8 Mental Model Interface Design

Putting Users in Control of Their Home-Heating Systems

8.1 INTRODUCTION

The final chapter in this book considers all the hypotheses described in Chapter 1 in an empirical study and draws together threads developed in the preceding chapters. It explores how a mental-model-driven design can influence the model held (hypothesis 3), which in turn influences behaviour patterns (hypothesis 1) to improve the achievement of home-heating goals (hypothesis 2). Linking these hypotheses together with the same data source provides evidence to test overall hypothesis 4 that 'the knowledge of existing mental models of devices can be used in device design to encourage patterns of device use that increase goal achievement'. This chapter performs statistical tests on data collected through application of QuACk developed in Chapter 3, as well as automated data collected from a home-heating simulation. The simulation design is documented in Chapter 7 and is informed by the design insights identified in Chapter 6. These, in turn, were based on evaluation of the omissions and errors in householders' mental models of home heating described in Chapters 4 and 5.

Energy consumption due to home heating is a key contributor to climate change, making up 58% of UK domestic energy consumption (Department of Energy and Climate Change 2011). It is easy to save energy in the home: just don't turn the heating on (Sauer et al. 2009). The real challenge is *using* energy effectively and efficiently, not just saving it. Using energy effectively to meet realistic heating goals is no mean feat (Revell and Stanton 2014, Chapter 4). Doing so in a way that minimizes waste is rife with difficulties when using devices that were not designed with this emphasis (Sauer et al. 2009, Revell and Stanton 2016).

Occupant behaviour is a key variable affecting the amount of energy used in homes (Raaij and Verhallen 1983, Emery and Kippenham 2006, Lutzenhiser and Bender 2008, Guerra-Santin and Itard 2010, Dalla Rosa and Christensen 2011). Kempton (1986) discovered that variations in the way occupants behaved with their home-heating thermostat could be explained by differences in their 'mental model' of its function. Different types of mental models held by occupants encouraged

different behaviour strategies for saving energy overnight. Kempton (1986) estimated considerable energy savings could result if specific mental models of thermostat function were promoted to domestic users. Householders' misunderstandings about thermostat function and the workings of the heating system as a whole is still a problem today, however (Brown and Cole 2009, Shipworth et al. 2010, Revell and Stanton 2014, 2015). Revell and Stanton (2014, 2015, 2016) extended the findings of Kempton (1986) to consider functional mental models of all home-heating controls present in the home, as well as their interactions at a system level. In addition to inappropriate mental models of device function, in Chapter 4 it was found that incomplete mental models at a system level explained differences in behaviour strategies that either wasted energy or jeopardized comfort goals.

Mental models are described as internal representations of the physical world (Veldhuyzen and Stassen 1976, Johnson-Laird 1983, Rasmussen 1983) and can act as an internal mechanism to allow users to understand, explain, predict and operate the states of systems (Craik 1943, Gentner and Stevens 1983, Kieras and Bovair 1984, Rouse and Morris 1986, Hanisch et al. 1991). The link between mental models and the operation of states of systems allows these 'internal representations' to help explain human behaviour (Gentner and Stevens 1983, Wickens 1984, Kempton 1986). There are many definitions of mental models and different perspectives from which to consider them (Wilson and Rutherford 1989, Richardson and Ball 2009, Revell and Stanton 2012), so specificity in definition is key (Bainbridge 1992, Revell and Stanton 2012). In this chapter, the concept is best understood in terms of a user mental model (UMM) (Norman 1983) and device model (Kieras and Bovair 1984). That is to say, a mental model held by a user of a specific technology that contains information about the operation and function of that device, and has been accessed and described by an analyst.

Lutzenhiser (1993) argues that human behaviour limits the efficiency of technology introduced to reduce consumption. However, it is often the choice and positioning of technology, as well as usability issues that impede discovery and use by householders (Glad 2012). As the case was made in Chapter 6, poor discoverability of controls is a credible cause of incomplete models (Shipworth et al. 2010, Revell and Stanton 2014) and usability issues of home-heating controls impede effective operation (Brown and Cole 2009, Combe et al. 2011, Shipworth et al. 2010, Peffer et al. 2011, 2013, Glad 2012). These can result in lack of use due to inconvenience (Chapter 5) or fear of complexity (Glad 2012). As background knowledge gained from experience affects the formation of the user's mental model (Johnson-Laird 1983, Moray 1990, Bainbridge 1992), lack of use can further impede appropriate models of heating controls, leading to inappropriate behaviour that further impacts home-heating goals (Kempton 1986).

Norman (1986) proposed that designers could help users operate technological systems more appropriately by designing interfaces that encourage a 'compatible' mental model of the way the system functions. Norman (1986, 2002) proposed that a compatible mental model was necessary to enable users to successfully navigate the 'gulf of evaluation and execution' when interacting with a system. In Chapter 6, the home-heating system was considered from the perspective of Norman's (1986) seven

stages of activity, to determine the components necessary in a compatible mental model for typical home-heating goals. Norman (1983, 1986) emphasized that whilst underpinned by their mental model, users' interaction with technology are ultimately driven by their goals. This is supported by Bainbridge (1992) and Moray (1990) who argue that users' mental model of the system constrains the performance with the system, but user goals influence the resulting strategy adopted.

The review so far has touched on system device design (in terms of discoverability and usability of controls), mental models (at device and system levels), user goals (both for comfort and energy saving) and the strategies adopted (in terms of operation of controls). These are all variables that influence the observed user behaviour with home-heating systems (for a simplified view of how these variables interact, see Figure 8.1). The consequences of user behaviour need to be understood in relation to the intended goals. However, as both Kempton (1986) and Sauer et al. (2009) note, goal achievement is also subject to variables acting in the broader system (e.g. building structure, infiltration, insulation, thermodynamics of the house, external temperature), as depicted in Figure 8.1.

