their peers' conservative drinking habits may lower their perceived and thus reported consumption (due to anchoring on the information), but that it would have no effect on their actual behaviour. Across three experiments, they asked students how much they drank in a week after providing a high or low false norm of their peers' consumption. In both cases, they found the normative anchor significantly affected the student's self-reported consumption (despite that students' actual level could not have changed).

Eyssel et al (2006), studying rape-myth acceptance, provided male students with either a descriptive social norm ("Male students answered, on average ___") or an explicitly "randomly-generated" anchor. They found a significant effect of the value's level, but no effect of the presentation format (norm versus anchor), suggesting the norm was functionally equivalent to the anchor.

6.4 Moving Forward

In this section, I looked at the literature needed to inform my design solution to the temperature-selection design space. I started with the literature on thermal comfort, which highlighted the subjectivity and interindividual variability of human comfort perception and how it stands in contrast with the engineering paradigm of comfort. This followed with the adaptive comfort standard for naturally ventilated buildings, which accounts better for this subjectivity and variability than standards for buildings with mechanized H&C systems. I then examined the literature on the labelling of controls, first mining the classical-human-factors and warnings literature for guidance, then exploring in depth a group of research done on enhanced control labelling. I finished with an overview of the cognitive anchoring effect and its potential interaction with social norms.

The findings of this review present opportunities for designers and researchers, as well as insights and guidance for my own investigation.

First, the research on the subjectivity and variability of comfort perception highlighted important areas for design consideration. Because of the wide variability of acceptable temperatures, PT designers may need to deemphasize the selection of specific temperatures and instead focus on other aspects of thermal control. Similarly, it may not be possible to satisfy all users and occupants with a single setting; building and PT designers may need to shift their focus to person heating rather than space heating. The subjectivity also presents an opportunity to design for the control of more efficient temperature settings by addressing people's perceptions and expectations.

Second, at least some of the guidelines from the classical human factors texts could be applied to CL design. In addition, the factors that improve compliance to warning labels may also improve compliance to control labels.

Third, the field of enhanced DCL is relatively young and thus presents opportunities for future research and design. Much investigation is needed on the effectiveness of quantitative enhanced DCLs in guiding users' control choices. DCLs could also be used to address some of the flawed mental models presented in Section 2.1.2.

Finally, the anchoring effect is of importance to my investigation. There is evidence presented here that the success of presenting social norms is due to people anchoring on the value, rather than being influenced by the message's social meaningfulness; therefore in any study that uses social norms, investigators should intend to separate the effect of the norms value from its meaningfulness, to rule out an anchoring effect. This is a primary reason for including an anchoring condition in my study.

Recall that the more people know about the value they intend to estimate, the less they will rely on anchors. This is relevant to my application because people tend to have some idea (even if faulty) of where they prefer to set their thermostats. Also, recall providing multiple anchors tends to reduce the effect of any one anchor (Switzer & Snizek, 1991;

Whyte & Sebenius, 1997). This may be relevant since I intend to anchor participant on a range rather than a single temperature.

Some noteworthy gaps and weaknesses in the literature deserve mention and are in need of further research:

- As I have pointed out, there is practically no human-factors research on control labelling for influencing and aiding the use of controls. Perhaps this is because typical users of human factors interventions are workers who are already motivated to take proper action.
- There is mixed evidence for effectiveness of eDCLs to influence behaviour. Some presentation factors deserve further study.
- I have found no research on anchoring to visual features nor on the use of anchored ranges, both of which are relevant to my design. There is also little research on the use of multiple anchors, which would be the closest thing to anchored ranges.

Through an iterative process of evaluating the above-mentioned literature and developing prototype ideas, in the temperature-ambiguity design space, I eventually settled on a simple design and experimental evaluation proposal. This is outlined in the following section.

7.0 Prototype Development and Experimental Evaluation

My investigations up until this point highlighted how people generally lack the understanding necessary to effectively manage their energy consumption, and how current technology insufficiently supports users to this end. In particular, people seem to have an ambiguous understanding of what temperatures are appropriate for comfort and energy efficiency.

My goal was to develop, and then evaluate, a prototype that would address this design space. The general idea was to provide thermostat users with information they would otherwise lack, thus filling in the anticipated gaps in their mental model of the relationship between indoor temperatures and thermal comfort; additionally, I intended to manipulate presentation format, to see if applying anchoring and social norms would influence these users to accept and act on said information. This section describes the prototype and then outlines an experimental method to evaluate it.

7.1 Research Questions

The research questions were:

- Does providing users of residential thermostats with a range of comfortable temperatures tend to decrease (in the heating season) the settings they choose for their own comfort?
- Is this mediated by whether they consider comfort alone, or in addition to considering energy efficiency?
- To what extent are the meaningfulness of the information or social normative influence factors in the selecting of a temperature setting?

7.2 Prototype Details and Research Method

I chose to create an enhanced CL by overlaying a simple PT interface with the seasonally-appropriate ASHRAE Adaptive-Standard (Brager & De Dear, 2000) comfort-range (in the 90%-acceptability limits)²⁰.

I decided that a winter scenario would be most appropriate for the current seasonal conditions²¹; heating is more likely than cooling in the autumn, and people would be more likely to set their thermostat for the winter (rather than summer) months at this time. Conveniently, the adaptive comfort standard (Brager & De Dear, 2000) levels off below a mean-monthly outdoor air temperature of 5°C, which is appropriate for Toronto winters between December and March (“Climate Data Online,” n.d.). This produced a comfort range of 17 to 22°C.

I found from the Usability and CM studies, people tended to choose wintertime temperatures of around 21 or 22°C, and this is supported by the literature (Karjalainen, 2007; Shipworth et al., 2010). Since this represents the upper end of my chosen range, I suspected my prototype would influence users to select lower temperatures than without enhanced labelling. This would have the beneficial effect of lowering energy consumption.

I used this particular “Adaptive” comfort standard (for naturally ventilated buildings)—compared to the widely-used “Variable” standard for central conditioning—for a few reasons:

- While homes are generally centrally conditioned, they are not done so in the constrained way that office buildings with mechanized HVAC²² systems are. Similar to naturally ventilated buildings, residents generally will experience more variability due to having more thermal control (i.e. temperature settings not constrained by any standard, and the choice to run the H&C at all).
- The winter range for the Adaptive standard (17 - 22 °C) is further away from the observed norm of 21 or 22°C, than the Variable standard is (21.3 – 23.7 °C). Thus, I

²⁰ recall Figure 6.1 in Section 6.1.1

²¹ late September, as the weather began to cool.

²² Heating, Ventilation, and Air Conditioning

anticipated that I would only be able to observe a change in chosen temperature using the Adaptive Standard.

- Having indoor temperatures more closely follow seasonal changes (as is so from the adaptive standard) presented an opportunity to influence residents to choose more energy-efficient temperature settings (due to a smaller indoor-outdoor heat differential).

Questions of this standard's appropriateness are addressed in Section 9.2, on study limitations.

Further, I was interested to see if the presentation format of this information—with different levels of meaningfulness and social relevance—would affect participant's choices. Thus, I included an anchoring condition, a meaningful non-social condition, a descriptive norm condition, and an injunctive norm condition.

The anchoring condition was included for two reasons. First, I wished to know if cognitive anchoring, in and of itself, would influence people to choose lower temperatures. Second (and related), if there is no effect of the anchoring condition (versus control) then I could infer that the success of any meaningful message framing would be due to the message itself and not the anchored numbers. Recalling that anchoring has been shown to be effective regardless of how meaningful the anchor is (Furnham & Boo, 2011; Switzer & Snizek, 1991), I decided to use non-meaningful message framing "Arbitrary Range" to isolate the effect.

The subsequent meaningful message framings would test the effectiveness of social norms on temperature selection. A non-social framing would be included to compare against the social normative framings. The final two conditions would be a descriptive norm (what others do) and an injunctive norm (what others approve of).

Recall again that comfort was found to be the main consideration determining people's PT use; more so than cost, and much more so than environmental concern (Rathouse & Young, 2004). Therefore, following Sauer et al (2002, 2004), I included two instruction conditions to see if the enhanced labelling could influence users to lower their

temperature settings (and thus energy consumption) without necessarily prompting them to be efficient. One condition would prompt users to choose the most energy-efficient setting they would still consider comfortable, while the other asked them to choose simply the most comfortable setting.

Concerned that habitual responses may detract from any effect of control labelling, and remembering that habits rely on familiar contextual cues (Verplanken & Wood, 2006), I chose a scenario in which the participant would imagine interacting with a new PT, programming it for the first time.

Since most thermostats allow a 10 to 30 °C range in selected temperatures (Shipworth et al., 2010), I used that as my scale's lower and upper limits.

7.3 Experimental Design

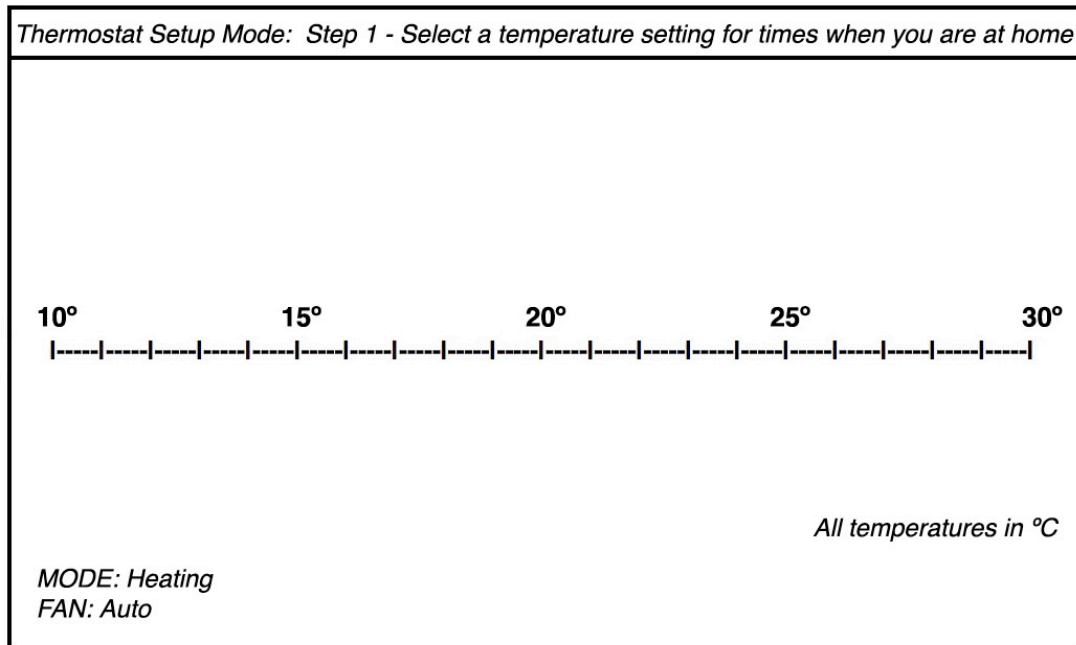
The experiment followed a 2x5 mixed factorial design. The within-subjects factor²³, Instructions, was the nature of instructions given to participants:

- Comfort-Only (C) instructions ("Choose a temperature setting that you would be most comfortable in."), and
- Comfort+Efficiency (C+E) instructions ("Choose the most energy-efficient temperature setting that you would be comfortable in.").

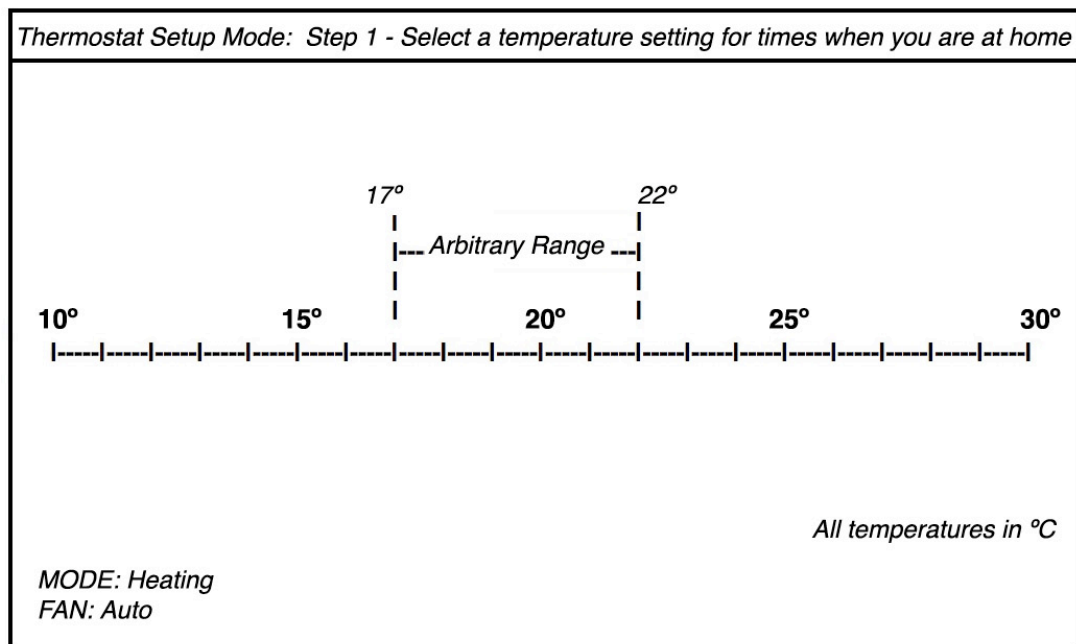
²³ Counter-balanced

The between subjects factor, Display Type, was the information level of display labelling of the thermostat prototype:

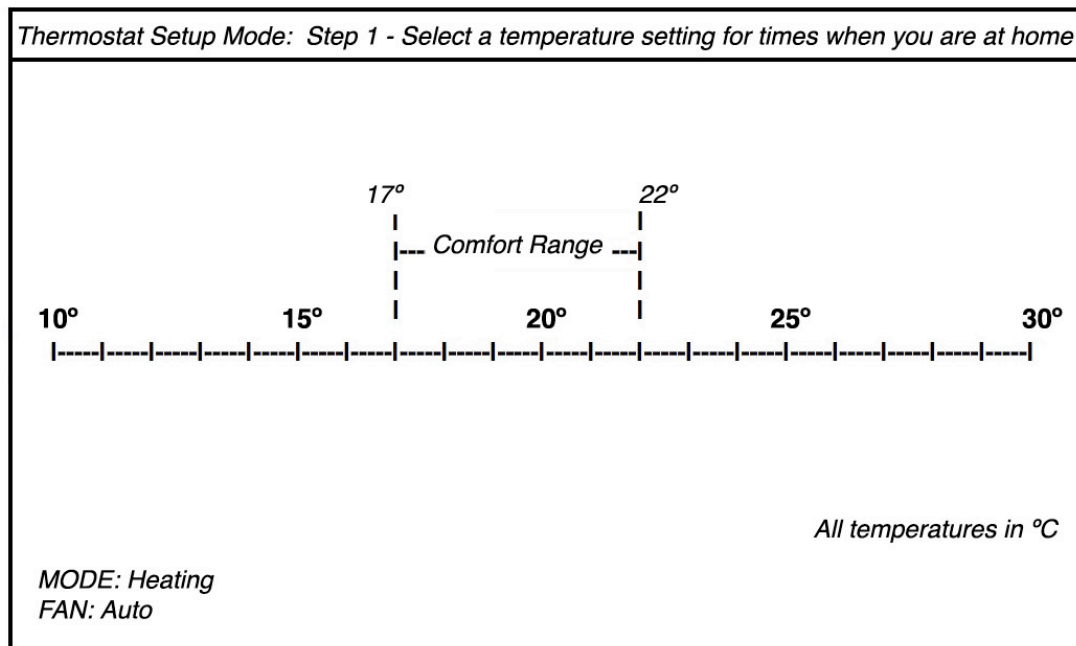
1. Scale Only: (no added information—just the temperature scale),



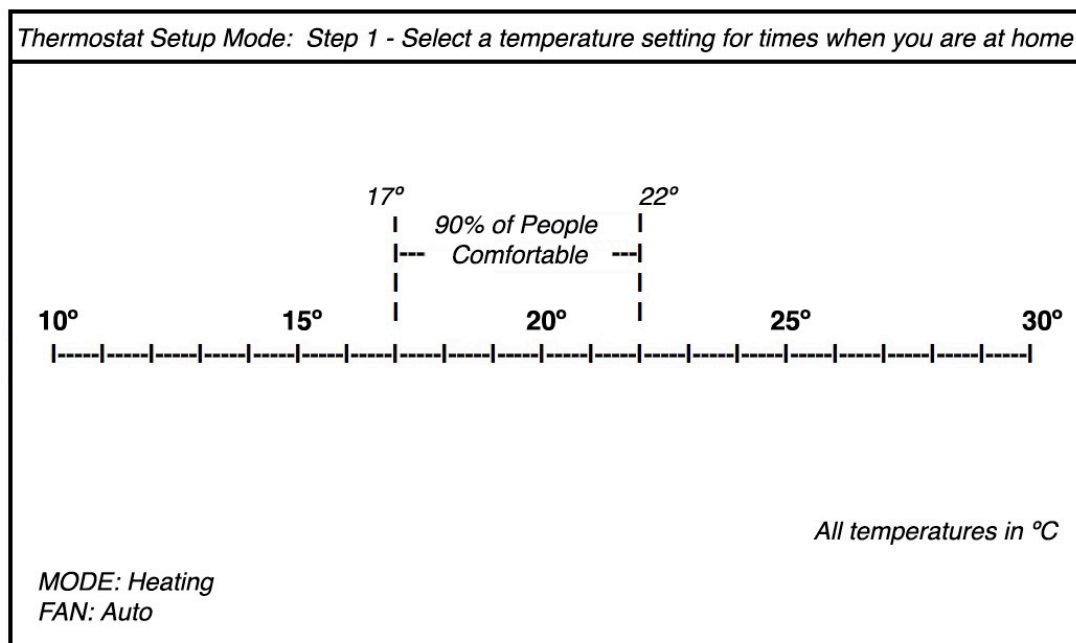
2. Anchor (overlaid temperature range, labelled “arbitrary range”. If participants asked, it was described as “randomly generated”),



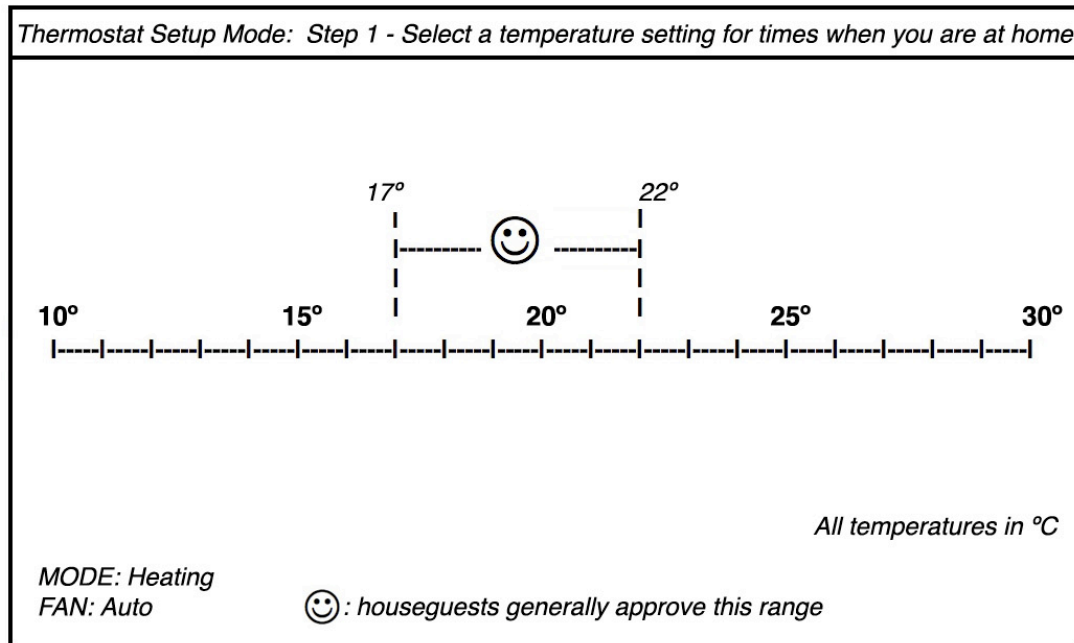
3. Meaningful-Non-Social (overlaid temperature range, labelled as the “comfort range”),



4. Descriptive Social Norm (overlaid temperature range, labelled “90% of People Comfortable”. If participants asked, it was described as “the research-verified range that 90% of people find comfortable in the summer”), and



5. Injunctive Social Norm (overlaid temperature range, labelled with a smiley face “😊”—to signify social approval²⁴—and described below²⁵ as “houseguests generally approve this temperature range”.)



²⁴ Recall that Schultz et al (2007) similarly used smiley faces to communicate injunctive norms

²⁵ this message framing included a written description to clarify the symbol's meaning.

7.4 Hypotheses

My hypotheses were:

1. Users will select lower temperatures with each of the enhanced labels than with the ‘control’ scale-only label.
2. Users will select lower temperatures with the efficiency instructions than with the comfort-only instructions.
3. Users will select lower temperatures with meaningful control labels than with the non-meaningful “arbitrary” anchor label. Further, users will select lower temperatures with the social normative labels than with the non-social “comfort range” label.

7.5 Participants

Participants were 144 University of Toronto students (75 male; 64 female; 5 undeclared) whose ages ranged from 17 to 45 ($M = 21.7$; $SD = 5.2$). They were recruited from passersby who consented to participate for 10 to 15 minutes. The current year of study ranged between first to PhD, with the median student—as well as 28% of all participants—being in their second year. Table 8.x1 shows the distribution of their academic program:

Table 7-1 – Participants’ Academic Programs

Academic Program	Count	Percent
Bachelor of Science	44	31%
Bachelor of Applied Science	28	20%
Bachelor of Arts	25	17%
Kinesiology & PhysEd	24	17%
Graduate and Professional Programs	12	8%
Other Bachelor's Degree	6	4%
Other and Undeclared	4	3%

7.6 Materials

The temperature selection task was presented on a paper prototype thermostat, mounted vertically on a poster-board, and was performed with a fine-tipped marker. The participant-information questionnaires (Appendix L) were completed using pen-and-

paper on a table. Indoor temperatures were measured on a portable digital temperature sensor.

7.7 Procedure

On each of the four days, an experiment station was set up in a different atrium of the following University of Toronto (St George Campus) buildings (Table 7-2), with the approval of Facilities and Services. The station occupied each location for one day.

Table 7-2 – Number and percent of participants recruited at each location.

Building Atrium	Number of Participants	Percent of Total Participants
Sydney Smith Hall	30	21
Medical Sciences Building	27	19
Bahen	52	37
Athletics Centre	35	24

Participants could sit to fill out the informed consent, but were asked to stand to complete the experimental task.

1. Introduction and Informed Consent

First, the investigator gave the prospective participant a description of the experiment and the experimental task. At the same time, the investigator presented them with the Informed Consent Form (Appendix J). They were given enough time to read the form thoroughly and ask any questions. If the prospective participant understood the form and wished to participate, then he/she was asked to sign the form. Participants received a completed copy of the Informed Consent Form.

2. Task

Subjects were presented with a paper-prototype thermostat (See Appendix K, and above), consisting of a horizontal linear scale of temperatures, ranging from 10 to 30 degrees Celsius, mounted on a vertical surface, approximately at eye level. Based on the task instructions, participants selected an appropriate temperature by marking an 'X' on the scale with a fine-tip marker. In some conditions of the display-type factor, the prototype

also included a graphical overlay of comfort-range information, highlighting the ASHRAE-recommended $\sim 17\text{-}22^\circ\text{C}$ (for winter conditions).

The participant was then brought to the vertically-mounted paper prototype, corresponding to their randomly assigned²⁶ condition. They were asked to:

Imagine you just bought a new thermostat; you are programming it for your home—choosing settings for different times of the day—to run your central heating system for the winter months. In the following two tasks, you will be asked to select indoor temperatures for times when you would be in the home. For each of the two tasks—based on the instructions I’m about to give you—make your selection by marking a single ‘X’ anywhere on the scale. Keep in mind that, there is no right or wrong answer²⁷, as long as you mark an X somewhere on the scale.

They were then presented, in one of the two Instructions conditions, with the paper-prototype thermostat (Appendix K) and asked to make a selection, followed by their level of confidence in their choice (on a 5-point scale between “very confident and not-at-all confident). Following that they were presented with the second paper-prototype thermostat (in the same Display Type level as the first), and were asked to make a selection with the other Instruction condition, followed by their level of confidence in their choice. Instruction conditions were counterbalanced (i.e. half the participants received C instructions first).

In the C instruction conditions, participants were told: “By marking an X on the scale, choose a temperature setting that you would be most comfortable in.” In the C+E condition, they were told: “By marking an X on the scale, choose the most energy-efficient temperature setting that you would be comfortable in.”

²⁶ Random assignment using Random Number Generator v3.4.0 for Android v2.3.3, by Brandao Apps © 2010

²⁷ To reduced the chance of response bias.

3. Questionnaire

Upon completion of the temperature-selection task, the investigator gave the participant a short paper questionnaire (Appendix L) to complete. The questionnaire asked for participants' age, gender, year and discipline of study, their current thermal comfort level (using the Bedford comfort scale ("Subjective measures," n.d.)), if they use central heating at home, if they control it, and what temperature they thought it was set to. The questionnaire then asked a few questions to test their mental model of thermal comfort, efficiency, and thermostat operation.

4. Debriefing

The participant was then verbally debriefed on the study and given a debrief form (Appendix M), including an explanation of any misrepresentation involved in their condition—for those participants in display conditions 2 through 5—with the option to withdraw their data. They were then informed that they may contact the investigator (through contact information provided on the informed consent form) with questions or concerns, and that their data may be withdrawn by request, if it has not already been analyzed and reported.

Data was also collected on conditions affecting thermal-comfort perception (current indoor and outdoor temperature, and outdoor humidity²⁸).

7.8 Moving Forward

In this section, I outlined the basis for my design intervention and details the experimental evaluation. The results are presented in the following section.

²⁸ Outdoor conditions were estimated based on the time recorded on the questionnaire, using a local weather database ("Climate Data Online," n.d.).

8.0 Experimental Results

This chapter presents the results of the study described in the previous section. Field (2009) was used as a reference for determining appropriate statistical analyses and reporting. All analyses were performed using SPSS v20.0 for Mac OS 10.5. For statistical tests, I am interpreting $p < 0.05$ to be strong evidence for an effect and $p \leq 0.1$ to be moderate evidence.

8.1 Normality and Homogeneity of Variance

First, outliers (any data with z-scores above 3.29, more than 1% of data above 2.58, and more than 5% above 1.96) were removed from the data as per recommended methods of Field (2009); four from C responses and two C+E responses were replaced with the next-highest (or lowest) score plus (or minus) one degree. The selected temperature data were then analysed for normality. Values of skewness and kurtosis for each sample group were used to determine Z-values (see Appendix O). Z-values for both parameters are well under the critical value 1.96 (at $p < .05$), therefore the groups are likely normally distributed.

According to the Kolmogorov-Smirnov (K-S) test (see Appendix O), selected temperatures for scale-only, anchor, and meaningful-non-social displays with C instructions, and the meaningful-non-social display with the C+E instruction, were significantly non-normal, $p < .05$. However, the K-S test is known to produce false-positives with large samples (Field, 2009). Given the results of kurtosis and skewness already mentioned, as well as inspection of the histograms (see Appendix N), I am confident in the normality of this data set.

According to Levene's test of homogeneity of variance, for user-selected temperatures, the variances were equal, $F(4,139) = 1.55$, ns, for C instructions and $F(4,139) = 0.77$, ns, for C+E instructions.

Accordingly, it was deemed appropriate to run parametric analysis (2-way mixed ANOVA in particular) on this data set.

8.2 Descriptive Statistics for the Independent Variables

Across all conditions, the selected temperatures—corrected for outliers—ranged from 13°C to 28°C ($M = 20.9$; $SD: 2.7$). Table 8-1 below shows descriptive statistics for each group of data.

Table 8-1 – Descriptive statistics for each condition.

Display Type	C	Std Dev _C	C+E	Std Dev _{C+E}	Avg	N
Scale Only	23.1	2	20.2	2.7	21.7	29
Anchor	22.7	1.9	20.2	2.7	21.4	30
Meaningful Non-Social	21.3	1.9	19.1	1.9	20.2	29
Descriptive Norm	22.1	1.8	19.4	2.5	20.8	29
Injunctive Norm	21.9	2.9	18.7	2.5	20.3	27
Average	22.2	2.2	19.5	2.5	<u>20.9</u>	144

8.3 Main Effects

A two-way mixed ANOVA was run on the selected-temperature data. The factors were Instructions (2 levels, within subjects) versus Display Type (5 levels, between subjects).

There was a significant main effect of Instructions on chosen temperature, $F(1,139) = 279.52$, $p < .001$. When presented with C instructions, participants chose temperatures ($M = 22.2$) significantly higher than when presented with C+E instructions ($M = 19.5$). A t-test on the order of instructions revealed there was no effect of presentation order for temperatures chosen under C instructions ($t = .38$, $p = .71$) nor C+E instructions ($t = -1.46$, $p = .15$).

There was a significant main effect of Display Type²⁹ on chosen temperature, $F(4,139) = 2.82$, $p < .05$. Contrasts³⁰ revealed that users chose significantly lower temperatures as compared to the scale-only display, for both the meaningful-non-social display, $p < .01$, and the injunctive-norm display, $p < .05$. There is also moderate evidence to suggest

²⁹ Recall, the levels of Display Type were: scale only (control), anchor (“arbitrary range”), meaningful non-social (“comfort range”), descriptive norm (“90% of people comfortable”), and injunctive norm (smiley face with legend: “houseguests generally approve this range”)

³⁰ There are no test-statistics for the contrasts. See Table O-10-15 for contrast results.

participants chose lower temperatures with the descriptive-norm display compared to scale-only, at $p = .10$. Differences were found to be non-significant between scale-only display and the anchor display, $p = .66$.

There was a non-significant interaction effect between Instructions and Display Type on chosen temperature, $F(4,139) = 1.15$, $p = .34$.

These results are depicted graphically in Figure 8.1. For reference, the graph is superimposed with the 17-22°C-comfort range provided to participants in conditions 1 through 5. See Appendix O for the output from the statistical test.

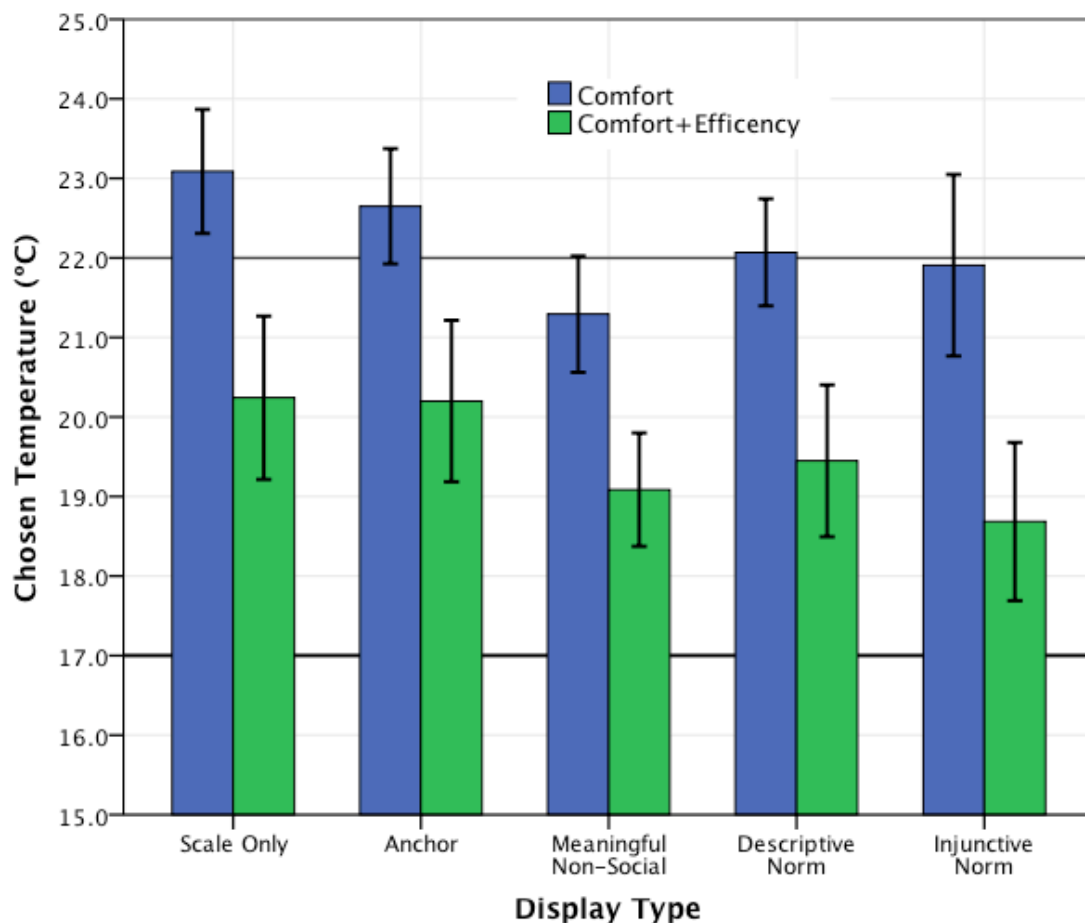


Figure 8.1 – User-selected temperature for each condition. Error bars represent the 95% confidence intervals.

There appears to be an overall downward trend as more meaningful and relevant information is presented to participants.

8.4 Confidence Scores

Participants were also asked to rate their level of confidence in their temperature selections, on a 5-point likert scale. Because the data are ordinal, non-parametric tests were used.

Table 8-2 describes the number (and percent of the sample) of participants who rated at different levels of confidence in their selected temperatures, between the instructions conditions. These results are depicted graphically in Figure 8.2. Note the majority rated themselves “Somewhat Confident” or higher.

Table 8-2 – Number and percent of participants who rated at different levels of confidence in their selected temperatures.

	C Instructions	C+E Instructions
“Very Unconfident”	5 (3%)	3 (2%)
“Somewhat Unconfident”	5 (3%)	9 (6%)
“Neither Confident nor Unconfident”	6 (4%)	12 (8%)
“Somewhat Confident”	67 (47%, median included)	88 (61%, median included)
“Very Confident”	59 (41%)	30 (21%)

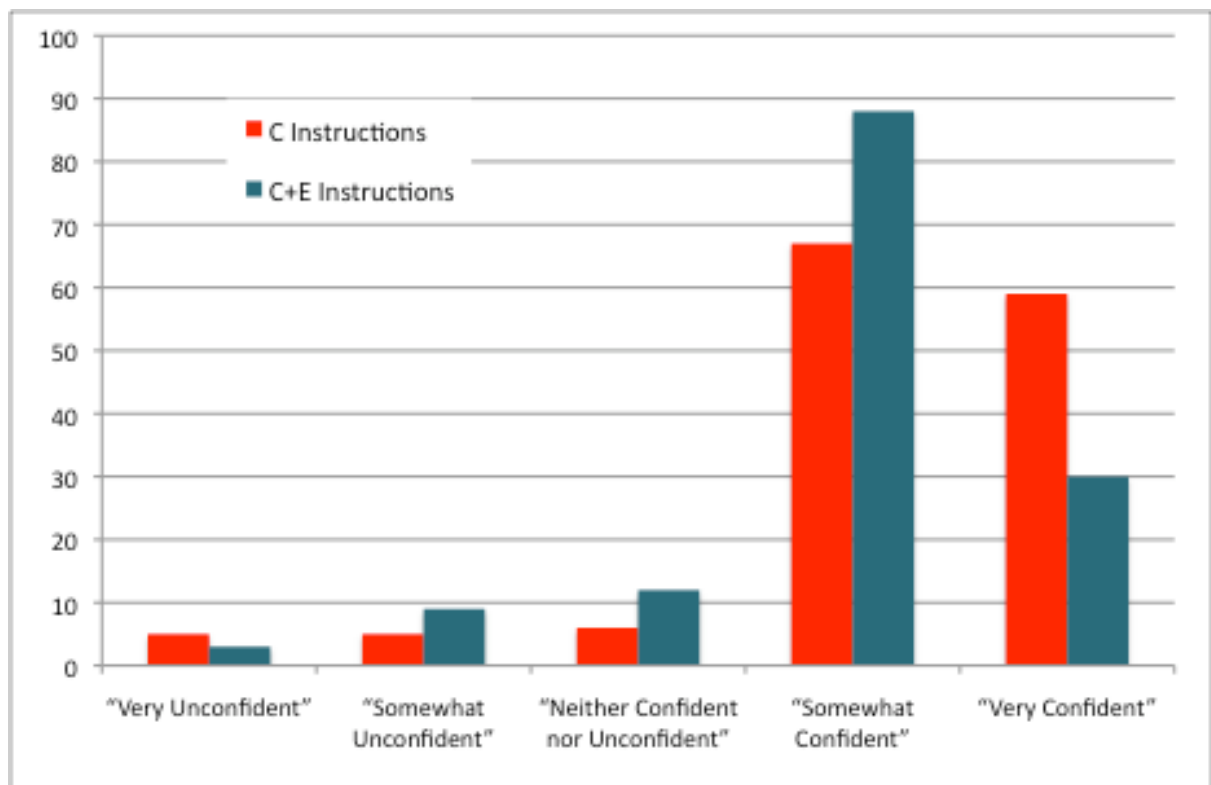


Figure 8.2 – Number of participants who rated at different levels of confidence in their selected temperatures

According to the Wilcoxon Signed-Rank Test, participants were significantly more confident with the C instructions (med = 4 “somewhat confident”, M = 4.2), than with the C+E instructions (med = 4 “somewhat confident”, M = 3.9), $z = -4.02$, $p < .001$. See Appendix O for the output from the statistical test.

The Kruskal-Wallis test revealed no significant effect of Display Type on confidence level, $H(4) = 4.33$, $p = .36$, for C instructions, and $H(4) = 3.36$ for C+E instructions, $p = .50$. See Appendix O for the output from the statistical test.

8.5 Supplementary Data Collected

During experimental trials, outdoor temperatures ranged from 15 to 20° C (average 17.2, SD: 1.5), and outside relative humidity ranged from 43 to 97% (average 69.7; SD: 17.3). Indoor temperatures ranged from 23.1 to 25.0° C (average: 23.0;SD: 0.9). None of these statistics significantly correlated with user-selected temperatures.

8.6 Questionnaire Data

Because a sense of control is thought to affect perceived comfort (Hedge et al., 2009), and because familiarity is thought to affect the efficacy of anchors (Furnham & Boo, 2011; Lombardi & Choplin, 2010; Tversky & Kahneman, 1974), users were asked about their current heating systems. When asked if they had central heating at home, 129 participants (90%) said yes, 12 (8%) said no, and 3 (2%) did not know. For those with central heating, when asked if they had control over it, about half, 69 participants (51%), said yes or sometimes, while 49% said no. When asked what it was currently set to, reports ranged from 12 to 30°C (average 21.2; SD 2.6)

When asked what temperature units they were more familiar with, 122 (90%) said Celsius, 12 (9%) said Fahrenheit, and one each said both and neither.

Asked what their current comfort level was, most participants (133, 93%) indicated they felt Comfortably Cool, Comfortable, or Comfortably Warm—most of whom (70; 49%), as well as the median participant, indicated feeling Comfortable (Table 8.3).

Table 8-3 – Participants' answer to the question "How comfortable are you right now?"

"Much Too Cool"	0
"Too Cool"	1 (0.7%)
"Comfortably Cool"	19 (13%)
"Comfortable"	70 (49%)
"Comfortably Warm"	44 (31%)
"Too Warm"	9 (6%)
"Much Too Warm"	0

8.6.1 Mental Model Checks

When asked “*To be comfortable, indoor temperatures should be _____ in winter than/as in the summer*”³¹, 92 (64%) and 33 (23%) incorrectly answered that indoor temperatures should be warmer or the same as indoor summer temperatures, respectively. Only 18 participants (13%) correctly answered that comfortable winter indoor temperatures are lower than in the summer.

When asked about energy-efficient heating, 14 participants (10%) incorrectly answered that heating is more efficient at higher temperature settings. Despite this misunderstanding, no participants chose higher settings with the C+E-instructions than those with the C instructions. However, as will be shown below, they did choose higher settings on average than participants with the correct mental model.

When asked if a thermostat works like a valve, 57 (40%) correctly answered that the statement was False, while 22 (15%) incorrectly indicated it was true, and 64 participants (45%) were unsure.

8.6.2 Predictors Analysis

Several other predictors had an effect on user-selected temperatures.

A two-way mixed ANOVA showed a significant main effect of control over heat on chosen temperature, $F(1,139) = 12.95$, $p < .001$. Participants who said they have some control over their home’s heat chose higher temperatures ($M = 22.6$ for C, and $M = 20.4$ for C+E), than those who said they had no control ($M = 21.8$ for C, and $M = 18.6$ for C+E). There was a significant interaction between instructions and control over heat, $F(1,132) = 9.07$, $p < 0.01$, suggesting larger difference in chosen temperature between C and C+E instructions when users had no control over their home’s heat. This is depicted in Figure 8.2.

³¹ I generated the each of the three mental-model questions myself, based on my review of the literature discussed in Section 2.1.2.

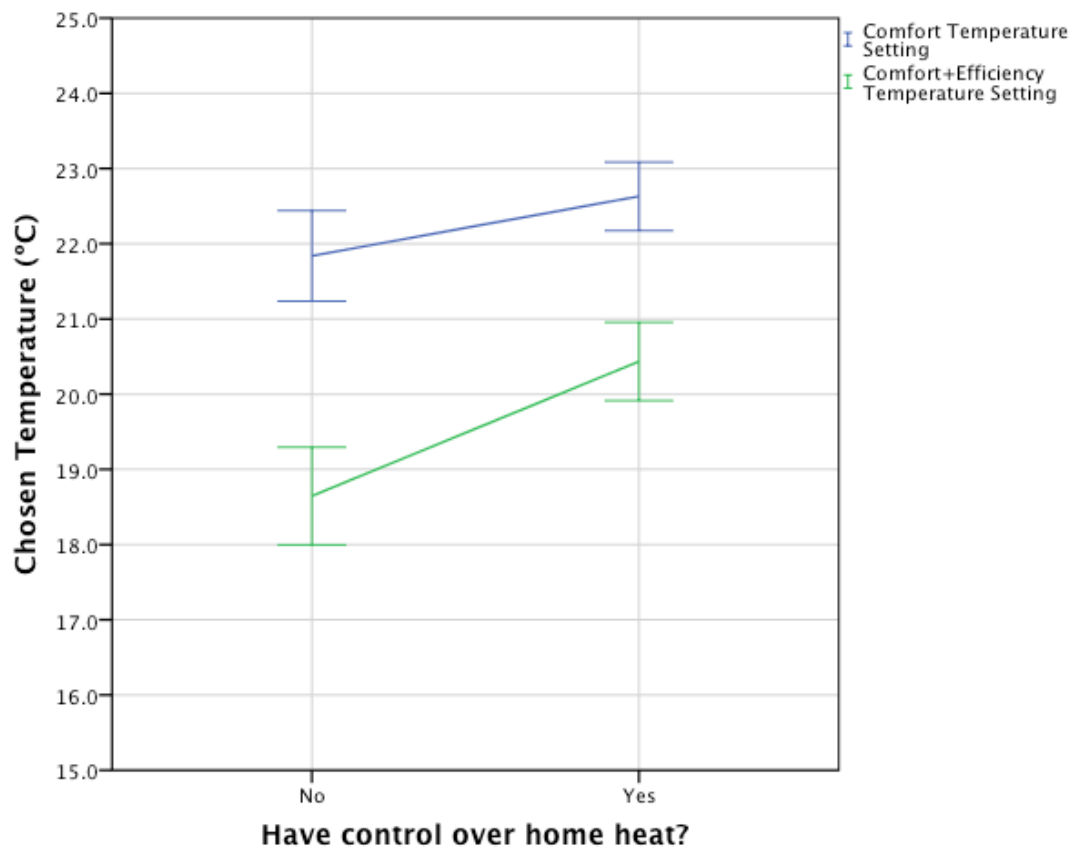


Figure 8.3 – User-selected temperature for participants with and without control over their H&C system. Error bars represent the 95% confidence intervals.

A two-way mixed ANOVA showed a significant main effect of the second mental-model check question³², $F(1,141) = 5.77$, $p < .05$. Participants who said heating is more efficient at higher temperatures also chose higher temperatures ($M = 23.1$ and 21.2 for C and C+E instructions) than those who correctly indicated it was more efficient at lower temperatures ($M = 22.1$ and 19.4 for C and C+E instructions). This is depicted in Figure 8.3.