In a realistic setting, there are clear barriers to controlling householders' goals and broader system variables, limiting the insights that can be drawn about the cause of differences in occupants' behaviour. Sauer et al. (2009) overcame these issues with a central heating system simulator that enabled control of house structure, external weather conditions, occupancy and regularity of arrivals and departures. Sauer's (2009) study compared the provision of different types of feedback on consumption levels (revealing predictive feedback to be the most effective at reducing consumption). Sauer's (2009) participants were all university students, and were tasked with creating heating profiles for each room within the simulation, by selecting heating

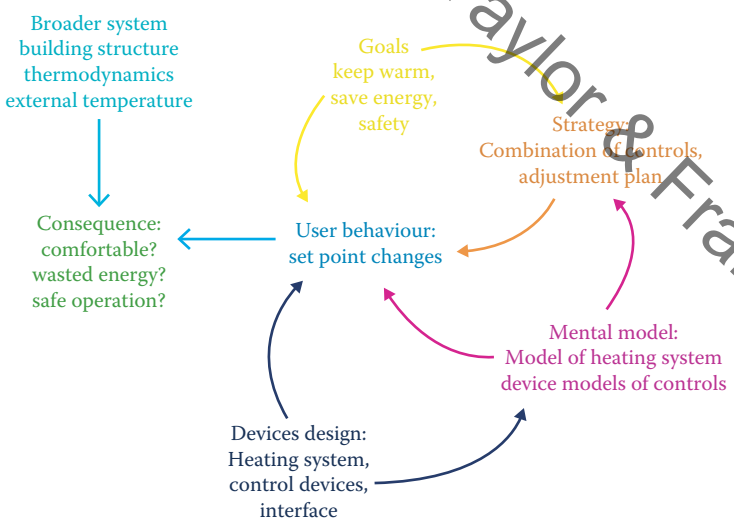


FIGURE 8.1 Different variables that effect home-heating behaviour and its consequences.

periods on a graph-style interface. This differed to the nature of typical home-heating goals, and for greater generalizability of results, Sauer et al. 2009 highlighted that a broader age group and more realistic goals were needed.

Previous work has revealed insights about UMMs of home heating at the individual level from a small, atypical sample selected for minimal experience with home-heating control (Revell and Stanton 2014). Use of a simulation in this study was adopted to allow a larger sample of typical home-heating users in the United Kingdom, enabling greater generalizability of conclusions. In addition, the simulation allowed control of goals, broader system variables and heating system type.

This chapter investigates how differences in interface design affect UMMs of home-heating systems. Using data from a simplified home-heating simulation, it explores how resulting changes in UMMs affect user behaviour with home-heating controls. Finally, differences in goal achievement resulting from different behaviour strategies enabled by different interfaces are compared. Classic studies focused on manipulating users' mental model and comparing their behaviour and performance (e.g. Kieras and Bovair 1984, Hanisch et al. 1991) with a single device. Using an extensive 'training' stage to alter the mental model, change is presumed to have occurred if the expected performance was observed. This study differs in that it relies not on prior training, but on the interface design (plus manuals) to alter the mental model held *at the time of interaction with the system*. This study also measures the success of this manipulation by analysis of user-verified UMM descriptions captured post study, behaviour related to statistically significant differences in UMMs and overall goal achievement. Following from the work of Norman (1986, 2002), the premise of the study predicts that an interface designed to promote a compatible mental model of the heating system, better enables users to achieve their home-heating goals.

8.2 METHOD

The findings over the preceding chapters summarized in the introduction require updates to the original hypotheses set out in Chapter 1. The perspectives by which mental models, behaviour and goal achievement are considered need to be better specified, and deserve definition to the reader. Each of the original hypotheses 1–3 is tackled in this chapter through focus on more granular 'sub-hypotheses' relating to the variables collected. Figure 8.2 maps the sub-hypotheses onto the updated original hypotheses, so it is clear how the results contribute to the original intentions of this research. This diagram also illustrates well the 'cascading' nature of the hypotheses, which informs the experimental design.

Whilst slightly beyond the scope of the main focus of this chapter, the study undertaken provided an opportunity to test an overriding assumption for this research: that the contents of UMM descriptions captured by the Quick Association Check (QuACK) developed in Chapter 3 (as opposed to other methods) are linked to behaviour with home-heating controls. The results of this investigation are also shown in this chapter, to add validity to the novel method deployed. The key focus, however, predicted that participants presented with an interface designed to promote