³² “Heating is more energy-efficient at _____ temperature settings.” (lower/higher)

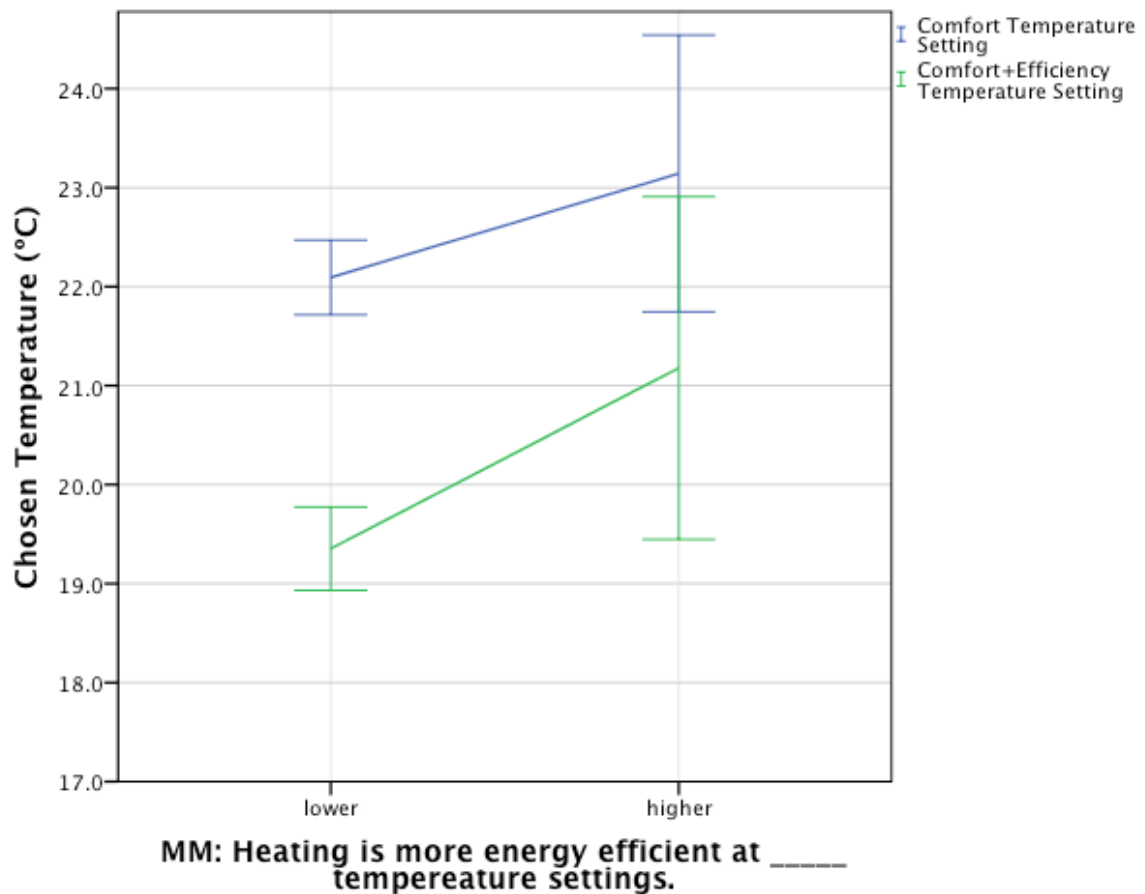


Figure 8.4 – User-selected temperature for mental-model question #2. Error bars represent the 95% confidence intervals.

A two-way mixed ANOVA showed a significant main effect of the third mental-model check question³³, $F(2,140) = 3.22$, $p < .05$. Post-hoc tests revealed that participants who were “unsure” about the Valve theory chose temperatures significantly lower ($M = 21.7$ and 19.0 for C and C+E instructions) than did both those who responded “true” ($M = 23.1$ and 19.9 for C and C+E instructions) or “false” ($M = 22.4$ and 19.9 for C and C+E instructions). There was no significant difference in responses between “true” and “false”. This is depicted in Figure 8.4.

³³ True or False: “A thermostat works like a valve: the more extreme the setting, the faster the temperature will change.”

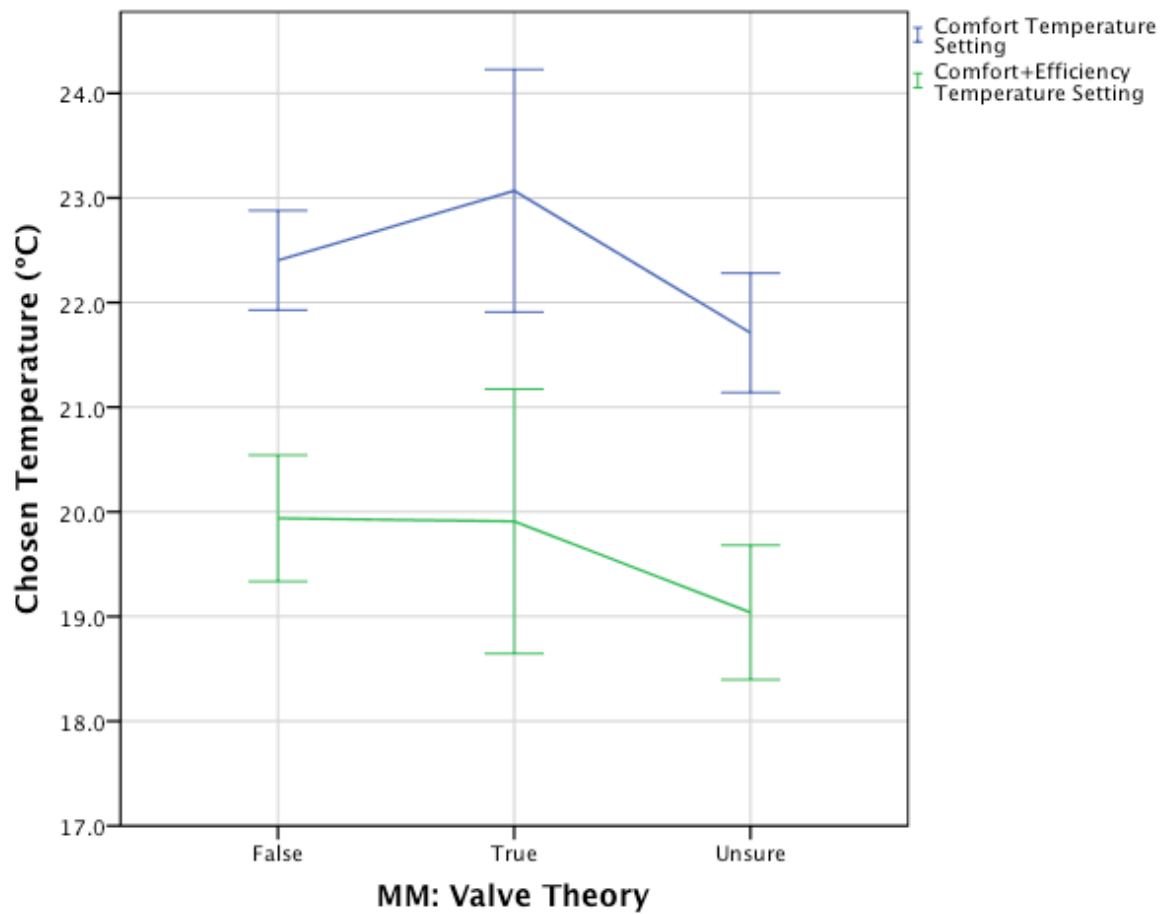


Figure 8.5 – User-selected temperature for mental-model question #3. Error bars represent the 95% confidence intervals.

8.7 Moving Forward

In this section, I presented the results from the temperature selection study. In Section 9.0, I provide a more in-depth analysis and discussion of these results.

9.0 Analysis and Discussion

In this section, I give my analysis, interpretation, and discussion of results presented in the previous section. I first examine the main effects on temperature setting and levels of participant confidence. Next, I look at the secondary results that stood out, including the results of the mental-model questions. I then summarize the limitations of this investigation. I finish with a conclusion and recommendations for future work.

The results suggest that enhanced control labelling had a significant and directional effect on user-selected temperatures. There is very strong evidence that the meaningful-non-social and injunctive-norm displays both influenced users to select significantly lower temperature settings than the scale-only display, and there is moderate evidence that the descriptive-norm display also had an effect, suggesting a general trend towards meaningful displays influencing the selection of lower temperatures. The anchor display showed no significant difference from scale-only, suggesting that any effect of display type was not due to any visual anchoring effect from the feature itself, but rather to each feature's meaningfulness. There was no significant interaction of instructions vs display type, suggesting that enhanced labels are effective regardless of the user's intention.

Two of the three meaningful display labels had a significant downward effect on temperature selection, namely the meaningful-non-social label ("Comfort Range") and the injunctive social norm label (The ☺ symbol, with subtext implying approval of houseguests). The descriptive social norm label ("90% of people comfortable") also showed moderate evidence of an effect on user-selected temperatures. I found this surprising, because the function of descriptive norms—in line with enhanced CL—is to provide information on what is appropriate and adaptive conduct in a particular situation. There are a few possible explanations for these observations:

- First, it is possible users were more sensitive to injunctive than to descriptive norms. If this were the case, participants may have viewed the "Comfort Range" as an injunctive norm, reflecting the experimenter's notion of the appropriate answer.

Further, the happy-face symbol may have implied the same sense of approval from the experimenters (rather than houseguests).

- Second, these two messages were the shortest and simplest, and thus most salient³⁴ of the four message conditions. Message brevity generally increases the likelihood of it being read (Laughery & Wogalter, in press), and greater salience generally improves the effectiveness of a normative message (Croy et al., 2010).
- Third, the descriptive norm may have been less effective because participants may not have believed it represented a 90%-acceptance range. Recall from Section 6.1.1 that this range is taken from the context of naturalistic field studies, *not* from asking people to control their thermal environment. While participants were not made aware of this during the experimental task, it is possible that they perceived such a strong claim (90% of people) to be inconsistent with the likely accepted norm of 21 or 22 °C. Perhaps reframing the message to “most (or many) people find comfortable” would seem more plausible and thus be more persuasive.
- Fourth, the descriptive norm’s message framing (“90% of people comfortable”) may have been too general or socially “distant”. Recall from Section 2.2.1 that norms tend to be more effective when the people feel more affiliation to or have a shared context with the described others (Berkowitz, 2004; Goldstein et al., 2007; Lewis & Neighbors, 2006; Neighbors et al., 2007). Perhaps a more specific framing like “90% of students in this building” would elicit a stronger effect.

Considering task instructions, I found that C+E instructions significantly lowered user-selected temperatures from C levels by 2.7°C on average³⁵. Although this is not surprising, it still suggests that the user’s intention to conserve energy creates more of a change in temperature settings than any influences from enhanced labelling that provides the comfort range. Given the C+E message framing that emphasized comfort³⁶, this also suggests that users are open to choosing significantly lower winter temperature settings without sacrificing their sense of comfort.

³⁴ Not considering the attached subtext, the happy-face symbol was the most salient message.

³⁵ C: 22.2°C vs C+E: 19.5°C, on average across all display types.

³⁶ “Choose the most energy-efficient temperature setting that you would be comfortable in.”

Another finding is that participants' levels of confidence in their selections were significantly lower when given C+E instructions (even though the median confidence was the same). This may be because people are generally more familiar with making decisions affecting their comfort than they are at improving their energy efficiency. There was no significant difference in confidence levels between display types, suggesting adding enhanced display labelling did not positively or negatively effect users' sense of confidence in their decision.

The questionnaire results suggest that participants who claimed to have no control over their home's thermostat chose significantly lower temperature settings than those who had some control. As discussed earlier, comfort research indicates that a sense of control does affect thermal comfort perception, influencing people to feel more comfortable than otherwise (Hedge et al., 2009). It is plausible that people with more control would be more comfortable at and thus choose lower temperatures, and so one might think that this finding is counter-intuitive; however, this effect does not dictate that people would choose lower temperatures. One interpretation is that these having-control users, supposedly with more thermostat-control experience (at least recently), may have learned that (in their homes) they need to set higher temperatures than they would otherwise estimate, due to perhaps inefficient home insulation (requiring more heat to rooms near the outside walls).

9.1 Mental Models

The mental-model questions helped illuminate a broad problem with people's understanding of how their systems affect their comfort and energy expenditures.

Similar to findings from Karjalainen & Vastamaeki (2007), most participants in this study incorrectly believed that comfortable indoor temperatures should be warmer (64%) or the same (23%) in the winter than as in the summer³⁷. This defies research and comfort standards showing that comfortable indoor temperatures tend to correspond to seasonal conditions rather than conflict with them (Brager & De Dear, 1998; Karjalainen &

³⁷ Recall, Karjalainen found 41% and 44% of participants falsely believed temperatures should be warmer or the same, respectively, in winter than in summer

Vastamaeki, 2007; Lovins, 1992). In my study, this may help explain why 1) display labelling had an effect on temperature selection (filling in the gaps of a flawed mental model for what is comfortable) and 2) average selected temperatures still hovered above or at the higher end of the 17-22°C comfort range when users considered only comfort.

Ten percent of participants incorrectly indicated that central heating would be more energy efficient at higher (rather than lower) temperatures³⁸, but strangely, *none of them chose a higher temperature with the C+E instructions*. Perhaps the wording of the question was misleading. “*Heating is more energy-efficient at _____ temperature settings.*” One explanation is these participants may have been confused around the word “settings”; some participants did report not understanding the question. If the interpretation was of higher outdoor temperatures, then indeed heating systems will use less energy to maintain a certain set point (minimizing the difference between indoor and outdoor temperatures). However, this does not explain why these individuals also chose significantly *higher* temperatures than those with the correct mental model. This requires further investigation.

Finally, in checking Kempton’s (1986) ‘Valve’-model, only 15% of participants believed—incorrectly—that it was true. This is promising compared to Kempton’s estimate of 25-50% of Americans holding ‘Valve’-theory (Kempton, 1986). However, 45% of participants indicated they were actually “unsure” if the statement “*A thermostat works like a valve: the more extreme the setting, the faster the temperature will change.*” was true or false³⁹. This may be due to an ambiguous interpretation of the statement (i.e. “valve” as an imperfect analogy) or a real lack of understanding. The results may have been much different if participants had been forced to choose between ‘true’ and ‘false’, or if there had been an additional instructions condition asking participants to demonstrate, say, adjusting the thermostat when they feel cold.

³⁸ This is not entirely surprising, given similar findings by Rathouse and Young (2004) mentioned in Section 2.1.3

³⁹ Kempton’s research was ethnographic and it appears that he did not classify any participants as being unsure, so a comparison cannot easily be made.

The analysis revealed that participants' answer to this question did account for a significant difference in user-selected temperatures, but the difference was between those who were and were not sure. It is unclear if this is a coincidence, or if there is some underlying phenomena that leads people who are unsure of how their thermostat works to choose lower temperatures. Further investigation is required.

9.2 Limitations

It is difficult to say how generalizable these observed effects are, due to the artificial nature of the experimental setup. Rather than interacting with a real touch-screen thermostat, in their home, controlling their home's heating system, participants used a marker to mark an X on a paper-prototype with no direct connection to their real systems. However, since participants did not convey a sense of difficulty with the task, it is fair to say I have a reasonable measure of their *intended* temperature settings for the given scenario; while execution may depend on the fidelity and usability of a real thermostat, this *intention* at least should be generalizable.

The primary limitation of this approach is the lack of thermal feedback from a heating system. Even if the enhanced labelling influenced a user to choose an appropriate lower temperature when programming their home thermostat, there is no evidence to say that the real-world feedback (ie. feeling too cold) would not undo this effect. For example, a user may initially choose a lower setting due to the enhanced label, but then override the initial program because they feel too cold.

However, I do think that, in a real-world setting, this enhanced labelling would still have a sustained influence. Since comfort-perception is highly subjective, I would predict that thermal feedback would have to compete with the other factors influencing comfort, and so it would not necessarily override the initial effect. If adjustments are made, I anticipate an anchoring effect on the programmed setting⁴⁰, so that overall, users would still choose lower temperatures than without the enhanced labelling. Also, programming typically sets a default, so that at very least, if a user is to manually alter the temperature, the controller

⁴⁰ Not to be confused with the non-significant anchoring effect on the comfort range.

would return to the pre-programmed setting soon after. Further investigation is required to see if the enhanced labelling would influence users of real-world systems in their manual changes to temperature settings.

It is also clear that the ASHRAE Adaptive Comfort Standard is taken out of context for my application, as can be seen from the literature (Brager & De Dear, 1998, 2000; De Dear & Brager, 1998). First, the standard addresses naturally ventilated buildings, not those that are centrally conditioned; as noted earlier, differences exist in thermal expectations between occupants of both building types. Second, the standard is based on naturalistic field studies, where participants—already exposed to a particular thermal condition—were queried on their resultant comfort level. Third, the standard is intended for the design of buildings, not for the control of heating or cooling systems.

However, I have also not found evidence to suggest that this model cannot be used for aiding control. It still provides some level of field-tested knowledge as to what temperatures people tend to be comfortable at, and thus may serve as a useful context to aid users in their decision-making processes. See Section 7.2 for my justification for using this particular comfort standard. Ultimately, I was more interested in the impact of an enhanced label with a suggested range than I was concerned about the numbers' validity.

A related limitation is that many regions have legal restrictions on minimum and maximum temperature settings (likely related to the engineering paradigm of thermal comfort). For instance, in rented residences in Ontario, landlords are required to keep winter temperatures above 20°C, ("Residential Tenancies Act, 2006 - O. Reg. 516/06," 2006). In Toronto, the requirement is above 21°C in winter, and below 26°C (when AC is available) in summer ("Toronto Municipal Code: Chapter 497," 2001). In both regulations however, there is an exception in cases where the tenant controls the heat ("Residential Tenancies Act, 2006 - O. Reg. 516/06," 2006, "Toronto Municipal Code: Chapter 497," 2001). This means that—in Ontario, unless regulations change—a thermostat equipped with enhanced labelling employing the 17 to 22 °C comfort range would be inappropriate in residences where the tenant has no heating control. There are still, however, enough owned and rented homes with resident control of heat to make such a design intervention

feasible in this region. Companies employing this approach should be aware of these legal restrictions and may also consider selling this feature in regions with less stringent regulations.

Another notable limitation was the choice of happy-face symbol for the injunctive norm condition. While it was inspired by similar social norms research (Schultz et al., 2007), it might have introduced too many messaging variables that may have confounded the results; it is thus unclear if the success of the injunctive-norm condition was due to the injunctive social norm, message brevity, or the symbolic representation.

On a lesser note, two of the mental-model questions were given a forced choice (“lower” vs “higher”), while another offered an “unsure” option. This inconsistency makes interpretation and comparison of these results less meaningful, and thus any conclusions drawn must be taken with some reservation.

9.3 Conclusions and Recommendations

Providing users of residential thermostats with enhanced thermostat control-labelling that meaningfully depicts a standard and seasonally-appropriate comfort-range appears to have encouraged them to select lower winter temperatures for a hypothetical heating system. This effect held regardless of whether their assigned motivation was solely comfort or also included energy efficiency.

The effect could not be explained by cognitive anchoring on the numbers themselves, and is thus attributed to the meaningfulness of the messaging.

The evidence suggests social norms were an effective form of messaging, but to what degree is still unclear. There was only moderate evidence that the descriptive norm produced a change from scale-only. The injunctive norm condition did have a significant effect on temperature settings. It is, however, unclear if this was due to its injunctive nature, or to its brevity and saliency. Regardless, the simple meaningful-non-social messaging “Comfort Range” produced a significant downward change of 1.5 °C in temperature setting.

Future Research and Recommendations

Future research could investigate the observed effects further in a few ways. First, more research should also be undertaken on message framing in control labelling, continuing to look at the effects of social norms; in particular it is important to see what effect perceived social proximity has on the efficacy of a normative CL. It is important to see if these observations would hold in different contexts. For example, the study could be run again in the context of choosing summer temperatures, since the equivalent summer comfort range is actually above (rather than below) typical settings. Second, to improve validity, a higher-fidelity prototype, in a more real-world simulated setting should be used to evaluate the effect of enhanced labelling. Further, it would be interesting to run a field study, connecting a fully-functional thermostat prototype to a real home's H&C system, to see if and how the observed effect interacts with the real system's thermal feedback to influence user's temperature selection choices; I am not aware of similar research at this time. Third, investigators should test conditions where the injunctive norms are as similar as possible to the descriptive norm and non-social messaging conditions.

10.0 References

A Municipal Guide on Economic Instruments to Support Municipal Waste Management

Programs. (1996). Federation of Canadian Municipalities.

Abrahamse, W., Steg, L., Vlek, C., & Rothengatter, T. (2005). A review of intervention studies aimed at household energy conservation. *Journal of Environmental Psychology*, 25(3), 273–291. doi:10.1016/j.jenvp.2005.08.002

Baekdal, T. (2005). Usability Severity Rating - Improved. Retrieved March 25, 2013, from <http://www.baekdal.com/insights/usability-severity-rating>

Bailey, J. J., Nofsinger, J. R., & O'Neill, M. (2004). 401(K) Retirement Plan Contribution Decision Factors: The Role of Social Norms. *Journal of Business and Management*. Retrieved from <http://www.highbeam.com/doc/1P3-629392061.html>

Berkowitz, A. (2004). The Social Norms Approach: Theory, Research, and Annotated Bibliography. Publisher: Author.

Black, J. S., Stern, P. C., & Elworth, J. T. (1985). Personal and contextual influences on household energy adaptations. *Journal of Applied Psychology*, 70(1), 3–21. doi:10.1037/0021-9010.70.1.3

Boff, K. R., & Lincoln, J. E. (1988). *Engineering Data Compendium. Human Perception and Performance. Volume 3*.

Brager, G. S., & De Dear, R. (1998). Thermal adaptation in the built environment: a literature review. *Energy and Buildings*, 27(1), 83–96. doi:10.1016/S0378-7788(97)00053-4

Brager, G. S., & De Dear, R. (2000). A standard for natural ventilation. *ASHRAE journal*, 42(10), 21–28.