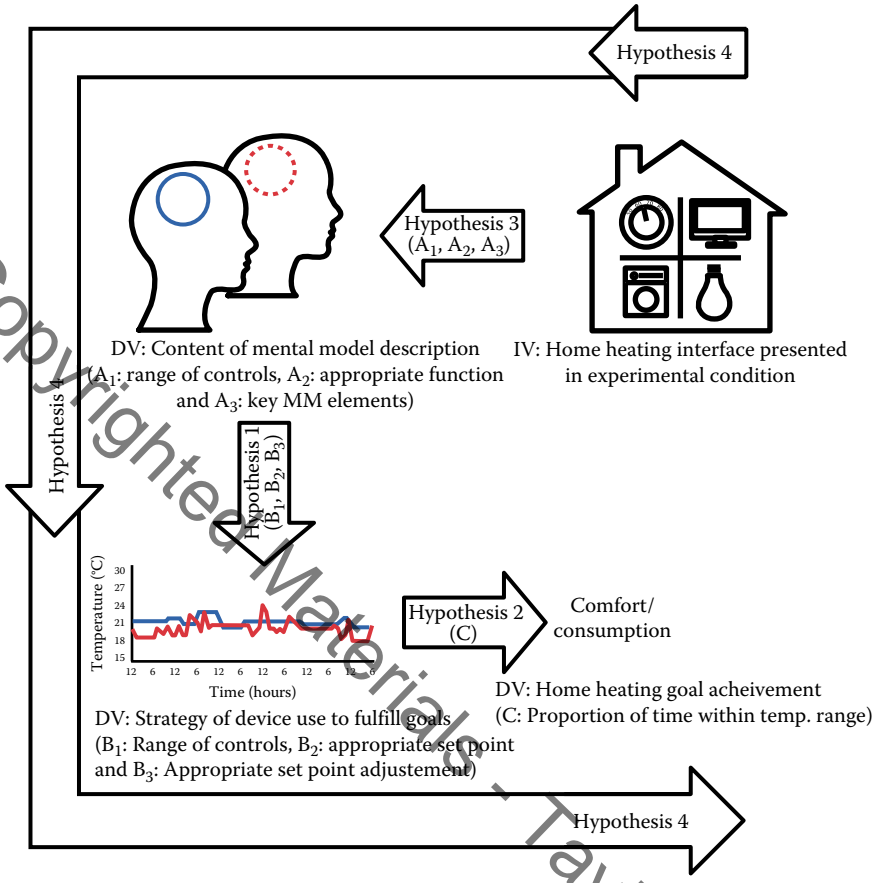


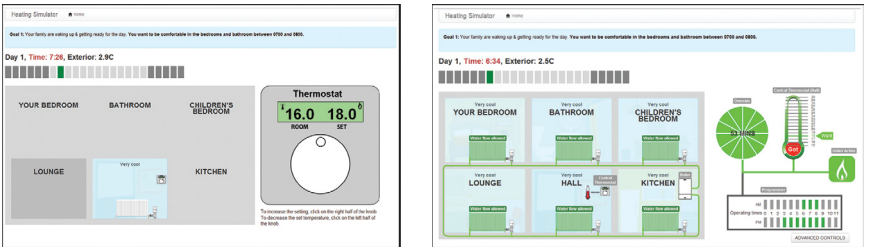
FIGURE 8.2 Diagram to show relationship between theses hypotheses and specific hypotheses tested in the thesis. The linked hypotheses informed the 'cascading' experiment design.

appropriate UMMs would exhibit positive differences in their mental models, behaviour strategies and goal achievement.

8.2.1 Experimental Design

The experiment was a between-subjects design. The independent variable was the version of the interface of a home-heating simulation presented to participants: either 'realistic' or 'design' (see Figure 8.3). Dependent variables related to participants' UMM, their behaviour with controls to achieve goals and the level of goal achievement in line with the hypotheses stated in the following.

Hypothesis 3 states that 'the device design influences mental models of those devices'. To test this hypothesis, dependent variables relating to UMMs included (1) the range of heating controls present in participants' UMM description, (2) the number of appropriate functional models of key controls and (3) the number of key



(a) (b)

FIGURE 8.3 Screen shots of realistic interface (a) and design interface (b).

mental model elements for home-heating operation present in participants' UMM descriptions (see Figure 8.1). Relating to hypothesis 3, this chapter tests the following sub-hypotheses:

1. *Hypothesis* (A_1): Users would have a greater range of heating controls in their mental model description (through enhanced discoverability recommended in Chapters 6 and 7).
2. *Hypothesis* (A_2): Users were more likely to hold an appropriate functional model of key devices (through redesign of controls, recommended by Norman (1986, 2002)).
3. *Hypothesis* (A_3): Users would describe mental models with a greater number of key home-heating system elements (through redesign of the home-heating interface at a system level (see Chapters 3 through 6).

Hypothesis 1 from Chapter 1 stated: 'Users' mental models of devices influence their pattern of device use'. Following the insights gained in the preceding chapters, this hypothesis has been amended to: 'Users' mental models of devices influence their *strategies* of device use'. Dependent variables relating to users' behaviour with the simulation included (1) the range of controls used in participants' strategies and (2) the number and value of set point adjustments time (see Figure 8.1). The following sub-hypotheses are tested to inform Hypothesis 1:

- *Hypothesis* (B_1): Where significant differences in the presence of specific devices in UMMs were found, associated differences would be found in the inclusion of those devices in behaviour strategies.
- *Hypothesis* (B_2): Where significant differences in the appropriateness of functional models at the device and system level were found, associated differences would be found in the adoption of set point values and frequency of adjustment for those devices (following from Kempton (1986) and Norman (1986, 2002).

Hypothesis 2 from Chapter 1 stated: 'Patterns of device use influence the amount of energy consumed over time'. This hypothesis has been amended to '*Strategies* of device use influence the amount of *goal achievement*'. This was done for two

reasons. First, the goals used in the simulation were derived from interview data for realism, and focused more on comfort than consumption. Second, the simulation was not sufficiently sophisticated in its modelling to provide a meaningful comparison of boiler 'on' times. The dependent variable used to represent goal achievement was that the proportion of time target room temperatures were within the target temperature range (see Figure 8.1). This chapter contributes evidence to hypothesis 2 with the following sub-hypothesis:

- *Hypothesis* (*C*): Duration of goal achievement would be greater for participant in the 'design' condition, than for participants presented with the 'realistic' interface (following from Norman 1986, 2002, Moray 1990, Bainbridge 1992).

Figure 8.2 illustrates how each hypothesis is linked to the other, resulting in a 'cascading' experiment design, where goal achievement variables are dependent upon the outcomes of the behaviour variables, which are in turn dependent upon the mental model variables. Where an unexpected outcome from the mental model or behaviour variables occurred, the expectations for behaviour and goal achievement outcomes, respectively, were therefore amended.

8.2.2 Participants

Forty participants took part in this experiment, 20 per condition; 10 males and 10 females were in each condition from ages ranging between 23 and 70 years (Mean = 38). Pairs in each condition were matched by gender, age category and the number of years' experience with central heating (±2 years). Experience with gas central heating (with radiators) ranged from 4 to 40 years, with a median of 12 years. Participants were all native English speakers and were recruited from staff, students and residents local to the University of Southampton. Participants were recruited through posters on university notice boards and website.

8.2.3 Apparatus and Materials

Development of the simulation allowed automatic data collection of energy use and behaviour strategies, and enabled two different interface views to be presented to participants. The simulation versions were presented on a Samsung LE40M67BD 40″ TV monitor attached to a DELL Latitude E6400 laptop connected to the Internet page hosting the simulation and controlled with a mouse. The 'realistic' version reflected the design of a typical gas central heating system, using the setup described in Chapter 3 and Revell and Stanton (2014). The 'design' interface was constructed to promote a compatible mental model of the home-heating system, following recommendations from Chapter 6. The duration of the simulation activity, goals presented and duration of goals were matched for both conditions. The simulation automatically collected data relating to (1) set point adjustments (2) 'boiler on' periods and (3) 'device control mouse clicks'. A 'play mode' was available for each version of the simulation that did not collect data. User manuals specific to each simulation

version were provided. A consent form, participant information sheet and participant instructions were provided along with a pen for the subject. Parts 1 and 3 of the QuACk interview script (Appendix A) were used to interview participants to collect background information and elicit the users' mental model. The former was recorded using a pen, and the latter was recorded on A3 plain paper and square post-it notes, using a marker pen. The interview was recorded on an Olympus VN-2100PC audio recorder.

8.2.4 Procedure

Subjects undertook the study individually, sitting at a desk in front of the monitor. After providing health and safety information, the subject was asked to fill in the consent form, read instructions for participation and fill in the demographic information section on this sheet using the pen provided. The experimenter then checked understanding, verbally reiterated the key points of the experiment, told the subject which condition they were allocated to and collected the completed consent forms.

Before starting data collection, the participant was provided with the user manuals for their experimental condition and exposed to a 5 minute practice session, using their simulation version in 'play' mode. At the start of the practice session, the experimenter pointed out key elements of the interface with reference to a script appropriate to the experimental condition. Those in the 'realistic' condition were asked to imagine they were operating the home-heating controls as if they were in their own home setting. For the 'design' condition, they were asked to imagine they had been provided with a digital interface to control the existing heating controls in their home setting. The experimenter then remained at a desk behind and to the left of the subject to allow the subject to practise independently. During the practice session only, the experimenter responded to any questions by the participant about the procedure for the study or layout of the controls and displays. Participants were referred to the user guides provided if they asked any questions about the function or operation of control devices. At the end of 5 minutes, the experimenter stopped the practice session and asked the subject if they understood what they were required to do and clarified any confusion.

After the practice session, the simulation appropriate to the experimental condition was started by the experimenter. It ran for 22 minutes with a home-heating goal presented textually at the top of the screen every 2 minutes. The goals represented typical home-heating goals for a family with young children and were the same in each condition. To direct action, specific time frames, rooms or objectives were provided in the goal description. To signal a change of goal, the screen flashed yellow. On reading the goal, the participant was required to decide what adjustment of heating controls was necessary to achieve the goal, and perform any operation they thought appropriate (even if this resulted in no adjustment). If a subject had not completed their intended adjustments before the next goal was presented, they were to move onto adjustments for the new goal. At the end of the experiment, the screen flashed yellow and text (where the goals had formally been presented) informed the subject that the simulation was over.

Once the paper-based questionnaires had been completed, the experimenter turned off the TV monitor and removed the questionnaires and user guides to prevent them

acting as prompts to the structured interview. The experimenter sat to the right of the subject at the same desk to conduct the structured interview. The audio recorder was switched on and placed on the desk. The experimenter explained the structure and purpose of the interview using the instructions for interviewer provided on QuACk. During part 1 of QuACk, the participants' answers to questions about their background experience with home-heating systems and attitudes to home-heating use were recorded by pen on the interview script by the experimenter. The participants' preferred terminology for home-heating components were written on individual post-it notes in marker pen. During part 3 of QuACk, the experimenter made clear to the subject that they should answer questions based on their experience with the simulation, not their own home-heating system. Participants were told they could refer to their own heating system as a means of comparison (e.g. 'like my heating system, it worked like….' or 'unlike my thermostat at home, it worked like….') but only descriptions relating to the simulation would be represented. Following the interview prompts, a diagram was constructed on the A3 paper by positioning post-it notes and linking and annotating by pen to represent the participants' UMM of the heating system presented on the simulation. On completion of this diagram, the experimenter paraphrased the mental model description to check for understanding and provide an opportunity for amendments. When any amendments were complete, a verification stage was undertaken, whereby each element, link and rule on the diagram was considered in turn and the subject asked to identify how confident they were that this represented what they believed. On completion of the verification stage, the subject was debriefed and paid £10 for their participation.

8.3 RESULTS

This section describes data gathered from UMM descriptions, variables relating to their behaviour with heating controls in the simulation and goal attainment based on room temperatures. The presentation of information will be split into three sections relating to these areas. In each section, the key hypotheses will be explored by tabulating descriptive statistics, then applying the Mann–Whitney U test for non-parametric data to determine the significance of differences in the realistic and design groups. Non-parametric tests were necessary as the data were not normally distributed (Field 2000). To drill into the detail of any significant differences, the data were illustrated in graphs and diagrams. Where graphs revealed the likely focus of differences, the Pearson's chi-square for categorical data was used to check the statistical significance. Where the data did not meet the criteria for Pearson's chi-square (i.e. expected values per cell were less than 5), the Fisher's exact test was applied.

8.3.1 User Mental Models of Home-Heating Simulation

Hypothesis A_1: Greater range of home-heating controls present in participants' UMMs when exposed to the design condition.

To determine if the design interface increased the discoverability of the home-heating controls, hypothesis A_1 predicted that the design condition would promote a greater

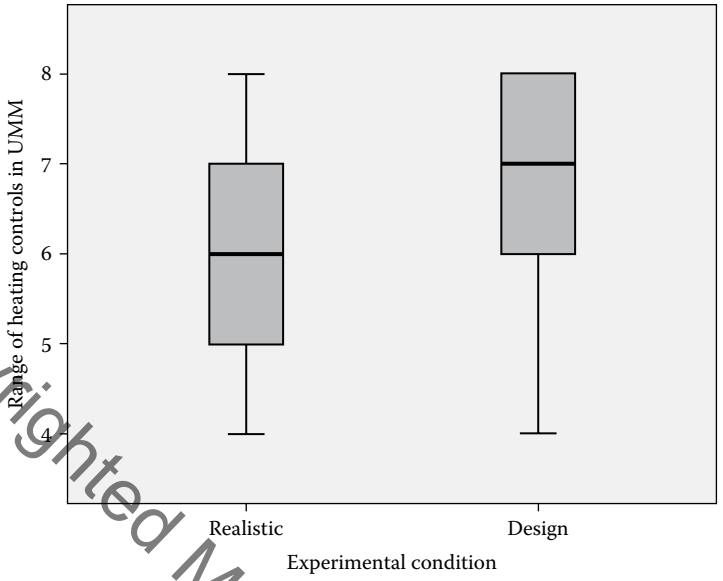


FIGURE 8.4 Range of heating controls found in user mental models.

range of heating controls in participants' UMM descriptions. Figure 8.4 shows a boxplot illustrating the median, interquartile range and minimum and maximum of the number of controls described by participants in their UMM descriptions. Using the Mann–Whitney test, it was found that the number of heating controls captured in participants' UMMs was significantly greater in the design condition than in the realistic condition ($U = 124.5$, $Z = 2.092$, $p < 0.05$, $r = -0.33$). The design condition therefore encouraged participants to include a greater range of heating controls, supporting hypothesis A_1.