- Carlsson-Kanyama, A., & Lindén, A.-L. (2007). Energy efficiency in residences—Challenges for women and men in the North. *Energy Policy*, 35(4), 2163–2172.
doi:10.1016/j.enpol.2006.06.018
- Cialdini, R. B. (2007). Descriptive Social Norms as Underappreciated Sources of Social Control. *Psychometrika*, 72(2), 263–268.
- City of Toronto. (2011). 2011 Residential Waste Diversion. City of Toronto. Retrieved from <http://www.toronto.ca/garbage/pdf/2011-chart.pdf>
- Climate Data Online. (n.d.). *National Climate Data and Information Archive*. Retrieved from http://climate.weatheroffice.gc.ca/climateData/canada_e.html
- Combe, N., Harrison, D., Craig, S., & Young, M. S. (2012). An investigation into usability and exclusivity issues of digital programmable thermostats. *Journal of Engineering Design*, 23(5), 401–417. doi:10.1080/09544828.2011.599027
- Croy, G., Gerrans, P., & Speelman, C. (2010). Injunctive social norms primacy over descriptive social norms in retirement savings decisions. *International journal of aging & human development*, 71(4), 259–282.
- De Dear, R., & Brager, G. S. (1998). Developing an adaptive model of thermal comfort and preference. Retrieved from <http://www.escholarship.org/uc/item/4qq2p9c6>
- Ehrhardt-Martinez, K., Donnelly, K. A., & Laitner, J. A. (2010). *Advanced Metering Initiatives and Residential Feedback Programs: A Meta-Review for Household Electricity-Saving Opportunities* (Research Report No. E105). American Council for an Energy-Efficient Economy. Retrieved from <http://aceee.org/research-report/e105>
- Eyssel, F., Bohner, G., & Siebler, F. (2006). Perceived rape myth acceptance of others predicts rape proclivity: Social norm or judgmental anchoring? *Swiss Journal of Psychology/Schweizerische Zeitschrift für Psychologie/Revue Suisse de Psychologie*, 65(2), 93–99. doi:10.1024/1421-0185.65.2.93

- Field, A. (2009). *Discovering Statistics Using SPSS* (3rd ed.). Sage Publications.
- Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics*, 40(1), 35–42.
- Goldstein, N. J., Griskevicius, V., & Cialdini, R. B. (2007). Invoking Social Norms A Social Psychology Perspective on Improving Hotels' Linen-Reuse Programs. *Cornell Hotel and Restaurant Administration Quarterly*, 48(2), 145–150. doi:10.1177/0010880407299542
- Government of Canada, N. R. C. (2013a, January 16). Residential Sector Canada Table 2: Secondary Energy Use and GHG Emissions by End-Use. Retrieved February 26, 2013, from <http://oee.nrcan.gc.ca/corporate/statistics/neud/dpa/showTable.cfm?type=CP§or=res&juris=ca&rn=2&page=4&CFID=29535801&CFTOKEN=48b3760032f26f36-1725C0D0-0828-0833-9C9D1C6C145DF715>
- Government of Canada, N. R. C. (2013b, January 17). Total End-Use Sector - Energy Use Analysis. Retrieved February 26, 2013, from <http://oee.nrcan.gc.ca/corporate/statistics/neud/dpa/showTable.cfm?type=AN§or=aaa&juris=ca&rn=1&page=4&CFID=29535792&CFTOKEN=c7ae6a04e91cf82e-172219CE-AC0C-5862-3B7D78059454EC9A>
- Government of Canada, N. R. C. (2013c, January 17). Total End-Use Sector – GHG Emissions. Retrieved February 26, 2013, from <http://oee.nrcan.gc.ca/corporate/statistics/neud/dpa/showTable.cfm?type=AN§or=aaa&juris=00&rn=1&page=4&CFID=29535801&CFTOKEN=48b3760032f26f36-1725C0D0-0828-0833-9C9D1C6C145DF715>
- Hargreaves, T., Nye, M., & Burgess, J. (2010). Making energy visible: A qualitative field study of how householders interact with feedback from smart energy monitors. *Energy Policy*, 38(10), 6111–6119. doi:10.1016/j.enpol.2010.05.068

- Hedge, A., Khalifa, H. E., & Zhang, J. (2009). On the Control of Environmental Conditions Using Personal Ventilation Systems. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 53(8), 536–540.
doi:10.1177/154193120905300811
- Ignite Toronto 2: Matt Ratto - Why Johnny Can't Read an iPhone: Using critical making to teach socio-technical literacy. (2009). Retrieved from <http://vimeo.com/8284854>
- Karjalainen, S. (2007). *The Characteristics of Usable Room Temperature Control*. VTT.
- Karjalainen, S., & Vastamaeki, R. (2007). Occupants Have a False Idea of Comfortable Summer Season Temperatures. In *Proceedings of Clima 2007 WellBeing Indoors*. Helsinki (Finland): FINVAC. Retrieved from <http://www.irb.fraunhofer.de/CIBlibrary/search-quick-result-list.jsp?A&idSuche=CIB+DC7482>
- Karlin, B., Au, A., Goleva, E., & Zinger, J. (2011, December). *Residential energy feedback: A Meta-analysis and methodological review*. Poster presented at the Behavior Energy and Climate Change (BECC) Conference.
- Kempton, W. (1986). Two theories of home heat control. *Cognitive Science*, 10(1), 75–90.
- Kempton, W., Feuermann, D., & McGarity, A. E. (1992). “I always turn it on super”: user decisions about when and how to operate room air conditioners. *Energy and Buildings*, 18(3–4), 177–191. doi:10.1016/0378-7788(92)90012-6
- Kempton, W., & Montgomery, L. (1982). Folk quantification of energy. *Energy*, 7(10), 817–827. doi:10.1016/0360-5442(82)90030-5
- Kempton, W., Reynolds, C., Fels, M., & Hull, D. (1992). Utility control of residential cooling: resident-perceived effects and potential program improvements. *Energy and Buildings*, 18(3–4), 201–219. doi:10.1016/0378-7788(92)90014-8
- Kort, de Y., McCalley, L. T., & Midden, C. J. H. (2008). Persuasive trash cans : activation of littering norms by design. *Environment and Behavior*, 40(6), 870null.

- Kuijer, L., & De Jong, A. (2011). Exploring Practices of Thermal Comfort for Sustainable Design. Presented at the CHI 2011.
- Kurz, T., Donaghue, N., & Walker, I. (2005). Utilizing a Social-Ecological Framework to Promote Water and Energy Conservation: A Field Experiment. *Journal of Applied Social Psychology*, 35(6), 1281–1300. doi:10.1111/j.1559-1816.2005.tb02171.x
- Laughery, K. R., & Wogalter, M. S. (In Press). A three-stage model summarizes product warning and environmental sign research. *Safety Science*. doi:10.1016/j.ssci.2011.02.012
- Lee, J. D. (2006). Human Factors and Ergonomics in Automation Design. In G. S. P. D. C., editor, and Head of the Department (Ed.), *Handbook of Human Factors and Ergonomics* (pp. 1570–1596). John Wiley & Sons, Inc. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1002/0470048204.ch60/summary>
- Lewis, M. A., & Neighbors, C. (2006). Social Norms Approaches Using Descriptive Drinking Norms Education: A Review of the Research on Personalized Normative Feedback. *Journal of American college health : J of ACH*, 54(4), 213–218.
- Lombardi, M. M., & Choplin, J. M. (2010). Anchoring and Estimation of Alcohol Consumption: Implications for Social Norm Interventions. *Journal of Alcohol & Drug Education*. Retrieved from <http://www.highbeam.com/doc/1P3-2125383991.html>
- Lovins, A. B. (1992). *Air Conditioning Comfort: Behavioral and Cultural Issues*. E Source.
- Lutzenhiser, L. (1993). Social and Behavioral Aspects of Energy use. *Annual Review of Energy and the Environment*, 18(1), 247–289. doi:10.1146/annurev.eg.18.110193.001335
- McCalley, L. T., & Midden, C. J. H. (2004). Goal Conflict and User Experience: Moderators to the Use of the Clock Thermostat as a Device to Support Conservation Behavior. In *Proceedings of the summer study on energy efficiency in buildings, American council for an energy efficient economy* (pp. 251–259). Retrieved from http://www.eceee.org/conference_proceedings/ACEEE_buildings/2004/Panel_7/p7_22

- McKenzie-Mohr, D. (2011). *Fostering Sustainable Behavior: An Introduction to Community-Based Social Marketing* (Third Edition.). New Society Publishers.
- McQuiston, F. C., Parker, J. D., & Spitler, J. D. (2004). *Heating, Ventilating and Air Conditioning Analysis and Design* (6th ed.). Wiley.
- Meier, A., Aragon, C., Pepper, T., Perry, D., & Pritoni, M. (2011). Usability of residential thermostats: Preliminary investigations. *Building and Environment*, 46(10), 1891–1898. doi:10.1016/j.buildenv.2011.03.009
- Military Standard 1472D: Human Engineer Design Criteria for Military Systems, Equipment, and Facilities. (1989). US Department of Defence. Retrieved from http://www.everyspec.com/MIL-STD/MIL-STD-1400-1499/MIL_STD_1472D_1209/
- Neighbors, C., Lostutter, T. W., Whiteside, U., Fossos, N., Walker, D. D., & Larimer, M. E. (2007). Injunctive norms and problem gambling among college students. *Journal of gambling studies / co-sponsored by the National Council on Problem Gambling and Institute for the Study of Gambling and Commercial Gaming*, 23(3), 259–273. doi:10.1007/s10899-007-9059-3
- Next Learning Thermostat. (2013). *Nest Labs*. Retrieved March 9, 2013, from <http://www.nest.com/>
- Nielsen, J. (1994). *Usability Engineering* (1st ed.). Morgan Kaufmann.
- Norman, D. A. (2002). *The Design of Everyday Things*. Basic Books.
- O'Hara, J. M., Brown, W. S., Lewis, P. M., & Persensky, J. . (2002). NUREG-0700 Human-System Interface Design Review Guidelines, rev 2. U.S. Nuclear Regulatory Commission. Retrieved from <http://www.nrc.gov/reading-rm/doc-collections/nuregs/staff/sr0700/>
- Parker, D., Barkaszi, S., Sherwin, J., & Richardson, C. (1996). Central Air Conditioner Usage Patterns in Low-Income Housing in a Hot and Humid Climate: Influences on Energy

- Use and Peak Demand. In *Proceedings of the 1996 ACEEE Summer Study on Energy Efficiency in Buildings*. Retrieved from <http://cgec.ucdavis.edu/ACEEE/1994-96/1996/VOL08/147.PDF>
- Peffer, T., Pritoni, M., Meier, A., Aragon, C., & Perry, D. (2011). How people use thermostats in homes: A review. *Building and Environment*, 46(12), 2529–2541.
doi:10.1016/j.buildenv.2011.06.002
- Pineau, C. (1982). The psychological meaning of comfort. *Applied Psychology*, 31(2), 271–282.
doi:10.1111/j.1464-0597.1982.tb00097.x
- Rathouse, K., & Young, B. (2004). *BNDH15: Use of domestic heating controls* (No. BNDH15). UK: Department for Environment Food and Rural Affairs (DEFRA).
- Ratto, M. (2011). Critical Making: Conceptual and Material Studies in Technology and Social Life. *The Information Society*, 27(4), 252–260. doi:10.1080/01972243.2011.583819
- Residential Tenancies Act, 2006 - O. Reg. 516/06. (2006). Retrieved May 1, 2013, from http://www.e-laws.gov.on.ca/html/regs/english/elaws_regs_060516_e.htm#BK0
- Sanders, M. S., & McCormick, E. J. (1993). *Human Factors In Engineering and Design* (7th ed.). McGraw-Hill Science/Engineering/Math.
- Sauer, J., Franke, H., & Ruettinger, B. (2008). Designing interactive consumer products: utility of paper prototypes and effectiveness of enhanced control labelling. *Applied ergonomics*, 39(1), 71–85. doi:10.1016/j.apergo.2007.03.001
- Sauer, J., & Rüttinger, B. (2007). Automation and decision support in interactive consumer products. *Ergonomics*, 50(6), 902–919. doi:10.1080/00140130701254266
- Sauer, J., Wiese, B. S., & Rüttinger, B. (2002). Improving ecological performance of electrical consumer products: the role of design-based measures and user variables. *Applied ergonomics*, 33(4), 297–307.

- Sauer, J., Wiese, B. S., & Rüttinger, B. (2003). Designing low-complexity electrical consumer products for ecological use. *Applied ergonomics*, 34(6), 521–531.
doi:10.1016/j.apergo.2003.07.001
- Sauer, J., Wiese, B. S., & Rüttinger, B. (2004). Ecological performance of electrical consumer products: the influence of automation and information-based measures. *Applied ergonomics*, 35(1), 37–47. doi:10.1016/j.apergo.2003.09.001
- Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological science*, 18(5), 429–434. doi:10.1111/j.1467-9280.2007.01917.x
- Seligman, C., Darley, J. M., & Becker, L. J. (1978). Behavioral approaches to residential energy conservation. *Energy and Buildings*, 1(3), 325–337. doi:10.1016/0378-7788(78)90012-9
- Shipworth, M., Firth, S. K., Gentry, M. I., Wright, A. J., Shipworth, D. T., & Lomas, K. J. (2010). Central heating thermostat settings and timing: building demographics. *Building Research & Information*, 38(1), 50–69. doi:10.1080/09613210903263007
- Smart Thermostat Quick Start Guide. (2008). ecobee Inc. Retrieved from <http://ebookbrowse.com/ecobee-smart-thermostat-quick-start-guide-pdf-d192876149>
- Stein, J. (2011). *Usability-Test Plan: ecobee Smart Thermostat and Accessories* (Usability Plan). Toronto, Canada: University of Toronto.
- Stein, J., Borg, M., Ratto, M., & Jamieson, G. A. (2011). *Usability-Test Report: ecobee Smart Thermostat and Accessories* (Usability Report). Toronto, Canada: University of Toronto.
- Subjective measures. (n.d.). *Comfortable Low Energy ARchitecture*. Retrieved November 19, 2012, from <http://www.new-learn.info/packages/clear/thermal/people/surveying/clomet.html>

- Switzer, F. S., & Sniezek, J. A. (1991). Judgment processes in motivation: Anchoring and adjustment effects on judgment and behavior. *Organizational Behavior and Human Decision Processes*, 49(2), 208–229.
- Syrek, D. B., & Michigan. Legislature. (1980). Michigan litter: after: a study of the impact of beverage container deposit legislation on street, roadside and recreation area litter in Michigan. Retrieved March 9, 2013, from <http://books.google.com/books?id=PpIQAQAAMAAJ>
- Thermostatic Radiator Valve and standard radiator valves. (2012). *Emergency Plumber Blog*. Retrieved March 19, 2013, from <http://www.plumber24hours.co.uk/plumber-blog/index.php/2012/12/21/thermostatic-radiator-valve/>
- Toronto Municipal Code: Chapter 497. (2001).
- Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 185(4157), 1124–1131.
- Verhallen, T. M. M., & Van Raaij, W. F. (1981). Household Behavior and the Use of Natural Gas for Home Heating. *Journal of Consumer Research*, 8(3), 253–57.
- Verplanken, B., & Wood, W. (2006). Interventions to break and create consumer habits. *Journal of Public Policy & Marketing*, 90–103.
- Whyte, G., & Sebenius, J. K. (1997). The Effect of Multiple Anchors on Anchoring in Individual and Group Judgment. *Organizational Behavior and Human Decision Processes*, 69(1), 74–85.
- Wickens, C. D., & Hollands, J. G. (1999). *Engineering Psychology and Human Performance* (3rd ed.). Pearson.
- Wiese, B. S., Sauer, J., & Rüttinger, B. (2004). Consumers' use of written product information. *Ergonomics*, 47(11), 1180–1194. doi:10.1080/00140130410001695951

Wilson, C. (1999). Reader's Questions: Severity Scales. Retrieved March 25, 2013, from
<http://www.stcsig.org/usability/newsletter/9904-severity-scale.html>

Woodson, W., Tillman, P., & Tillman, B. (1992). *Human Factors Design Handbook* (2nd ed.).
McGraw-Hill Professional.

Appendix A – User Recruitment Letter

“Invitation to Usability Study: energy-feedback technologies”

Dear ThingTank members,

As part of a research project investigating the socio-economic implications of emerging innovations in energy monitoring and data visualization, we would like to obtain a picture of the current state-of-the-art in the usability of energy-feedback technologies. We recognize that the ThingTank community embodies the demographics required to gain an accurate understanding of human-computer interaction issues within this context. As such, we would like to invite you to participate in a usability study.

This study should take place in the **month of October**, according to your availabilities, and **should not take longer than one hour**. The location of the study will be at the ThingTank Lab at 376 Bathurst St, Toronto, ON.

We are well aware that some of the conversations may contain sensitive or confidential information. Any confidential or identifying information will be made anonymous and only the research team will have access to the information collected. I would like to assure you that the study has been reviewed and received ethics clearance through the Office of Research Ethics at the University of Toronto. It should also be disclosed that this project is being carried out in partnership with Ecobee, who's mission is to help homeowners and businesses conserve energy, save money and reduce their environmental impact. Note that all audio and video recordings will be encrypted as per the University of Toronto's standards.

The following is a brief description of the demographic profile of participants. **You may still be considered if you meet some but not all criteria.

- between 30-55 years old
- lives in a household with 2-4 people (participant included)
- total household income exceeds \$100k
- typically on the internet 2+ hours each day
- familiar with touch-screen devices
- pays their own home electricity and heating bills and has control over their thermostat
- is not an HVAC professional or usability expert

If you or someone you know would be an ideal candidate for this study please contact Josh Stein.

Refer to the following links for more information:

- <http://www.research.utoronto.ca/ethics/pdf/human/nonspecific/datasecurity.pdf>
- http://www.utoronto.ca/security/UTORprotect/encryption_guidelines.htm

If you are interested in participating, or would like to further discuss this letter, please do not hesitate to contact me at matt.ratto@utoronto.ca.

Thank you for your consideration,

Signature:

Matt Ratto
Assistant Professor
Director, Critical Making Lab
Faculty of Information
University of Toronto

Appendix B – Letter of Consent for Usability Study

Letter of Consent: Research Interview

Research project title: Usability and Design of Ecobee Accessory Products for Residential and Commercial Energy Use.

Principal investigator:

Dr. Matt Ratto

Assistant Professor, Faculty of Information, University of Toronto

416-946-5415

matt.ratto@utoronto.ca

Summary of the research project

Monitoring energy consumption has been a long time necessity for utility companies; however that data has largely been gathered at the unit of a household level. Ecobee, has developed and commercialized a WiFi enabled ecobee Smart Thermostat (ST) for homeowners and the Energy Management System for commercial applications. These technologies, along with commercial accessories will enable more granular energy monitoring capabilities. When these granular data collection techniques are combined with rich visualization tools there is the possibility that innovative standards of practice will emerge.

We posit that increased awareness of granular and personalized energy consumption habits as facilitated by smarter technology and visualization tools will promote greater concern for reducing energy consumption habits. This awareness will have value within residential and commercial properties, and will assist in mitigating the inefficiencies that exist as a result of imperfect or nonexistent information.

The objective of this study is to assess the relationship between energy consumption and the communication of its granular metrics at the household and commercial levels. Greater understanding of the behavioral and habitual processes that surround issues of energy consumption could lead to decreasing inefficiencies. The potential for bottom-up habitual change has potential value that has been scarcely researched within the realm of human-computer interaction, data visualization and electricity consumption.

Invitation to participate and respect of ethical principles

As a significant contributor in this area, you are invited to participate in this study. We would like you to participate in an interview session, to be conducted via online videoconference (Skype) or in person.

Please be assured that all collected information will be treated confidentially. Recordings, transcripts and notes from the session will be kept in a secure location for at most two years and then destroyed.

While information collected during the interview may be used for publication purposes, the anonymity of individual participants can be ensured upon request by designating stakeholders by surrogate names in interview transcripts, field notes, and any other collected data. Considering the focus of the study and the confidentiality measures that will be taken, participating in this study should not cause you any prejudice.

You should not, under any circumstance, feel obliged to participate. Moreover, once you have agreed to participate, you may withdraw at any time without further justification. Your participation

should be completely voluntary. You may also decline to answer any question during the interview session.

If you require further information on any of the above, please feel welcome to contact principal investigator Matt Ratto. If you want more information regarding your rights as a participant, you may contact the Office of Research Ethics by phone at 416-946-3273 or by email at ethics.review@utoronto.ca. You may also keep a copy of this letter for your records.

Participant signature

Having read and understood the above text, and having had the possibility to ask and receive complementary information on the study, I consent to participate to the following activities in this research:

- Research Interview: _____

I would like to receive a copy of the final report document and a copy of any published materials using results from the study:

- Yes _____ No _____

Participant name: _____

Participant signature: _____

Date: _____

Appendix C – User Aides

This image describes the most important ecobee Smart-Thermostat controls. This material has been removed due to copyright restrictions. It can be found here:
<http://www.ecobee.com/wp-content/themes/ecobee/etc/file/web-QS-30-03-093.pdf>

Figure B.10.1 – Using the Touch Screen

This image describes the ecobee Smart-Thermostat Feature buttons: Register, Weather, Details, Quick Save, Program, and Vacation. This material has been removed due to copyright restrictions. It can be found here:
<http://www.ecobee.com/wp-content/themes/ecobee/etc/file/web-QS-30-03-093.pdf>

Figure B.10.2 – Feature Buttons

Appendix D – List of Tasks

Table D-10-1 – List of Tasks in the Usability Study

Task	Task Description
1	Report the current indoor temperature.
2	Report the current outdoor temperature
2a	Return to the home screen
3	Turn on the furnace.
4	Create a new program using the Program Wizard, using your typical workday, and choose your own preferred temperature settings when prompted.
5a	Edit-Task: On Mondays and Wednesdays, you don't come home until 7:30.
5b	Edit-Task: On Tuesdays and Thursdays, you come home for lunch, 12:30 to 1:30, then return to work.
6	[alter current setpoint] Imagine it is currently too cold (cold enough to wear a thick sweater); adjust the thermostat so you are more comfortable.
6a	Return to the preprogrammed temperature set point [from Hold]
7	Enter Quick-Save mode (note: task ends in quick save mode)
7b	Interpret
8	Install the 2nd temperature sensor
9	Check the 2nd temperature sensor
10x	Find the Smart Plug Dashboard
10	Turn a smart plug off (or on)
10a	Turn the smart plug back on (or off)
11	Set a smart-plug program so it is turned on only when you are at home, consistent with your temperature program.
12	[Interpret]Compare and interpret usage of two plugs from the smart-plug dialogue
13	Find hourly electrical use data (note: task ends in hourly report dialogue)
13a	Interpret

Appendix E – Usability Study Participant Questionnaire

Please answer the following questions honestly for demographics purposes. Your answers will be kept completely confidential and you will not be identified by name.

Age: _____

Occupation: _____

Number of people currently living in your household (including you): _____

Total household income:

- ☐ < \$50k
- ☐ \$50-100k
- ☐ \$100-150k
- ☐ > \$150k
- ☐ Prefer not to say

Number of hours typically spent on the Internet each day: _____

Are you currently responsible for your home's electric and heating bills? _____(yes/no)

Do you have control over your home thermostat? _____(yes/no)

Are you an expert in HVAC or usability? _____(yes/no)

Do you have any previous experience/knowledge of the ecobee Smart Thermostat?

_____.

Describe your current thermostat: _____

_____.

(analogue vs. digital display / programmable? /straight-line vs circular dial / etc)

If price were not an issue, would you prefer to use ecobee's thermostat to your current one?

_____ (yes/no)

Participant Code: _____

Appendix F – Quantitative Data

Table F-10-2– Errors and task time, by task, for each participant – dashes indicate tasks not recorded/performed

Task	P1 Time	P2 Time	P3 Time	Novice Average	Expert Time	P1 Errors	P2 Errors	P3 Errors	Novice Average	Expert Errors
1	0:20	0:02	-	0:11	-	0	0	-	0.0	-
2	0:15	0:03	-	0:09	-	0	0	-	0.0	-
2a	0:05	-	-	0:05	-	0	-	-	0.0	-
3	0:58	1:14	-	1:06	-	4	3	-	3.5	-
4	3:10	2:35	1:37	2:27	0:57	5	2	3	3.3	2
5a	1:32	1:00	0:44	1:05	1:28	3	2	1	2.0	3
5b	2:39	3:14	3:46	3:13	1:55	5	6	8	6.3	7
6	0:21	-	-	0:21	0:10	1	-	-	1.0	0
6a	0:20	0:07	0:30	0:19	0:09	0	1	2	1.0	1
7	0:11	0:05	0:06	0:07	0:05	1	1	1	1.0	0
8	3:59	9:40	-	6:49	1:55	3	8	-	5.5	6
9	0:17	-	0:44	0:30	0:05	0	-	3	1.5	0
10x	0:16	0:11	-	0:13	-	0	1	-	0.5	-
10	0:28	0:09	0:20	0:19	0:16	1	0	0	0.3	0
10a	0:20	0:14	-	0:17	0:13	1	1	-	1.0	1
11	2:04	-	-	2:04	6:30	3	-	-	3.0	9
13	0:20	-	-	0:20	0:10	0	-	-	0.0	0

Appendix G – Subjective Usability Metrics

Each usability issue was rated by severity (1: minor, 4:critical), expected frequency (1: rare issue, 4: frequent issue), and test occurrence (% of participants effected). The priority score describes the product of these three metrics.

$$\text{Priority Score} = \text{Severity} \times \text{Expected Frequency} \times \text{Occurrence}$$

Severity⁴¹: how bad are the implications?

1. **Minor**: problem is rare and causes no data loss or major time loss. Minor cosmetic or consistency issue. Issues like cosmetic errors, spelling problems, non-critical workflow issues
2. **Moderate**: Minor but irritating problem. No data loss but the problem slows users down slightly, minimal violations of guidelines that affect appearance or perception, and mistake that are recoverable.
3. **Major**: Moderate problem causing wasted time, but no *permanent* data loss. A workaround exists. Internal inconsistencies result in increased learning/error rates. An important function/feature does not work as expected. Loss of functionality, problematic impact user's workflow
4. **Critical**: Severe problem causing possible loss of data. User has no workaround to the problem. Performance is so poor that the system is universally regarded as 'pitiful'. System crashes, workflows breaks down, complete loss of focus for a specific task. Loss of information

Expected Frequency: how often will this be encountered in regular usage?

1. **Rare**
2. **Occasionally**
3. **Somewhat Often**
4. **Very Often**

Occurrence: of those novices who could have encountered this issue, what fraction (between 0 and 1) of participants were affected? Includes the expert only for instances wherein he encountered the same problem.

⁴¹ I created this scale by combining two existing scales, from Wilson (1999) and Baekdal (2005)

Appendix H – Full List of Qualitative Usability Issues

Table H-10-3 – Qualitative Usability Issues – Category: Confusion

Issue	% Occurrence	Severity	Expected Frequency	Priority Score
Not sure what "hold" means.	100%	3	4	12.0
Quick-Save mode confusion: what does "set-point" mean? Why 2.2° and what does it have to do with set-point and current temperature?	100%	3	3	9.0
Confuses current with set temperature.	67%	3	4	8.0
Does altering the temperature change it now or in the future?	67%	3	4	8.0
Did not notice accidental change in editor overview chart.	50%	3	3	4.5
Does not know what bottom-right indicator icons mean (other than wifi).	100%	1	3	3.0
Sensors: "EI" vs "Remote Sensor Board"?	100%	3	1	3.0
No idea what "Follow Thermostat" means with respect to a plugged-in device.	100%	3	1	3.0
Not sure if Current temperature is indoor or outdoor, because it's above the weather icon.	33%	2	4	2.7
FIRE symbol is very unsettling; seems like something is wrong or burning.	33%	2	4	2.7
Gets lost in Settings menu structure/hierarchy.	50%	2	2	2.0
Sensors: "Dry Contact" vs "Temperature"?	50%	3	1	1.5
Unclear that "System" is the furnace.	33%	2	2	1.3
"Hold event ends at ____" sounds like engineer speak.	33%	1	4	1.3
Sensor screen very confusing.	33%	2	2	1.3
Not clear, in Smart Plug Dashboard, that plugs can be switched on/off.	33%	2	2	1.3
Unclear if an item pre-highlighted in a choice menu is selected or not.	67%	2	1	1.3
Diff' between "Monitoring" vs "Control" sensor? Wouldn't they overlap?	50%	2	1	1.0
Do 2 sliders indicate a range or 2 separate controls, in selecting temp in Wizard?	33%	3	1	1.0
After done with Wizard, "Congratulations" screen is confusing. Wants to see programmed times/temperatures.	33%	2	1	0.7
"Is the area occupied?" sounds overly formal.	33%	1	1	0.3

Table H-10-4 – Qualitative Usability Issues – Category: System Errors

Issue	% Occurrence	Severity	Expected Frequency	Score
Unresponsive screen buttons.	100%	3	4	12.0
Editor: trying to select "Add New Item" accidentally goes to editing bottom list item.	100%	3	2	6.0
Temperature selector erratically 'jumps' around, making precise selection difficult.	67%	2	4	5.3
Time wheel selector spins erratically, making selection difficult	75%	2	2	3.0
Editor: hard to select "Add New Item", because screen shifts up/down.	50%	3	2	3.0
Keyboard is error-prone	100%	2	1	2.0
"Add New Item" selects (goes blue) but no action follows. Functionality freezes.	33%	3	1	1.0
Selected a time "to go to work", clicked "Next", screen did not advance, but time wheel scrolled on its own to 5am!	33%	3	1	1.0

Table H-10-5 – Qualitative Usability Issues – Category: User Errors

Issue	% Occurrence	Severity	Expected Frequency	Score
Pressed OK in Quick-Save mode, not realizing it disables QS.	100%	3.5	3	10.5
Not clear to press "Resume" to return to present temp. Moved temp slider instead.	67%	3	4	8.0
Pressed physical button to finish. Lost changes.	75%	3.5	3	7.9
In turning a plug off, did not hit "Resume", but chose "indefinite hold".	100%	2.5	3	7.5
Pressed Weather icon to see outdoor temp, despite it written on the icon itself.	67%	1	3	2.0
Presses "Edit" button despite seeing that pressing the Program schedule diagram does the same (and easier).	100%	1	2	2.0
Goes to the Sensor screen to install the sensor	100%	2	1	2.0
Looked for 2nd temperature-sensor reading in Details screen.	33%	2	2	1.3
Pressed plug-switch icon (in the Plug Dashboard), looking to set a program.	33%	2	2	1.3
Pressed "More" icon accidentally instead of "Program"	25%	2	2	1.0

Table H-10-6 – Qualitative Usability Issues – Category: Action Difficulty

Issue	% Occurrence	Severity	Expected Frequency	Score
They thought there were too many steps to turn a plug on/off.	100%	2	3	6.0
Tried to touch and interact with text for more info.	100%	2	2	4.0
In trying to add a new sensor, goes to Sensor screen, and presses "Configure". Tells him to go somewhere else!	100%	3	1	3.0
Difficulty finding "Add New Item" in the Editor screen.	67%	2	2	2.7
Cannot see the temp when he is adjusting it because it's blocked by his finger.	33%	2	4	2.7
Gave up on programming smart plug, due to confusion.	33%	4	2	2.7
In programming a plug, UI asks "computer: schedule no/yes". He thinks there was a mistake.	67%	3	1	2.0
Unclear how to set temperature for items in the Editor screen.	33%	3	2	2.0
"Select the temperature you want" BUT there are two sliders!	67%	3	1	2.0
Adding a "lunch" item to the schedule was so difficult, he gave up .	33%	4	1	1.3
Editor: he thinks selecting "Home" (in "Add New Item") will disrupt previous settings under "Home"	33%	2	2	1.3
No explicit Exit button from the Program screen.	33%	2	2	1.3
Temp-sensor installation: "Done" is grayed out (no explanation) after he selected something, but more configuration was needed. Tries to press.	67%	2	1	1.3
Failed attempt to find smart-plug installation menu.	33%	3	1	1.0
Turning a plug on, it's not clear that sliding to "Indefinitely" selects it.	33%	1	3	1.0
Wizard doesn't easily support working from home, or non-standard schedules.	33%	3	1	1.0

Table H-10-7 – Qualitative Usability Issues – Category: System Quirks and Pitfalls

Issue	% Occurrence	Severity	Expected Frequency	Score
Plug historical graphs lack a point of reference or comparison	100%	3	3	9.0
energy "Consumption" (Smart-Plug Dashboard) has no units!	100%	3	3	9.0
Time wheel defaults at 12 AM for new items, causing confusion and error.	100%	3	2	6.0
Plugs don't show current power usage (watts) right away (long delay)	100%	2	3	6.0
Trying to explore Program screen, unintentionally activates the Editor (and may bypass the Wizard)	100%	2	3	6.0
Screen timeout causes lost changes.	67%	3	2	4.0
In plug historical graph, it's not very clear that x-axis is time of day	50%	2	3	3.0
Trying to swipe to explore Editor chart, unintentionally activates items to be edited	67%	2	2	2.7
Looking at display from above (being tall), colour differentiation is lost!	33%	2	3	2.0
Disabling WiFi puts "Weather" icon off of home-screen.	1	1	2	2.0
In plug historical graph, text on bottom (time & cost) are too small to read easily.	50%	1	3	1.5
New temp-sensor doesn't show temp immediately, so he thinks install failed	50%	2	1	1.0
Over-pushing "Resume" led to accidental selection, due to poor screen responsiveness.	0.25	2	2	1.0
Asterisk on sensor readings (as a note in the Sensor screen) looks "very bad". "I'm not reading a magazine"	33%	1	2	0.7

Table H-10-8 – Qualitative Usability Issues – Category: Accessory-Specific Issues (not mutually exclusive from above categories)

Issue	% Occurrence	Severity	Expected Frequency	Score
Plug historical graphs lack a point of reference or comparison	100%	3	3	9.0
energy "Consumption" (Smart-Plug Dashboard) has no units!	100%	3	3	9.0
In turning a plug off, did not hit "Resume", but chose "indefinite hold".	100%	2.5	3	7.5
Plugs don't show current power usage (watts) right away (long delay)	100%	2	3	6.0
They thought there were too many steps to turn a plug on/off.	100%	2	3	6.0
Sensors: EI vs Remote Sensor Board?	100%	3	1	3.0
No idea what "Follow Thermostat" means with respect to a plugged-in device.	100%	3	1	3.0
In trying to add a new sensor, goes to Sensor screen, and presses "Configure". Tells him to go somewhere else!	100%	3	1	3.0
In plug historical graph, it's not very clear that x-axis is time of day	50%	2	3	3.0
Doesn't know what bottom-right indicator icons mean (other than wifi).	100%	1	3	3.0
Gave up on programming smart plug, due to confusion.	33%	4	2	2.7
In programming a plug, UI asks "computer: schedule no/yes". He thinks there was a mistake.	67%	3	1	2.0
Gets lost in Settings menu structure/hierarchy.	50%	2	2	2.0
Goes to the Sensor screen to install the sensor	100%	2	1	2.0
Sensors: "Dry Contact" vs "Temperature"?	50%	3	1	1.5
In plug historical graph, text on bottom (time & cost) are too small to read easily.	50%	1	3	1.5
Sensor screen very confusing.	33%	2	2	1.3
Not clear, in Smart Plug Dashboard, that plugs can be switched on/off.	33%	2	2	1.3
Looked for 2nd temp-sensor reading in Details screen.	33%	2	2	1.3
Pressed plug-switch icon (in the Plug Dashboard), looking to set a program.	33%	2	2	1.3
Temp-sensor installation: "Done" is grayed out (no explanation) after he selected something, but more configuration was needed. Tries to press.	67%	2	1	1.3
Failed attempt to find smart-plug installation menu.	33%	3	1	1.0
Turning a plug on, it's not clear that sliding to "Indefinitely" selects it.	33%	1	3	1.0
Diff' between "Monitoring" vs "Control" sensor? Wouldn't they overlap?	50%	2	1	1.0
New temp-sensor doesn't show temp immediately, so he thinks install failed	50%	2	1	1.0

Asterisk on sensor readings (as a note in the Sensor screen) looks <i>"very bad"</i> . <i>"I'm not reading a magazine"</i>	33%	1	2	0.7
--	-----	---	---	------------

Appendix I – Critical Making Character Questionnaire

Critical Making Questionnaire As Your CHARACTER

Character Name: _____

Group: _____

I am _____ years old.

I am responsible for paying my household's energy bills: Yes / No

Gender: _____

[Below, mark an X on the line in the position that best describes you]

My ideal home-energy system offers me:

100% Personal Control |-----|-----| 100% Automation

My decision-making style of energy use is based on:

'Rational Choice' (cost/benefit) |-----|-----| 'Norm Activation' (my values & social-influence)

[Below, circle the choice that best represents you]

Compared to most people, I generally prefer the temperature to be:

warmer	about the same	cooler
--------	----------------	--------

Feeling comfortable in the home is more important to me than saving energy

1. Strongly Agree	2. Agree	3. Neutral	4. Disagree	5. Strongly Disagree
-------------------	----------	------------	-------------	----------------------

It is more important to take care of others than to take care of myself.

1. Strongly Agree	2. Agree	3. Neutral	4. Disagree	5. Strongly Disagree
-------------------	----------	------------	-------------	----------------------

I am motivated to reduce my energy consumption.

1. Strongly Agree	2. Agree	3. Neutral	4. Disagree	5. Strongly Disagree
-------------------	----------	------------	-------------	----------------------

I am willing to sacrifice some convenience to use less electricity.

1. Strongly Agree	2. Agree	3. Neutral	4. Disagree	5. Strongly Disagree
-------------------	----------	------------	-------------	----------------------

When it is cold, I am more likely to put on warm clothes before choosing to turn up the heat.

1. Strongly Agree	2. Agree	3. Neutral	4. Disagree	5. Strongly Disagree
-------------------	----------	------------	-------------	----------------------

I am _____ to changes in indoor temperature:

1. Very Sensitive	2. Somewhat Sensitive	3. Neutral	4. Somewhat Tolerant	5. Very Tolerant
-------------------	-----------------------	------------	----------------------	------------------

I prefer to run my dishwasher when it's most *convenient*, rather than waiting until it's fully loaded.

1. Strongly Agree	2. Agree	3. Neutral	4. Disagree	5. Strongly Disagree
-------------------	----------	------------	-------------	----------------------

Additional Character Details?

Comments?

Appendix J – Informed Consent

Informed Consent Form

Study Title: Effects of Display Type on Thermostat Temperature Selection

Principal investigator: Joshua Stein (416-978-0881; josh.stein@utoronto.ca)
Faculty supervisor: Prof. Greg A. Jamieson (416-946-8504; jamieson@mie.utoronto.ca)
Department of Mechanical and Industrial Engineering, University of Toronto

The purpose of this study is to evaluate the effects of different display types (of a mock thermostat) and instructions on users' selection of an appropriate number on a graphical scale of temperatures.

You are invited to participate in this study because you are a student at the University of Toronto. There will be approximately 150 participants involved.

Procedure

You will be asked to perform the following tasks if you choose to participate in this experiment. The experiment is expected to last 15 to 20 minutes.

1. You will perform two sets of a simple selection task—choosing a number on a graphical scale—each time with a similar but different instruction.
2. You will complete a questionnaire for personal statistics (age, gender, and year and discipline of study), and a few questions related to the task domain (thermostats and thermal comfort).

Risks

There are no known or anticipated risks for participants in this study. There are no physiologically- or psychologically-demanding tasks in this study. Your performance on the tasks will not be used for any other purpose than to study human interaction with the user interface.

Benefits

You may benefit from gaining exposure to information about thermostats and thermal comfort.

Compensation

You will receive \$5 for completing this study of approximately 15-20 minutes. The total compensation amount will be paid in one full payment at the end of the study. If you wish to withdraw from the study before completion, you will still be compensated for the full amount.

Confidentiality

Your privacy and identity will be carefully protected in this study. Your data will be kept securely, and will only be accessed by the investigator and future researchers within the investigator's lab. In any publication, information will be provided in such a way that you cannot be identified. Your participation and identity will not be disclosed to others. You may request that your data be withdrawn and destroyed after your participation and debriefing, provided that it has not yet been analyzed and reported.

Participation

Your participation in this study is completely voluntary. You may decline to participate or decline to answer any question without any negative consequences. In addition, you may withdraw from the study at any time without any penalty, and request your data be destroyed.

Contacting the investigators

M.A.Sc. candidate Joshua Stein is undertaking this study in partial fulfillment of his degree requirements. If you have any additional questions later about this study or would like a summary of the research results, Joshua (josh.stein@utoronto.ca; 416-978-0881) will assist you. For information about participants' rights in scientific study, you can contact the University of Toronto's Ethics Review Office (ethics.review@utoronto.ca; 416-946-3273).

You will be given a completed copy of this form to keep.

Participation Consent

I have read this Informed Consent Form. I have had the opportunity to ask any questions that I had regarding the study, and I have received satisfactory answers to those questions. By my signature I affirm that I agree to take part in this study with the understanding that I can withdraw at any point. I have received a copy of this Informed Consent Form.

Participant Signature

Investigator Signature

(Please PRINT name)

(Please PRINT name)

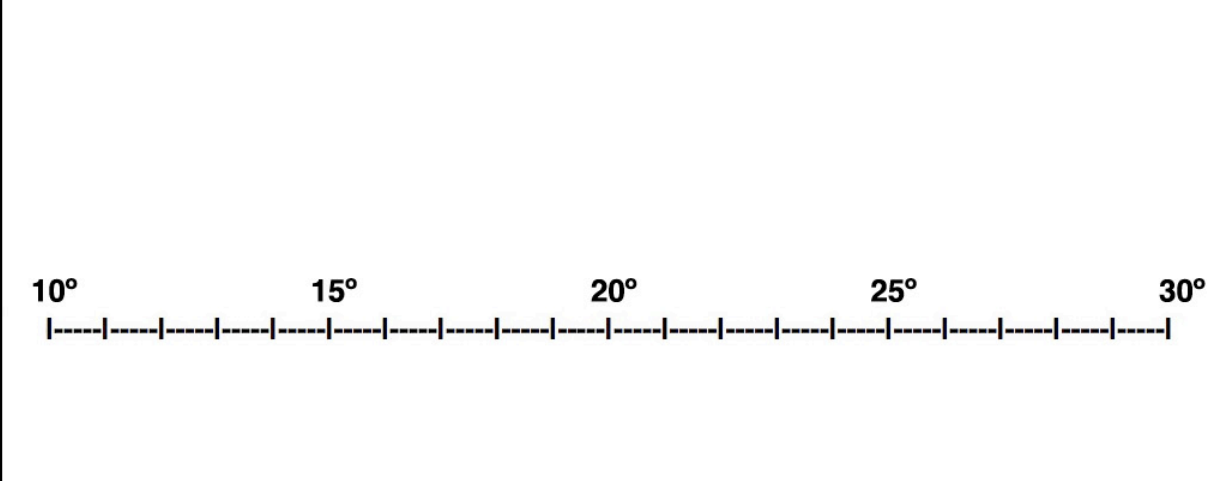
Date

Date

Appendix K – Thermostat Paper-Prototypes

Display Condition 1: Scale-only

Thermostat Setup Mode: Step 1 - Select a temperature setting for times when you are at home



10° 15° 20° 25° 30°

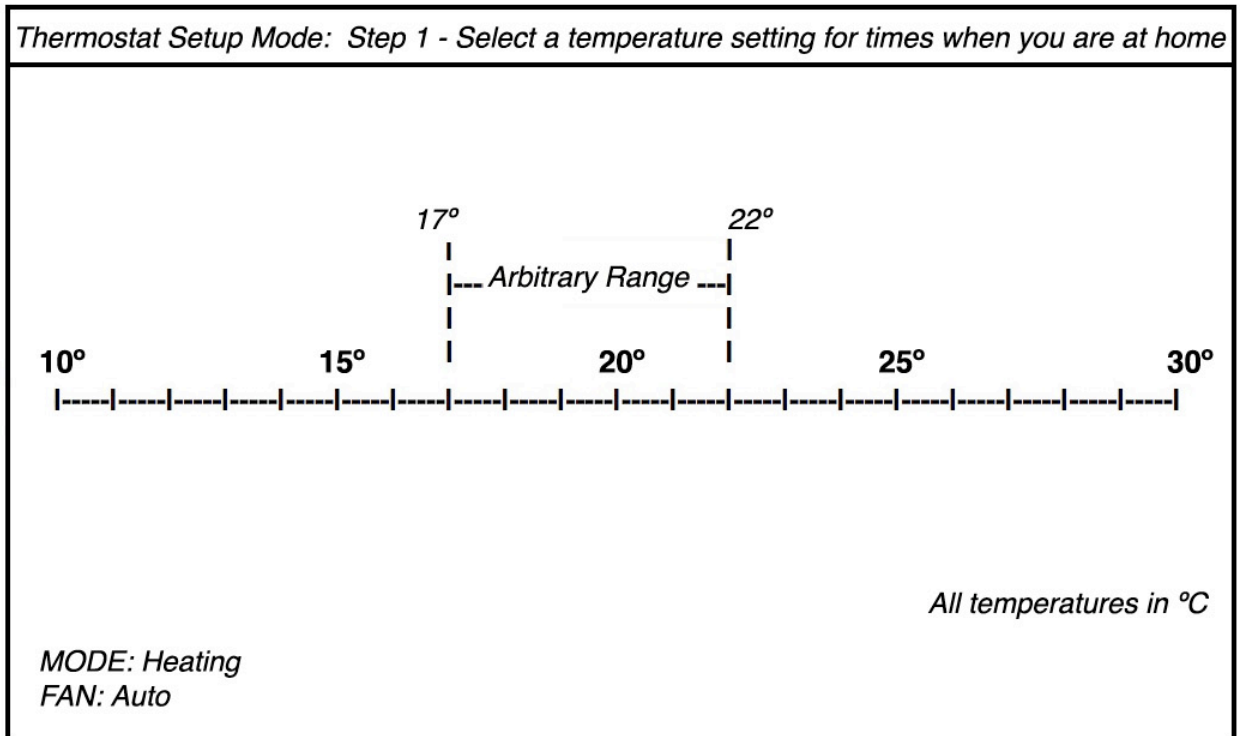
All temperatures in °C

MODE: Heating
FAN: Auto

How confident are you that your choice is optimal for you?
(circle your answer):