Figures 8.5 and 8.6 group data indicating the presence of heating controls in UMMs by 'key controls' and 'advanced controls'. Whilst Figure 8.5 revealed little difference in the discoverability of key controls between conditions, greater variation was seen in number and type of advanced controls present in UMMs (Figure 8.6). Chi-square tests revealed a significantly greater use of the 'holiday button' ($\chi^2 = 10.99$, d.f. = 1, $p < 0.001$) and 'frost protection' ($\chi^2 = 32.40$, d.f. = 1, $p < 0.0001$) controls in the design condition.

Hypothesis A_2: Improved functional models of key devices are held by participants in the design condition.

Hypothesis A_2 predicted that participants in the design condition would hold more appropriate functional models of key devices than those in the realistic condition. Figure 8.7 shows a boxplot illustrating the median, interquartile range and minimum and maximum of appropriate functions described by participants to key controls in their UMM diagrams. Results of the Mann–Whitney U test found the number

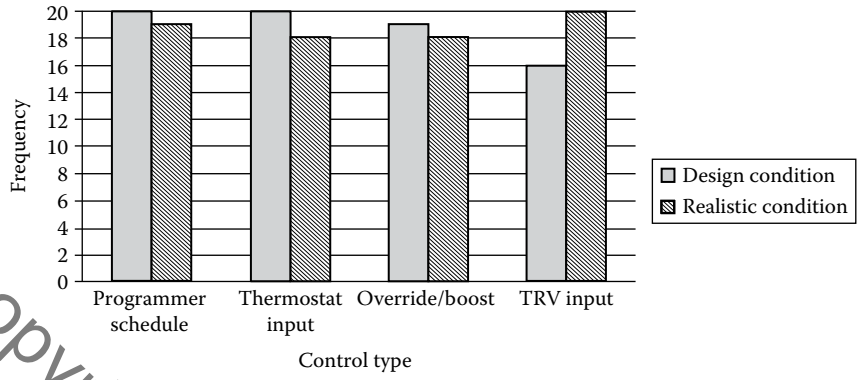


FIGURE 8.5 Frequency of participants who described key heating controls in user mental model descriptions.

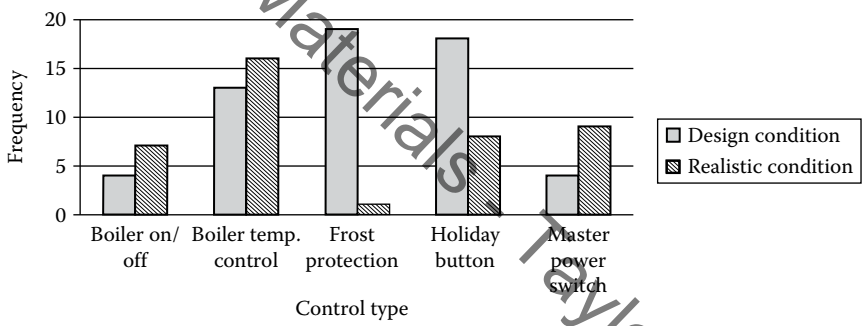


FIGURE 8.6 Frequency of participants who described advanced heating controls in user mental model description.

of appropriate functional models of key controls captured in participants UMMs was significantly greater in the design condition than in the realistic condition ($U = 108.00$, $Z = -2.617$, $p < 0.01$). To take into account variations in the number of key controls present in UMMs, a chi-square test was also performed comparing appropriate and inappropriate models, revealing a statistically significant difference ($\chi^2 = 7.335362$, d.f. = 1, $p < 0.01$). Both tests support hypothesis A_2, that participants in the design condition were more likely to have an appropriate functional model of key devices.

The graph in Figure 8.8 illustrates differences in the appropriateness of functional models for the different controls described in UMMs. This shows that the programmer schedule is functionally understood by all participants, regardless of condition. In this sample, more participants in the design condition held an appropriate functional model for the boost and thermostat controls, but in both conditions,

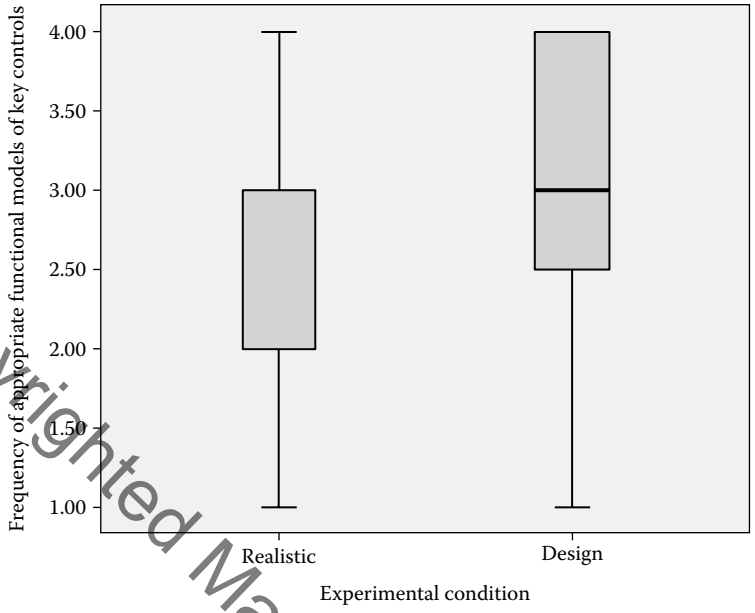


FIGURE 8.7 Frequency of appropriate functional models for key controls.