Very unconfident	Somewhat unconfident	Neither confident nor unconfident.	Somewhat Confident	Very confident
---------------------	-------------------------	--	-----------------------	-------------------

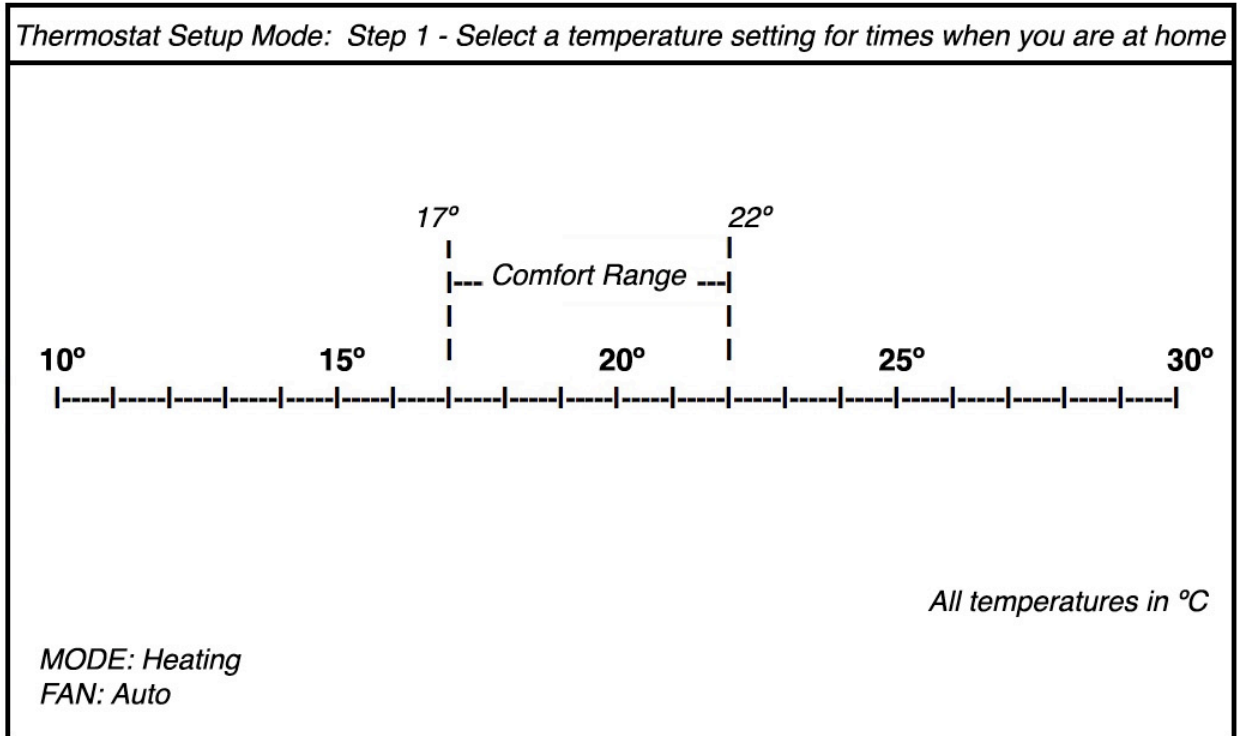
Display Condition 2: Anchor



How confident are you that your choice is optimal for you?
(circle your answer):

Very unconfident	Somewhat unconfident	Neither confident nor unconfident.	Somewhat Confident	Very confident
---------------------	-------------------------	--	-----------------------	-------------------

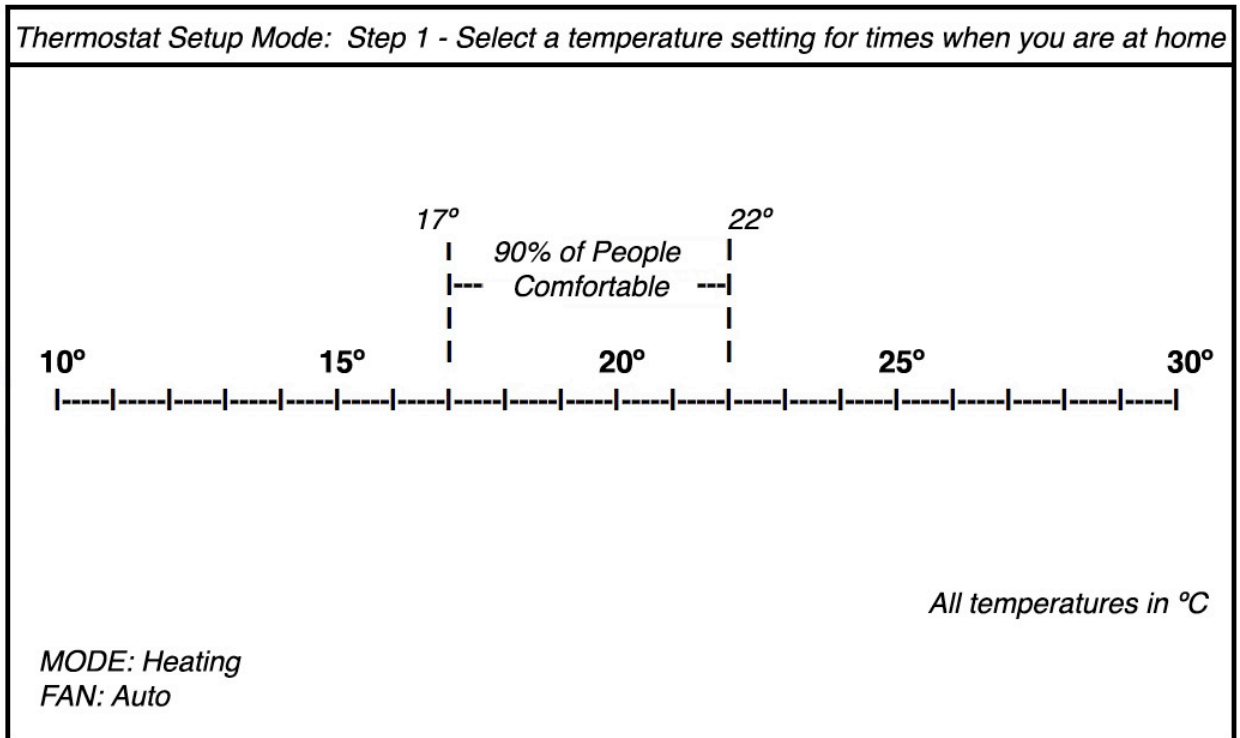
Display Condition 3: Meaningful-non-social



How confident are you that your choice is optimal for you?
(circle your answer):

Very unconfident	Somewhat unconfident	Neither confident nor unconfident.	Somewhat Confident	Very confident
---------------------	-------------------------	--	-----------------------	-------------------

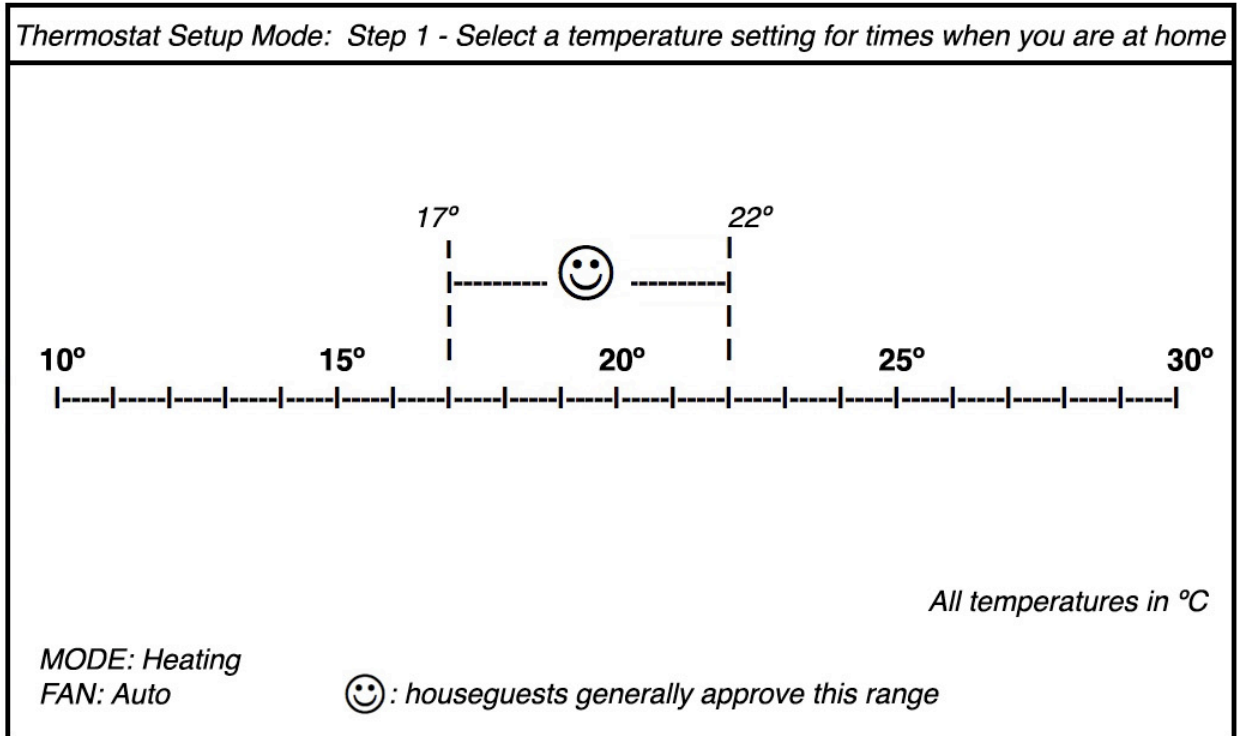
Display Condition 4: Descriptive Norm



How confident are you that your choice is optimal for you?
(circle your answer):

Very unconfident	Somewhat unconfident	Neither confident nor unconfident.	Somewhat Confident	Very confident
---------------------	-------------------------	--	-----------------------	-------------------

Display Condition 5: Injunctive Norm



How confident are you that your choice is optimal for you?
(circle your answer):

Very unconfident	Somewhat unconfident	Neither confident nor unconfident.	Somewhat Confident	Very confident
---------------------	-------------------------	--	-----------------------	-------------------

Appendix L – Participant Questionnaire

Participant Questionnaire

Age: _____

Gender: _____

Degree and program: _____ (e.g., BSc in mechanical engineering)

Year of study: _____ (e.g., third year student)

Do you have central heating at home? _____

If so, do you control it? _____

What temperature do you think it is usually set to when you are around and awake? _____

Are you more familiar with °C or °F? _____

For the following statements, *circle* the answer that you most agree with:

How comfortable are you right now?:

Much Too warm	Too warm	Comfortably warm	Comfortable	Comfortably cool	Too cool	Much Too Cool
---------------------	-------------	---------------------	-------------	---------------------	-------------	---------------------

To be comfortable, indoor temperatures should be _____ in winter than/as in the summer.

warmer	cooler	the same
--------	--------	----------

Heating is more energy-efficient at _____ temperature settings.

lower	higher
-------	--------

A thermostat works like a valve: the more extreme the setting, the faster the temperature will change.

True	False	Unsure
------	-------	--------

Any comments about anything in this study?

Appendix M – Debrief Form

Study Title: Effects of Display Type on Thermostat Temperature Selection

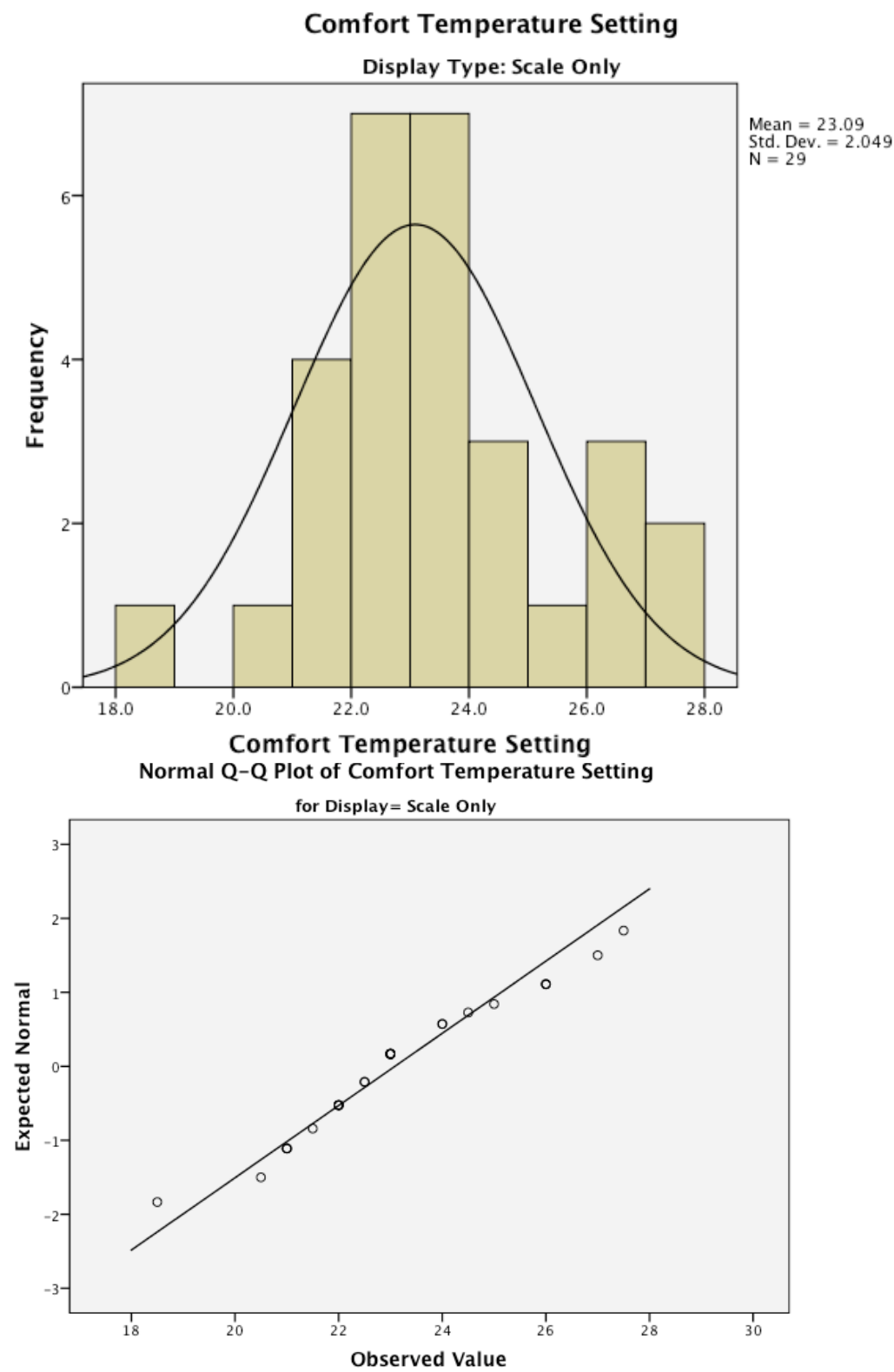
Some misrepresentation was used in this study.

The range you were presented with (17-22 °C) actually represents a range in which 90% of participants (of naturalistic field studies) in naturally-ventilated office buildings in the winter, already exposed to their specific thermal conditions, found acceptable. This range applies specifically to conditions in which the Mean Monthly Outdoor Temperature (average of daily highs and lows for a particular month), is approximately 5° C or less (typical for winter in Toronto). The way in which the range was described for your particular condition was used to see if the presentation format would affect your choice of a comfortable indoor temperature.

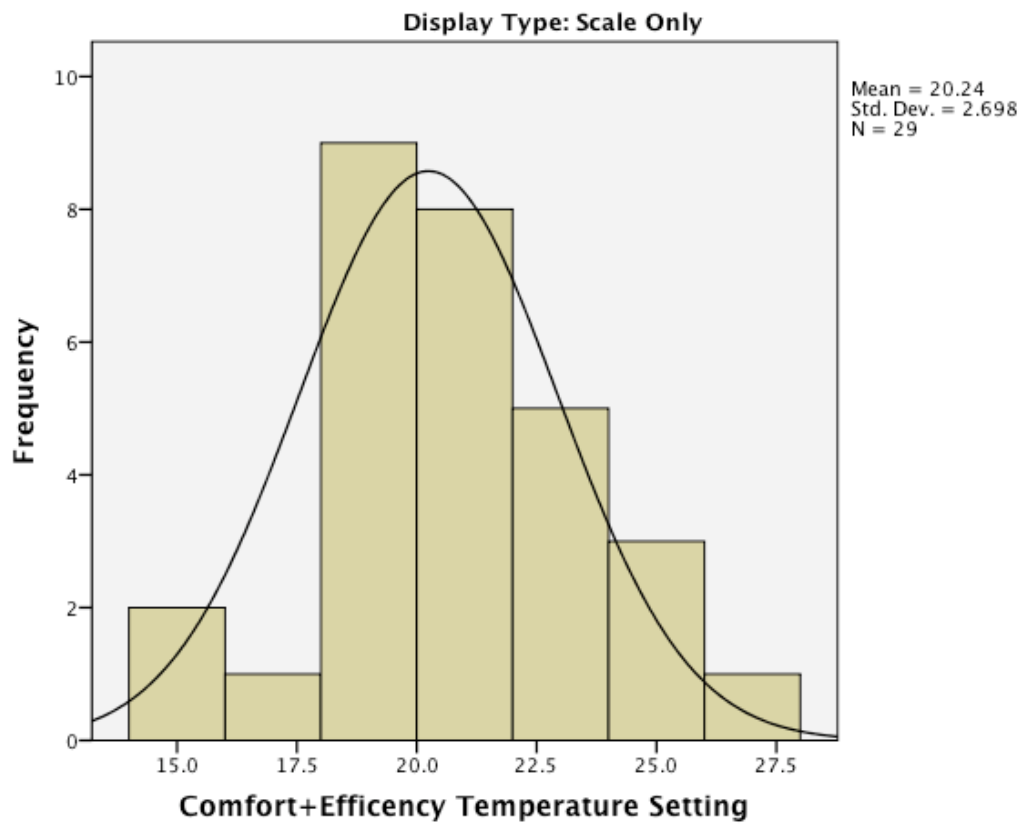
Please let the experimenter know if you have any questions or wish to withdraw your data from the study. You may do so without consequence, and still collect your \$5 for your time. You may also request that your data be withdrawn and destroyed anytime after your participation and debriefing, provided that it has not yet been analyzed and reported; see your Informed Consent Form for the investigator's contact information.

Thank you for participating!

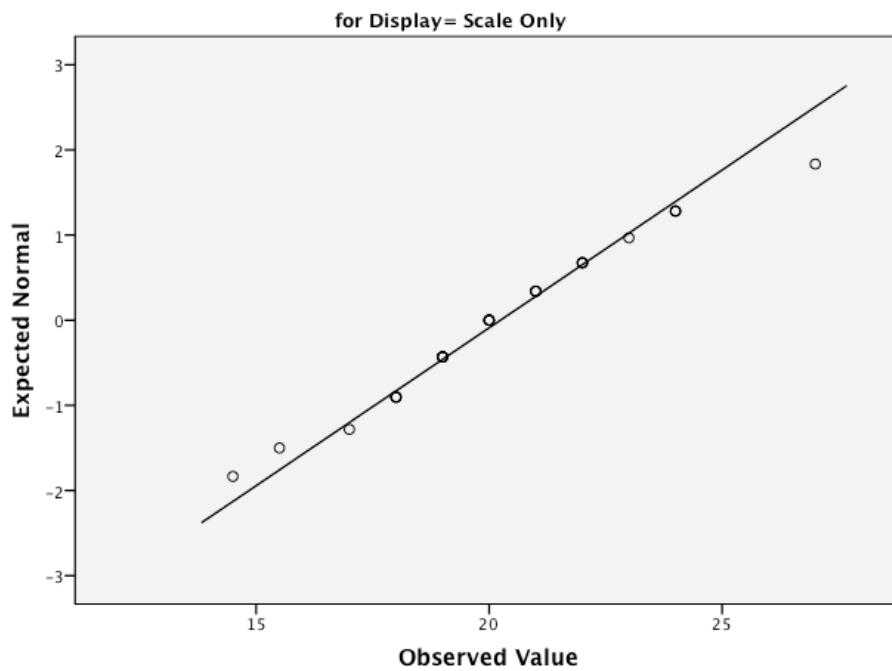
Appendix N – Normality Plots

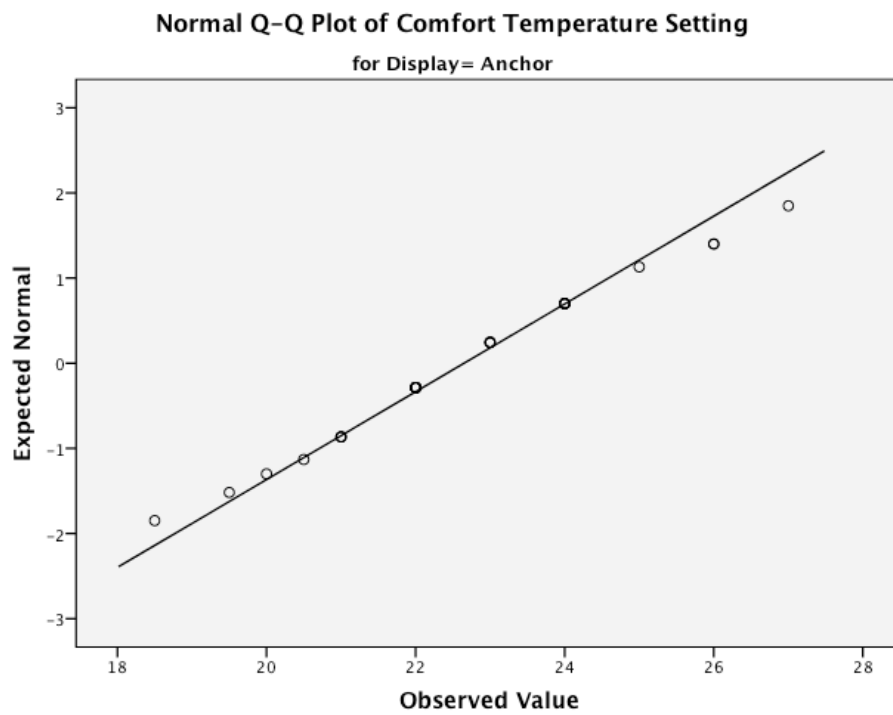
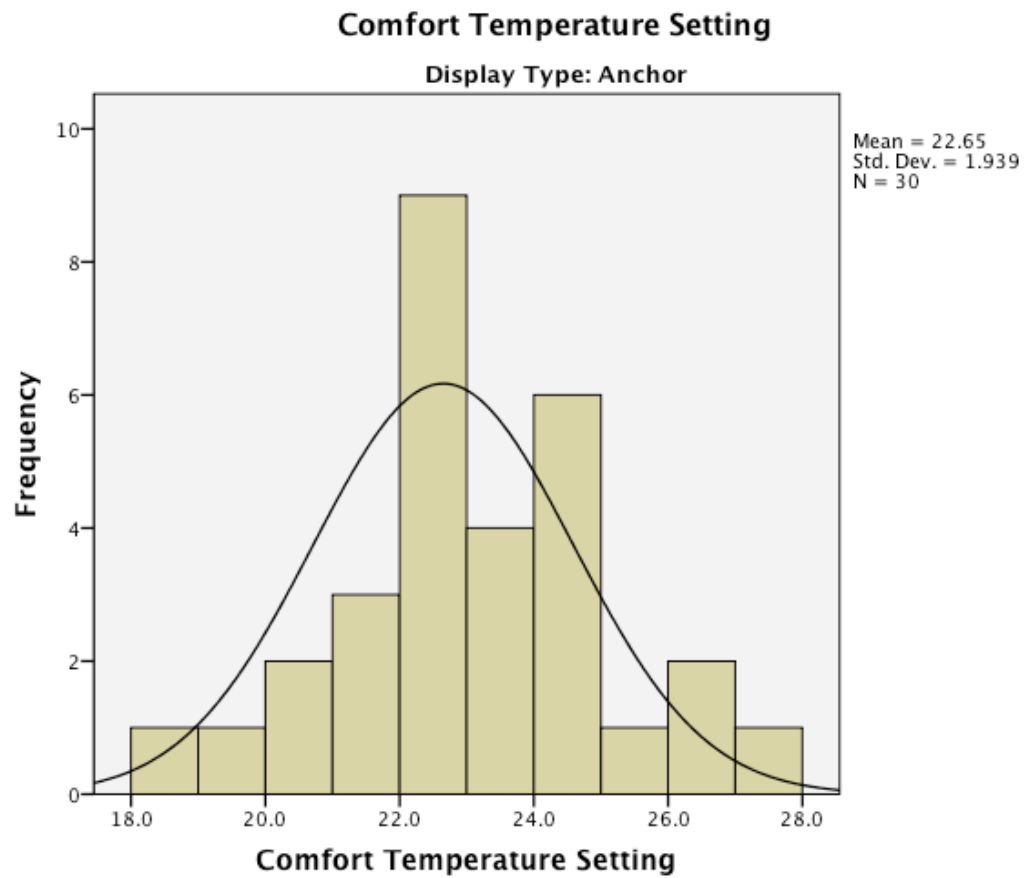


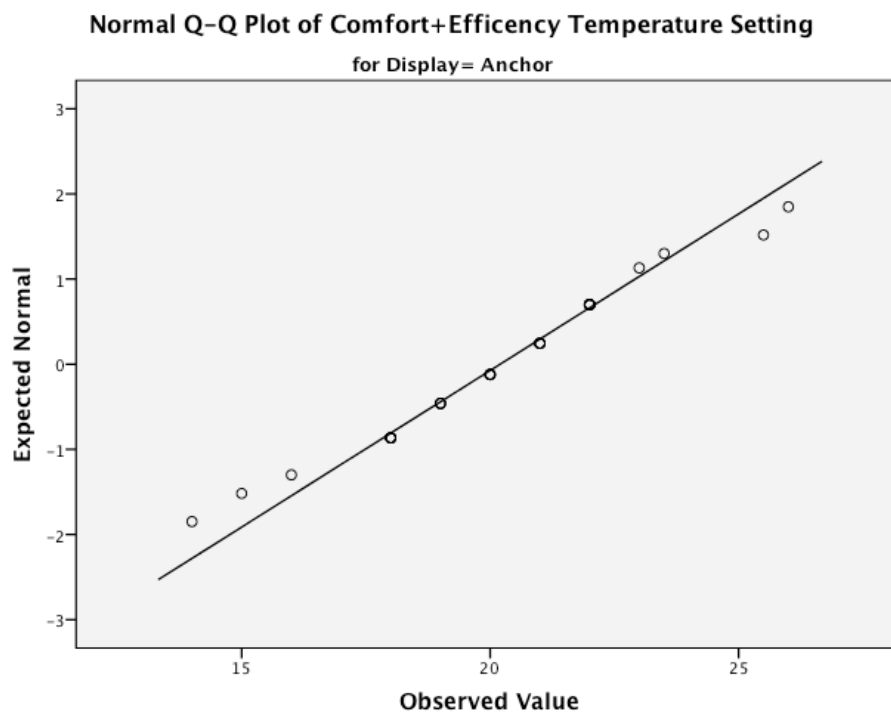
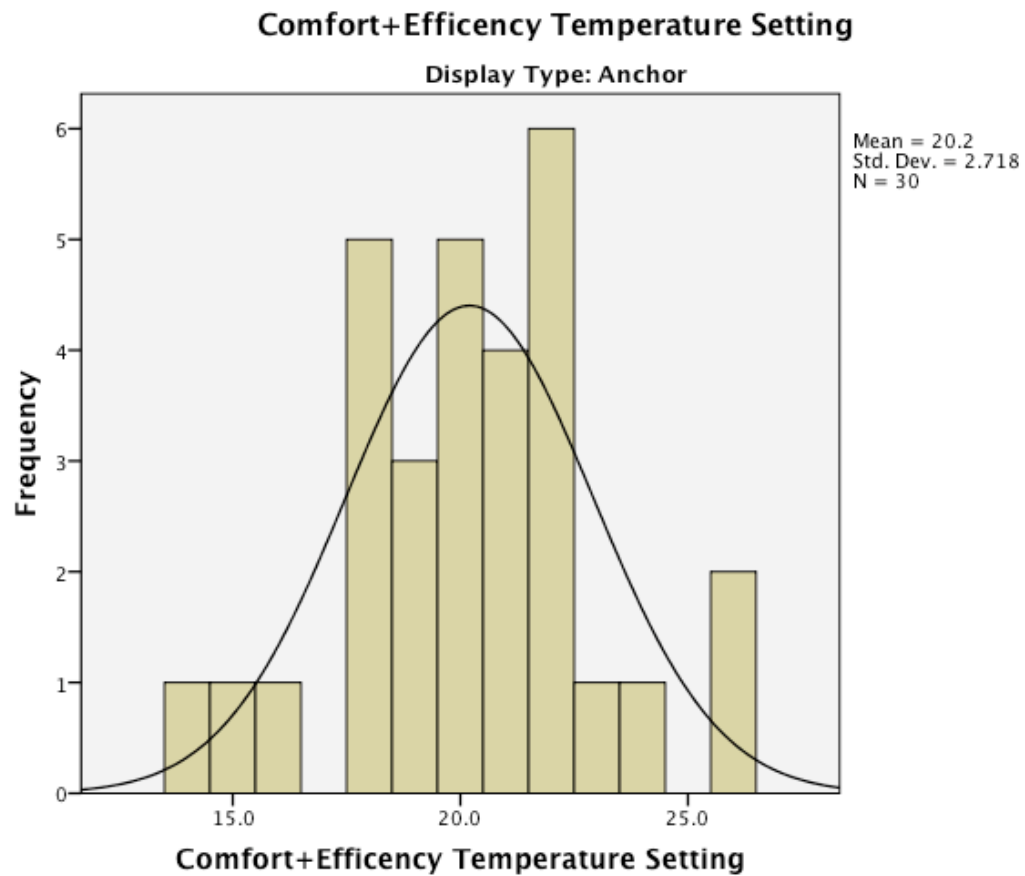
Comfort+Efficiency Temperature Setting

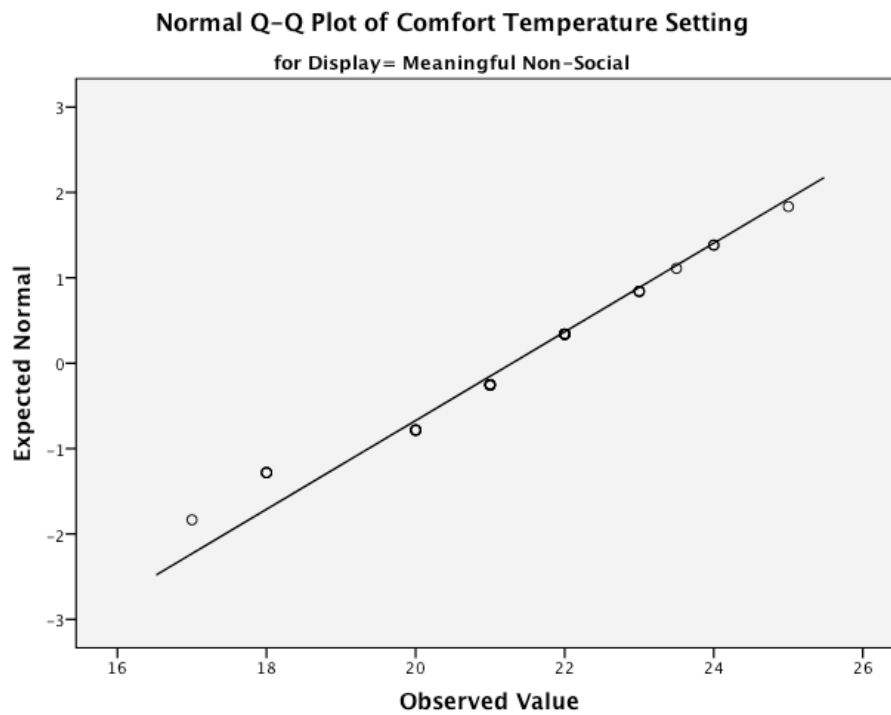
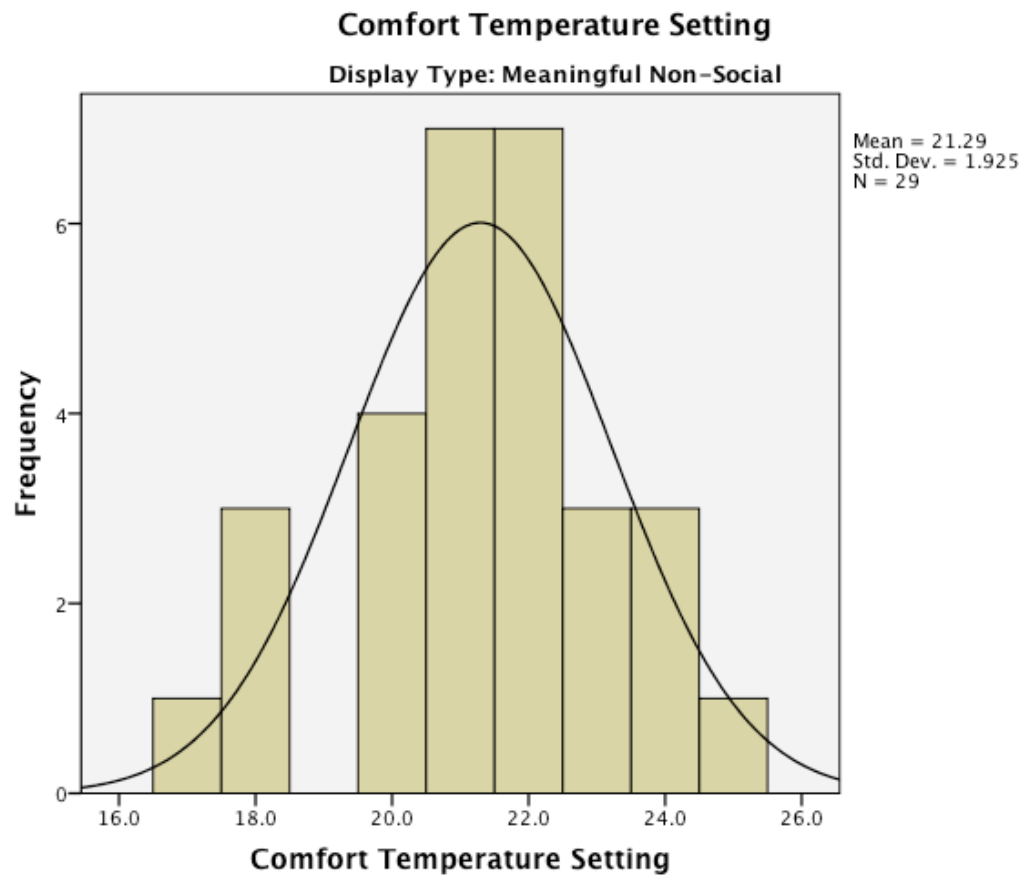


Normal Q-Q Plot of Comfort+Efficiency Temperature Setting

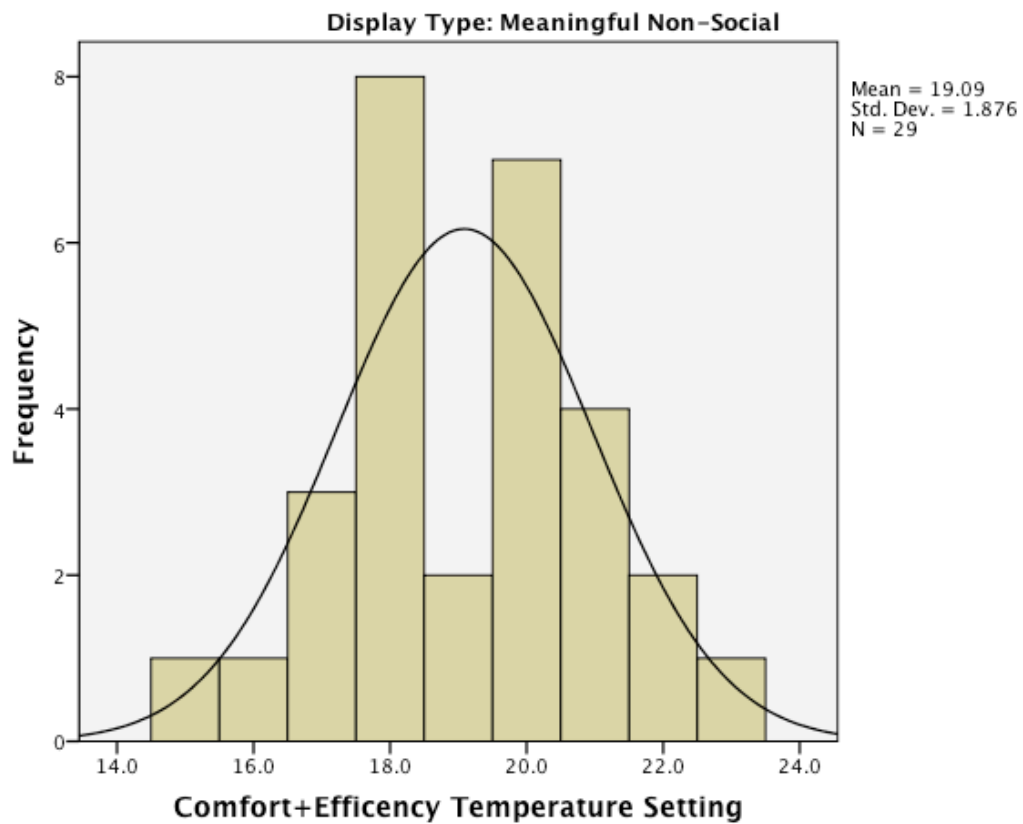




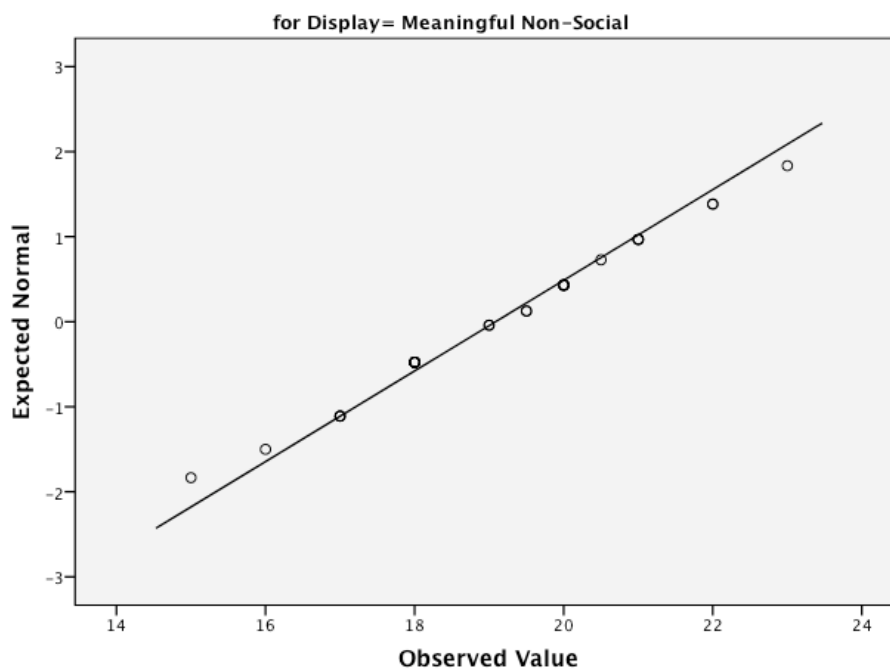


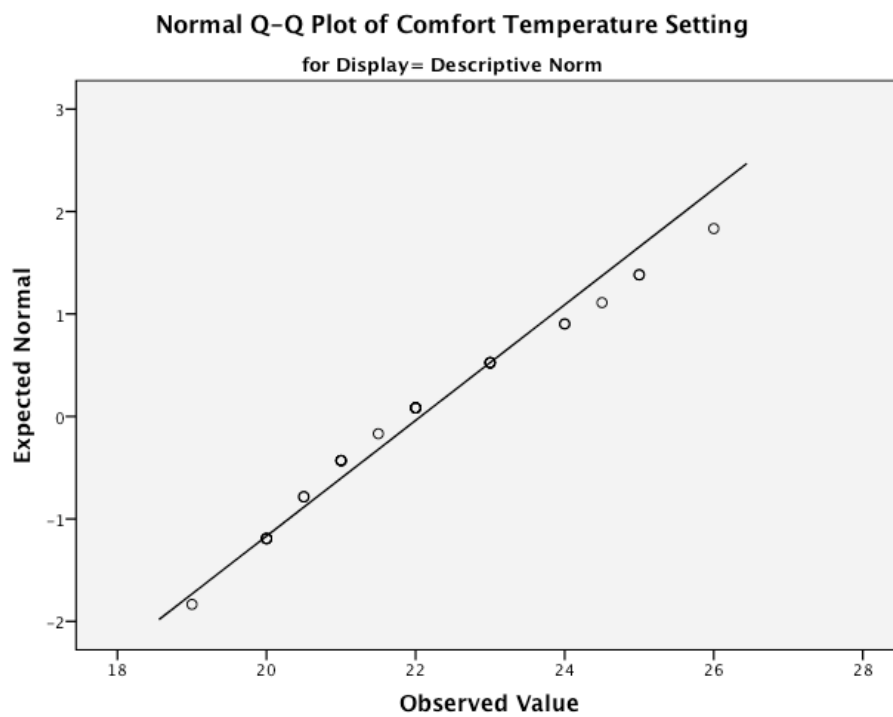
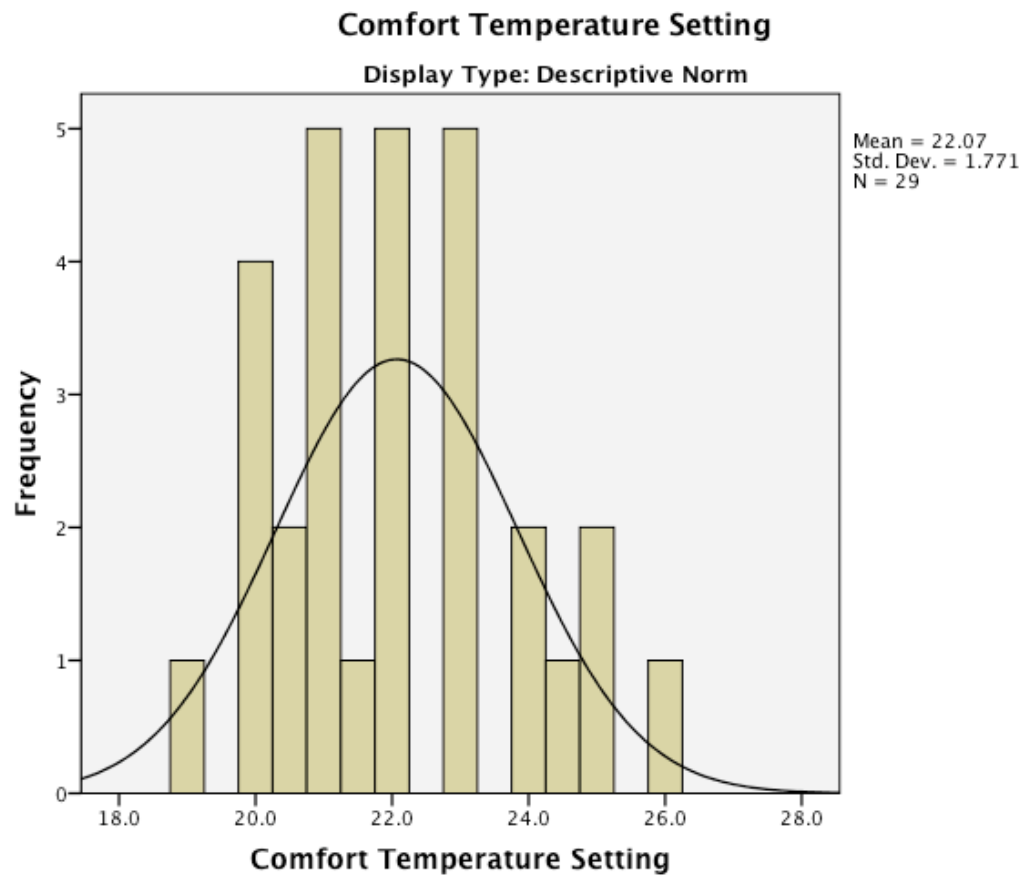