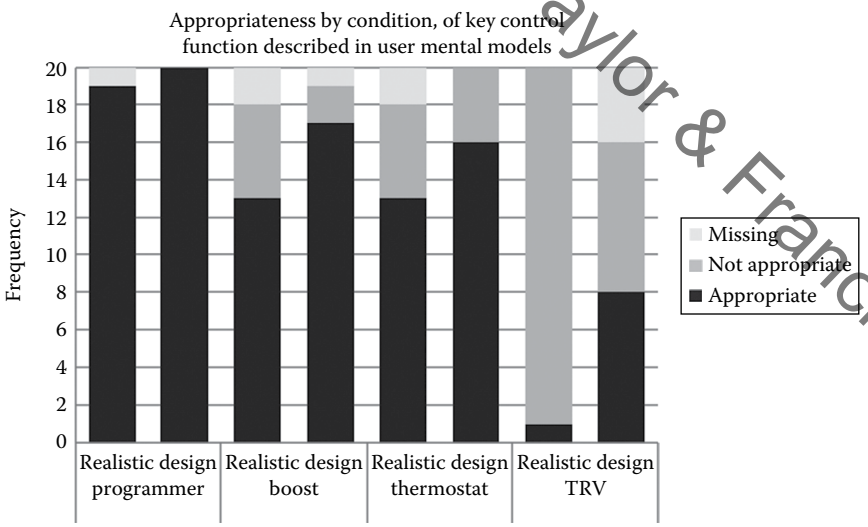


FIGURE 8.8 Graph to compare the frequency of appropriate and inappropriate functions assigned to key controls.

the majority had an appropriate model. Figure 8.8 shows that a statistically significant difference was found in the functional model held for the TRV control (χ^2 = 9.60, d.f. = 1, $p < 0.01$). Almost half of participants in the design condition had an appropriate device model of the TRV however, compared to a single participant in the realistic condition.

Hypothesis A_3: Improved number of key home-heating system elements described in UMMs of participants in the design condition.

Hypothesis A_3 predicted that participants from the design condition would describe a greater number of key home-heating system elements. Figure 8.9 shows a boxplot illustrating the median, interquartile range and minimum and maximum of key system elements described by participants in their UMM diagrams. It was found, using the Mann–Whitney test, that the number of key system elements present in UMMs was significantly greater in the design condition than in the realistic condition (U = 124.5, Z = 2.092, $p < 0.05$, r = −0.33), supporting hypothesis A_3.

The graph in Figure 8.10 compares the frequency of each key element found in UMMs. The largest differences relate to increases in design condition in the presence of the 'conditional rule', 'TRV feedback link' and TRV active indicator. Chi-square tests showed these differences were significant for the conditional rule (χ^2 = 5.226667, d.f. = 1, $p < 0.05$) and the TRV feedback link (χ^2 = 7.025090, d.f. = 1, $p < 0.01$). Fisher's exact test showed a significant difference for presence of the TRV active indicator in UMMs ($p < 0.01$). This indicates the design condition was more effective at encouraging increases in the presence of these elements.

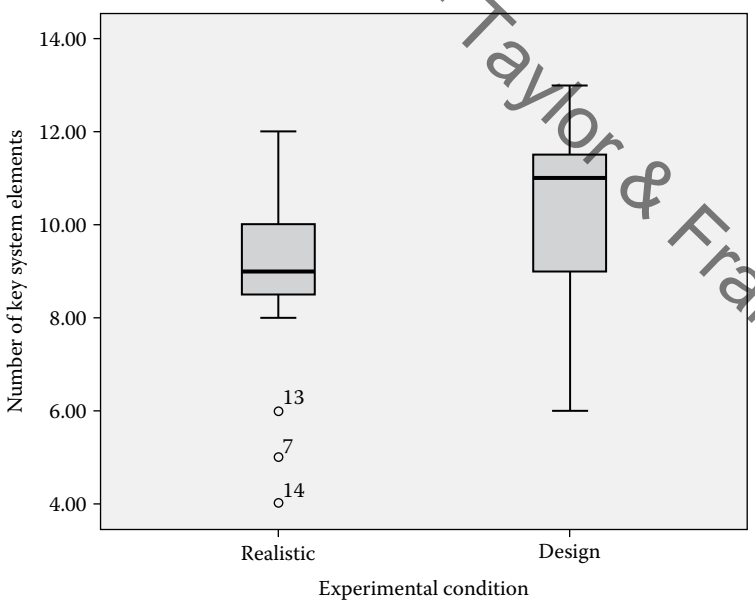


FIGURE 8.9 Number of key system elements present in UMM descriptions.

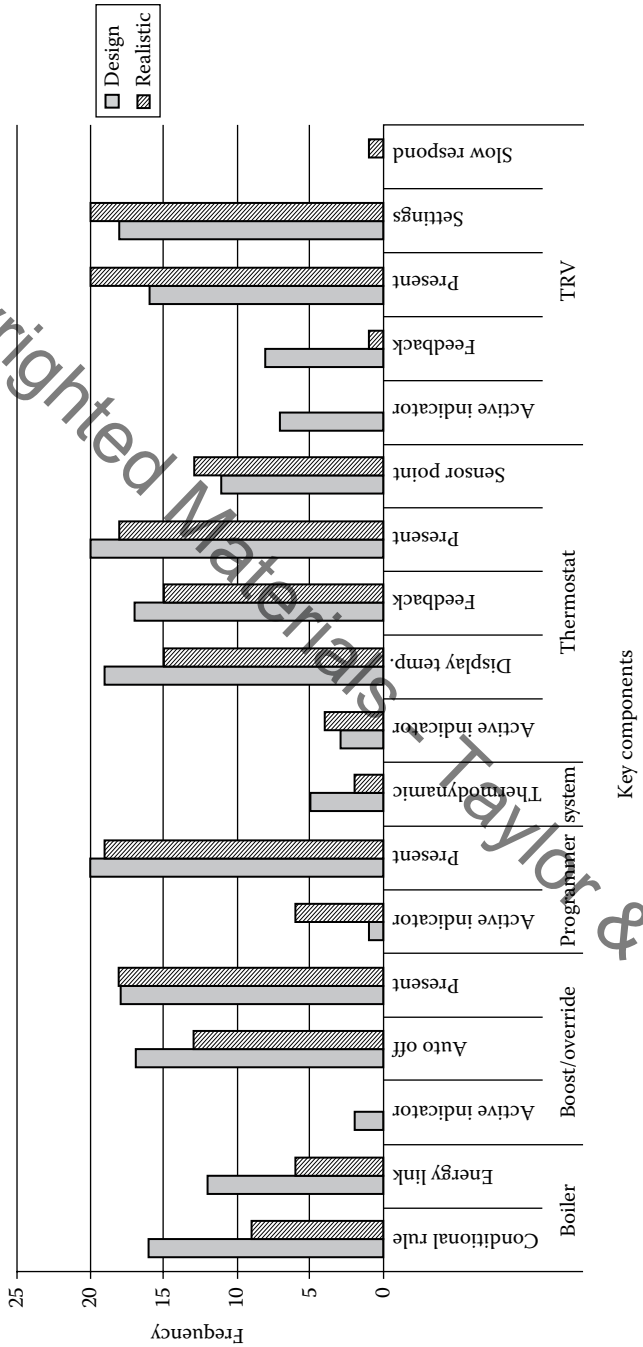


FIGURE 8.10 Frequency of key system elements present in UMM descriptions.