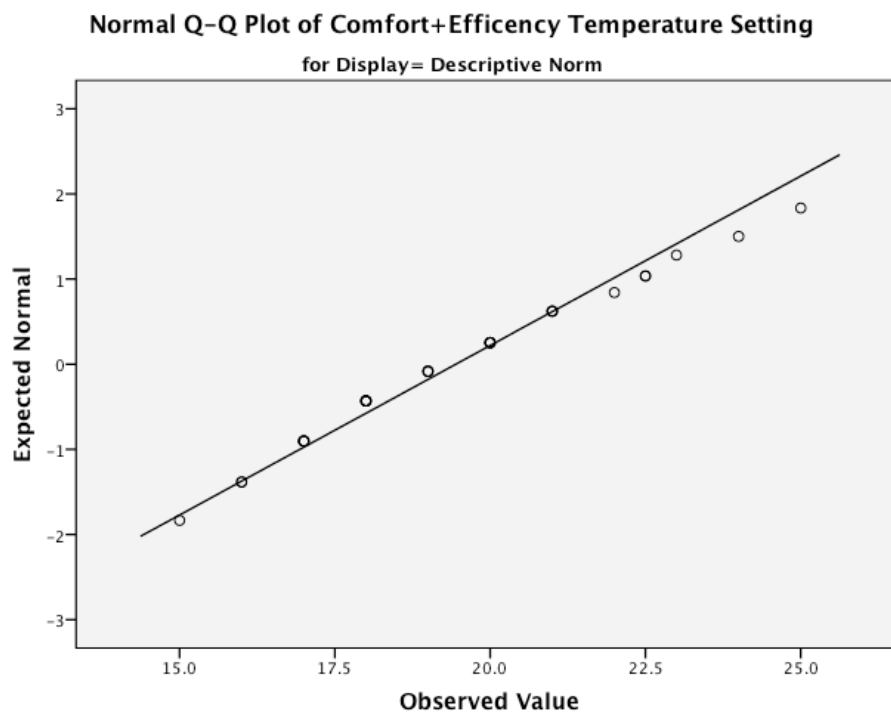
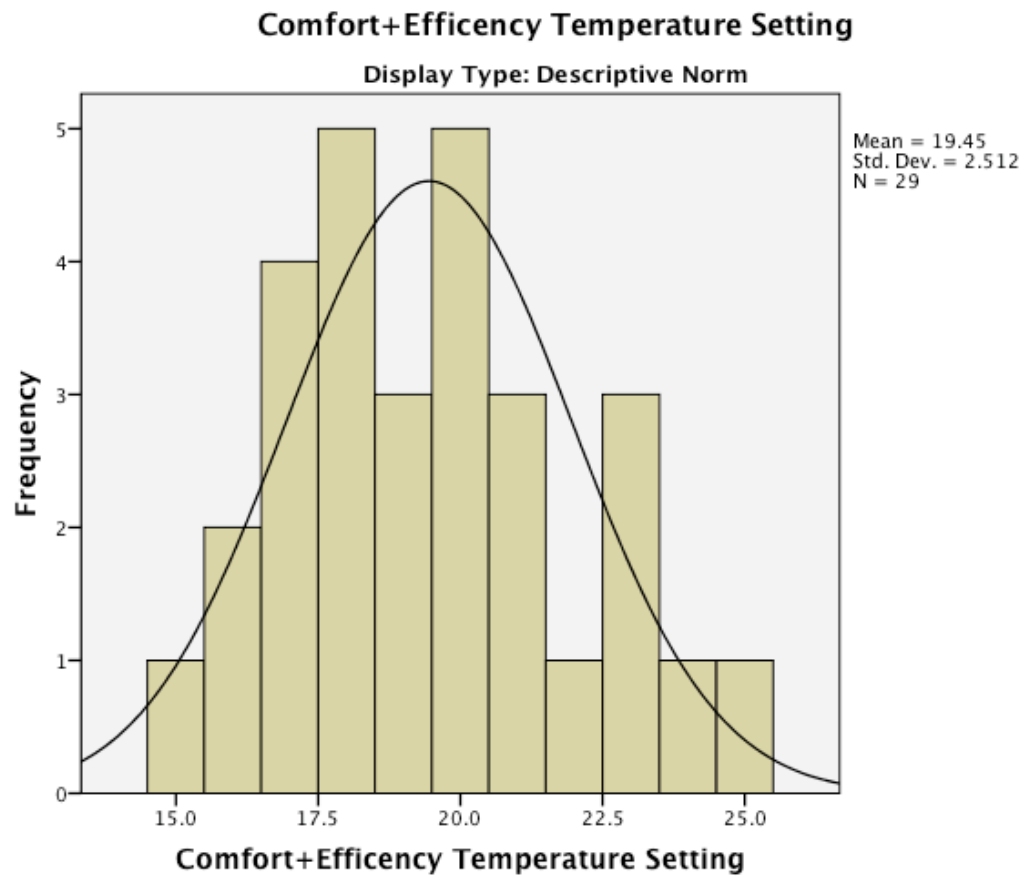
Comfort+Efficiency Temperature Setting

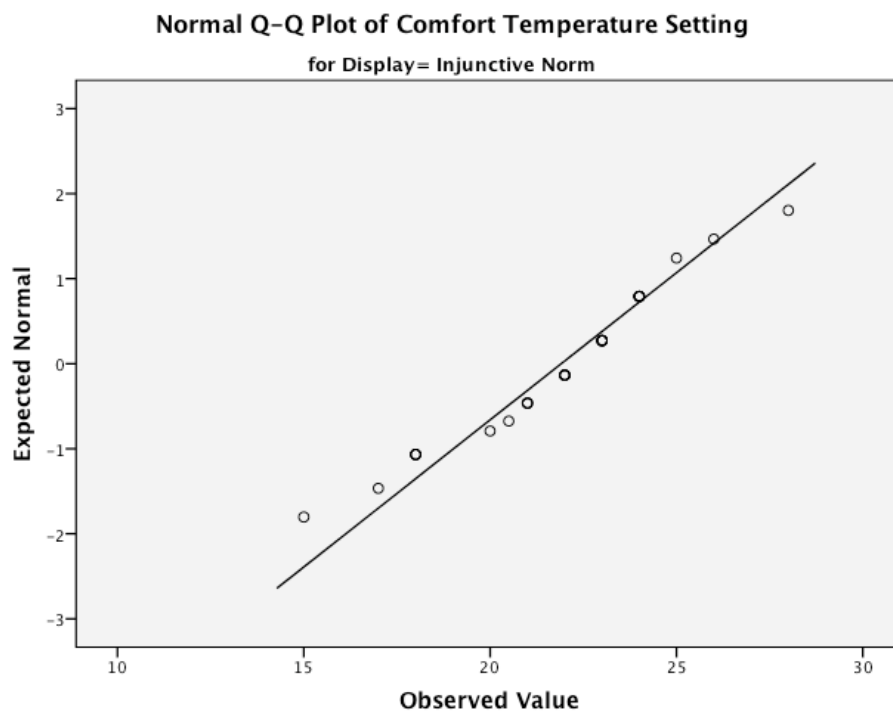
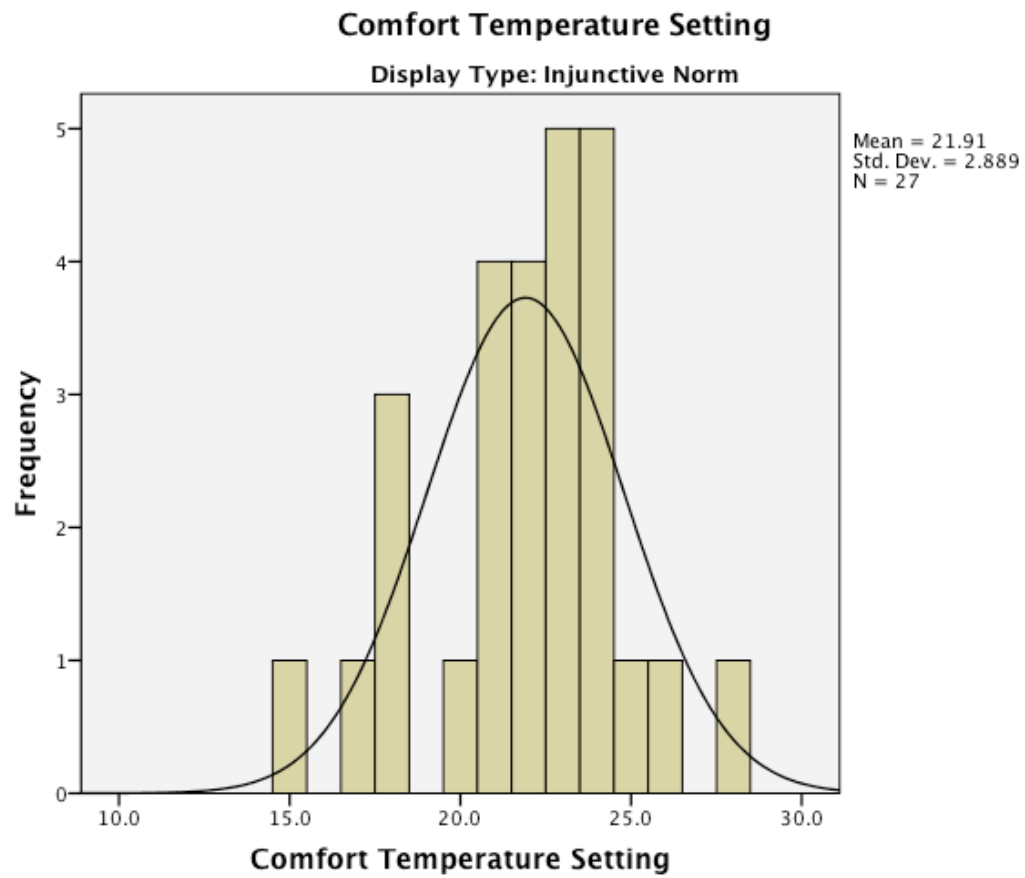


Normal Q-Q Plot of Comfort+Efficiency Temperature Setting

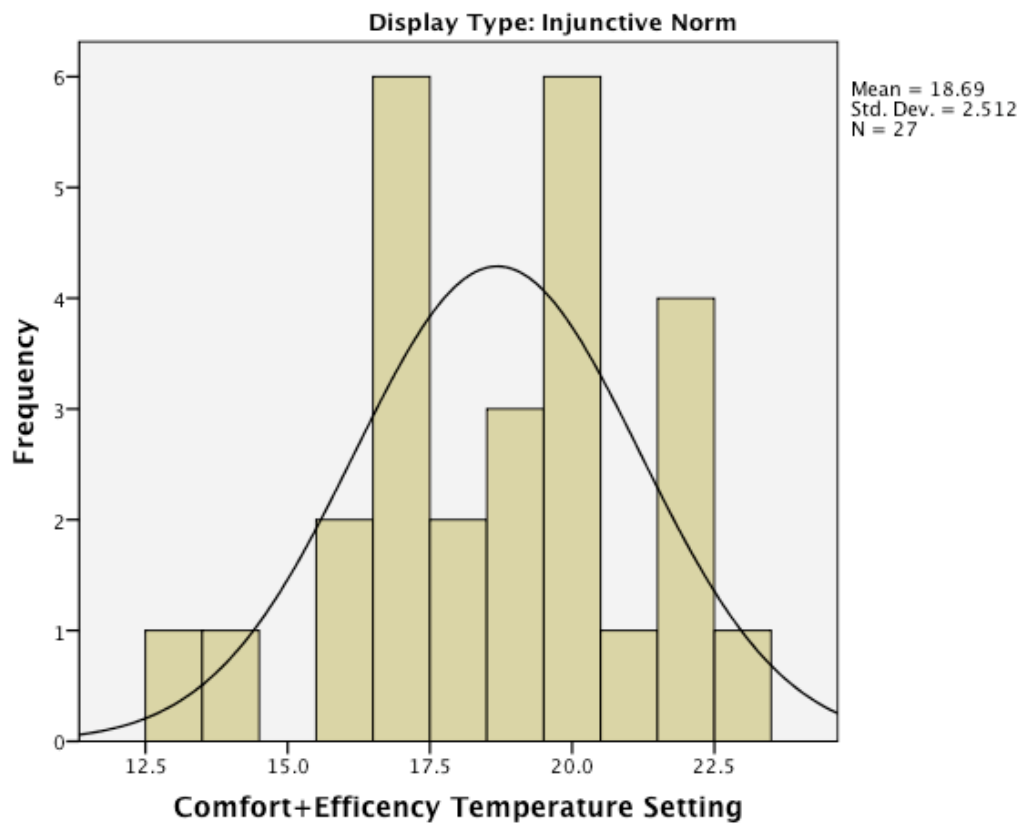




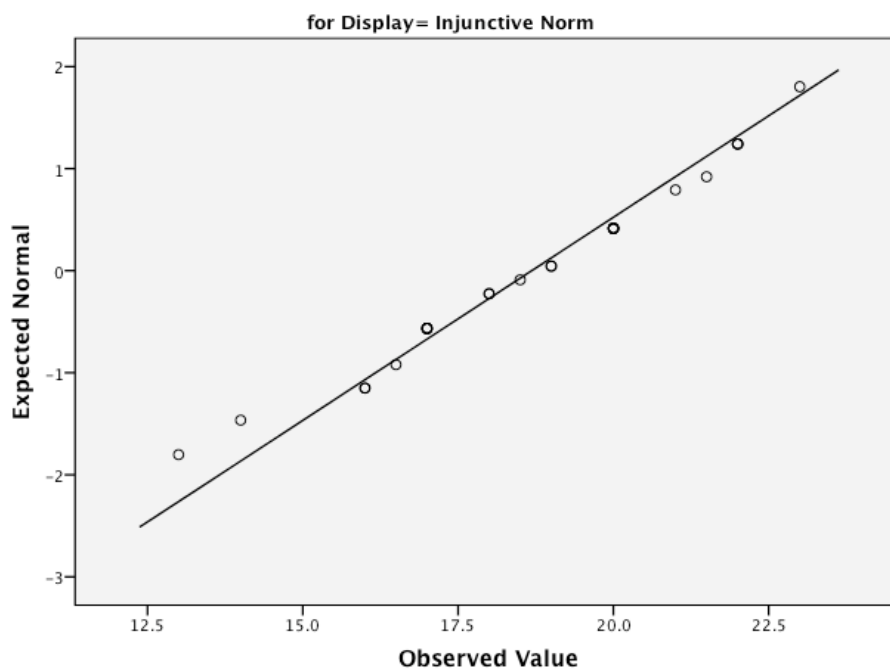




Comfort+Efficiency Temperature Setting



Normal Q-Q Plot of Comfort+Efficiency Temperature Setting



Appendix O – Study Calculations and Statistical Tests

Normality

Tables O-1 and O-2 show values of skewness and kurtosis for each sample group. Z values for both parameters are well under the critical value 1.96 (at $p < .05$), therefore the groups are likely normally distributed.

Table O-10-9– Tests of Skewness

Instructions	Display	Skewness	Standard Error	Z _{skewness}
C	Scale Only	0.356	0.434	0.820
C	Anchor	0.193	0.427	0.452
C	Meaningful Non-Social	-0.411	0.434	-0.947
C	Descriptive Norm	0.452	0.434	1.041
C	Injunctive Norm	-0.451	0.448	-1.007
C+E	Scale Only	0.249	0.434	0.574
C+E	Anchor	-0.133	0.427	-0.311
C+E	Meaningful Non-Social	0.012	0.434	0.028
C+E	Descriptive Norm	0.367	0.434	0.846
C+E	Injunctive Norm	-0.321	0.448	-0.717

Table O-10-10– Tests of Kurtosis

Instructions	Display	Kurtosis	Standard Error	Z _{kurtosis}
C	Scale Only	0.183	0.845	0.217
C	Anchor	0.111	0.833	0.133
C	Meaningful Non-Social	0.037	0.845	0.044
C	Descriptive Norm	-0.502	0.845	-0.594
C	Injunctive Norm	0.396	0.872	0.454
C+E	Scale Only	0.49	0.845	0.580
C+E	Anchor	0.398	0.833	0.478
C+E	Meaningful Non-Social	-0.262	0.845	-0.310
C+E	Descriptive Norm	-0.446	0.845	-0.528
C+E	Injunctive Norm	-0.323	0.872	-0.370

Table O-3 shows the results for the Kolmogorov-Smirnov test produced by SPSS.

Table O-10-11– SPSS Output for Tests of Normality

		Kolmogorov–Smirnov ^a			Shapiro–Wilk		
	Display Type	Statistic	df	Sig.	Statistic	df	Sig.
Comfort Temperature Setting	Scale Only	.206	29	.003	.954	29	.229
	Anchor	.165	30	.037	.968	30	.477
	Meaningful Non–Social	.164	29	.046	.953	29	.219
	Descriptive Norm	.141	29	.148	.957	29	.273
	Injunctive Norm	.142	27	.170	.960	27	.372
Comfort+Efficiency Temperature Setting	Scale Only	.122	29	.200 [*]	.976	29	.735
	Anchor	.121	30	.200 [*]	.970	30	.551
	Meaningful Non–Social	.167	29	.038	.973	29	.644
	Descriptive Norm	.132	29	.200 [*]	.971	29	.587
	Injunctive Norm	.144	27	.158	.964	27	.452

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Table O-4 shows the results of Levene’s test of homogeneity of variance.

Table O-10-12 – SPSS Output for Levene’s Test

		Test of Homogeneity of Variance			
		Levene Statistic	df1	df2	Sig.
Comfort Temperature Setting	Based on Mean	1.548	4	139	.192
	Based on Median	1.388	4	139	.241
	Based on Median and with adjusted df	1.388	4	119.625	.242
	Based on trimmed mean	1.501	4	139	.205
Comfort+Efficiency Temperature Setting	Based on Mean	.766	4	139	.549
	Based on Median	.668	4	139	.615
	Based on Median and with adjusted df	.668	4	126.232	.615
	Based on trimmed mean	.757	4	139	.555

ANOVA Results

Table O-10-13 – SPSS Output for the mixed-ANOVA within-subjects effects, for instruction type

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Instructions	Sphericity Assumed	512.256	1	512.256	279.515	.000
	Greenhouse-Geisser	512.256	1.000	512.256	279.515	.000
	Huynh-Feldt	512.256	1.000	512.256	279.515	.000
	Lower-bound	512.256	1.000	512.256	279.515	.000
Instructions * Display	Sphericity Assumed	8.423	4	2.106	1.149	.336
	Greenhouse-Geisser	8.423	4.000	2.106	1.149	.336
	Huynh-Feldt	8.423	4.000	2.106	1.149	.336
	Lower-bound	8.423	4.000	2.106	1.149	.336
Error(Instructions)	Sphericity Assumed	254.740	139	1.833		
	Greenhouse-Geisser	254.740	139.000	1.833		
	Huynh-Feldt	254.740	139.000	1.833		
	Lower-bound	254.740	139.000	1.833		

Table O-10-14 – SPSS Output for the mixed-ANOVA between-subjects effects, for display type

Tests of Between-Subjects Effects

Measure: MEASURE_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	125250.672	1	125250.672	14064.012	.000
Display	100.346	4	25.087	2.817	.028
Error	1237.900	139	8.906		

Table 0-10-15— SPSS Output for between-subjects planned contrasts, for display type

Contrast Results (K Matrix)

		Averaged Variable
Display Type Simple Contrast ^a		MEASURE_1
Level 2 vs. Level 1	Contrast Estimate	-.239
	Hypothesized Value	0
	Difference (Estimate - Hypothesized)	-.239
	Std. Error	.550
	Sig.	.665
	95% Confidence Interval for Difference	Lower Bound Upper Bound -1.325 .848
Level 3 vs. Level 1	Contrast Estimate	-1.474
	Hypothesized Value	0
	Difference (Estimate - Hypothesized)	-1.474
	Std. Error	.554
	Sig.	.009
	95% Confidence Interval for Difference	Lower Bound Upper Bound -2.570 -.378
Level 4 vs. Level 1	Contrast Estimate	-.905
	Hypothesized Value	0
	Difference (Estimate - Hypothesized)	-.905
	Std. Error	.554
	Sig.	.105
	95% Confidence Interval for Difference	Lower Bound Upper Bound -2.001 .191
Level 5 vs. Level 1	Contrast Estimate	-1.367
	Hypothesized Value	0
	Difference (Estimate - Hypothesized)	-1.367
	Std. Error	.564
	Sig.	.017
	95% Confidence Interval for Difference	Lower Bound Upper Bound -2.483 -.252

a. Reference category = 1

Table O-10-16 — SPSS Output for the Kruskal-Wallis Test of confidence scores by display type

Ranks			
	Display Type	N	Mean Rank
Comfort Confidence	Scale Only	29	81.33
	Anchor	30	66.12
	Meaningful Non-Social	28	70.48
	Descriptive Norm	29	75.72
	Injunctive Norm	26	63.13
	Total	142	
Comfort+Efficiency Confidence	Scale Only	29	76.90
	Anchor	30	73.33
	Meaningful Non-Social	28	63.04
	Descriptive Norm	29	76.59
	Injunctive Norm	26	66.81
	Total	142	

Table O-10-17 – SPSS Output of the test statistics for the K-W test

Test Statistics ^{a,b}						
					Comfort Confidence	Comfort+Efficiency Confidence
Chi-Square					4.334	3.360
df					4	4
Asymp. Sig.					.363	.499
Monte Carlo Sig. Sig.					.360 ^c	.502 ^c
99% Confidence Interval Lower Bound					.347	.489
Upper Bound					.372	.515

a. Kruskal Wallis Test

b. Grouping Variable: Display Type

c. Based on 10000 sampled tables with starting seed 2000000.

Table O-10-18 – SPSS Output from the Wilcoxon Signed-Rank test of confidence scores by instruction type

Ranks		N	Mean Rank	Sum of Ranks
Comfort+Efficiency Confidence – Comfort Confidence	Negative Ranks	50 ^a	32.58	1629.00
	Positive Ranks	15 ^b	34.40	516.00
	Ties	77 ^c		
	Total	142		

a. Comfort+Efficiency Confidence < Comfort Confidence

b. Comfort+Efficiency Confidence > Comfort Confidence

c. Comfort+Efficiency Confidence = Comfort Confidence

Table O-10-19 – SPSS Output of the test statistics for the WS-R test

Test Statistics ^{a,c}				Comfort+Efficiency Confidence – Comfort Confidence
Z				-4.023 ^b
Asymp. Sig. (2-tailed)				.000
Monte Carlo Sig. (2-tailed)	Sig.	99% Confidence Interval		.000
		Lower Bound		.000
		Upper Bound		.000
Monte Carlo Sig. (1-tailed)	Sig.	99% Confidence Interval		.000
		Lower Bound		.000
		Upper Bound		.000

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

c. Based on 10000 sampled tables with starting seed 624387341.