8.3.2 User Behaviour with Home-Heating Simulation

Hypotheses B_1 and B_2 predicted that participants in the design condition would adopt more appropriate behaviour strategies in line with the content of UMMs. For this study, significant differences were found between conditions with the prevalence of the frost protection and holiday buttons devices in UMMs. Hypothesis B_1 therefore predicts that the (i) frost protection and (ii) holiday buttons are included in more behaviour strategies in the design condition than in the realistic condition. Significant increases in the appropriateness of the functional model (at the device level) of the TRV and inclusion of the TRV feedback link (at the system level) were found in the design condition. Finally, a significant increase in the occurrence of the 'conditional rule' at the system level was found in the UMMs of participants in the design condition. Understanding the conditional rule enables deliberate control of boiler activation. Hypothesis B_2 therefore predicts that (i) the TRVs in the design condition will be operated in a way consistent with a temperature-sensing feedback device, and that (ii) effective boiler control would occur more in the design condition.

8.3.2.1 Underlying Assumption for Study

This study is underpinned by an assumption that if a control is present in a UMM description, this control is available for participants to include in behaviour strategies. This underlying assumption predicted that if a heating control is present in a UMM, it will be present (where the goal requires), in a home-heating behaviour strategy. Chi-square tests revealed a highly significant difference for both the design condition ($\chi^2 = 78.268$, d.f. = 1, $p < 0.0001$) and the realistic condition ($\chi^2 = 90.496$, d.f. = 1, $p < 0.0001$). For the design condition, 133 control elements were present in UMMs and 90.2% of these controls were used during the simulation; 47 controls were absent from UMMs, of which 76.6% were also absent in participants' behaviour in the simulation. The same trend was found in the realistic condition, with 89.7% of 116 controls present in UMMs being used in the simulation. Similarly, of the 111 controls absent from the UMMs, 77.9% were also missing from behaviour strategies (see Figure 8.11).This supports the assumption that contents of UMM descriptions captured by QuACK are linked to behaviour with home-heating controls.

Hypotheses B_1(i) and (ii): Differences in the inclusion of specific devices in behaviour strategies.

The graph in Figure 8.12 shows that the majority of all participants used all four key controls (programmer, thermostat, boost and TRV). It also reveals a considerable difference in use of the frost protection and holiday buttons. No participants in the realistic condition used the frost control button, compared to almost all participants in the design condition. Chi-square tests showed these differences were statistically significant for the frost control button ($\chi^2 = 36.190$, d.f. = 1, $p < 0.0001$) and for the holiday button ($\chi^2 = 7.619$, d.f. = 1, $p < 0.01$), supporting hypotheses B_1(i) and (ii).

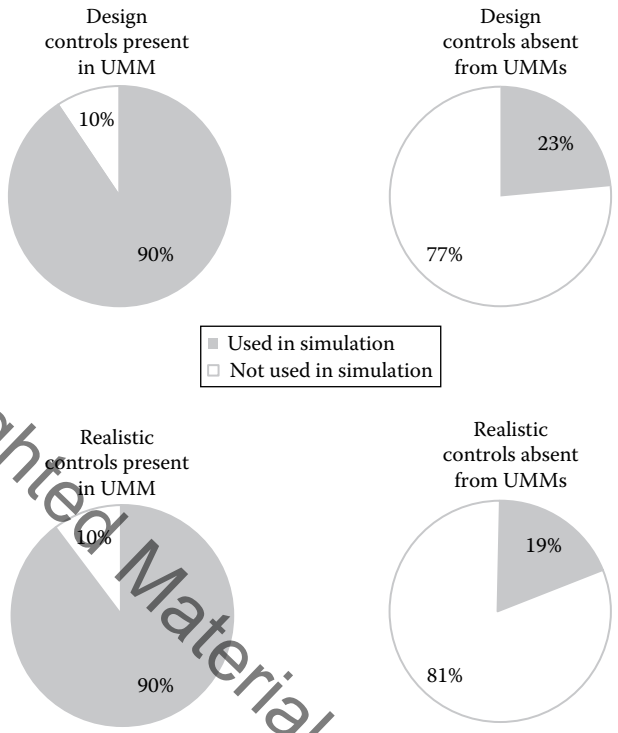


FIGURE 8.11 Proportion of controls used in simulation, depending on presence in UMM.

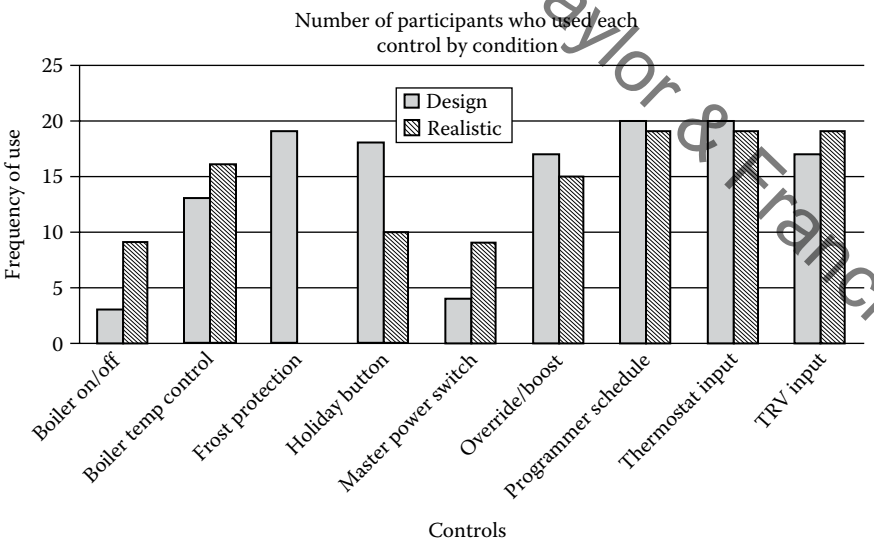


FIGURE 8.12 The frequency of use for controls.

Hypothesis B_2: The adoption of more appropriate set point values and frequency of adjustment for specific devices.

Hypothesis $B_2(i)$: TRV operation consistent with a temperature-sensing feedback device.

Hypothesis B_2(i) predicted that TRV operations in the design condition will be operated consistent with a temperature-sensing feedback device. Less frequent and less extreme set point adjustments are more consistent with appropriate operation of a temperature-sensing feedback device for typical home-heating goals (Kempton 1986).

Figures 8.13 and 8.14 show boxplots illustrating the median, interquartile range and minimum and maximum of the total frequency of TRV set point adjustments, and mean range of TRV set point choices, respectively. Performing a Mann–Whitney test for non-parametric data failed to reveal a significant result (U = 158.000000, Z = −1.137224, p = not significant) for frequency of adjustment, but showed a statistically significant difference in the range of TRV set points (U = 110.500, Z = −2.428742, $p < 0.05$).

Figure 8.15 shows the adjustment strategies of TRVs for each condition. The set point choices in the realistic condition are more extreme and vary in direction far more than in the design condition, which shows a more subdued pattern. The greatest variation can be seen with the lounge, kitchen and children's bedroom, reflecting the target rooms in the majority of the provided goals.

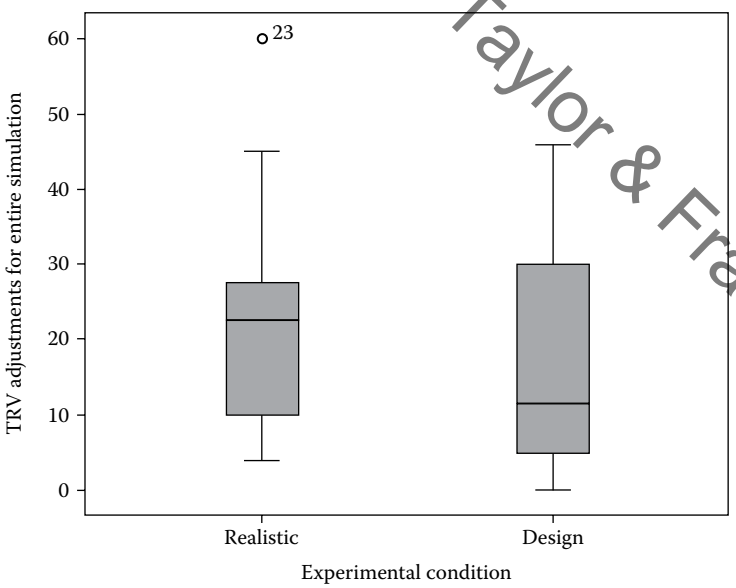


FIGURE 8.13 Frequency of TRV set point adjustments.

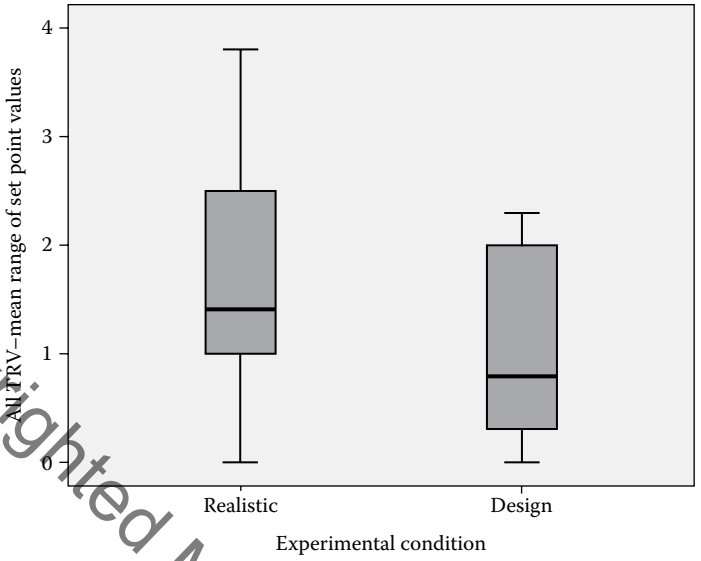


FIGURE 8.14 Mean range of TRV set point values.

Hypothesis B_2(ii): Differences in control of boiler activation.

Hypothesis B_2(ii) predicts that effective boiler control would occur more in the design condition. An essential prerequisite for boiler activation is for the thermostat to hold a higher set point than the hall room temperature, and a lower set point for deactivation. To test the statistical significance of this hypothesis, an independent samples t-test for parametric data was performed to compare the percentage of thermostat set point value changes that crossed the current hall temperature value. The results showed a statistically significant increase in control of boiler activation in the design condition ($t = 3.296$, d.f. = 37, $p < 0.01$), than in the realistic condition, supporting *hypothesis B_2(ii) (Figures 8.16 and 8.17).

Hypothesis C: Data relating to goal achievement through target temperature durations.

Goal achievement was based on target rooms achieving room temperatures within a target temperature range during a target time period. Where the target related to multiple rooms, the median room was used as the basis for measuring the duration of goal achievement as it reflected central tendency for non-normally distributed data. As target goal durations differed, to prevent this becoming a confounding variable, the proportion of time each goal was achieved was used. These were summed for 18 goals and converted into a percentage of overall goal achievement. Figure 8.18 shows boxplots illustrating the median, interquartile range and minimum and maximum of goal achievement. A Mann–Whitney test was undertaken, showing a statistically significant increase in goal achievement in the design condition ($U = 125.500$, $Z = -2.015$, $p < 0.05$), supporting hypothesis C.

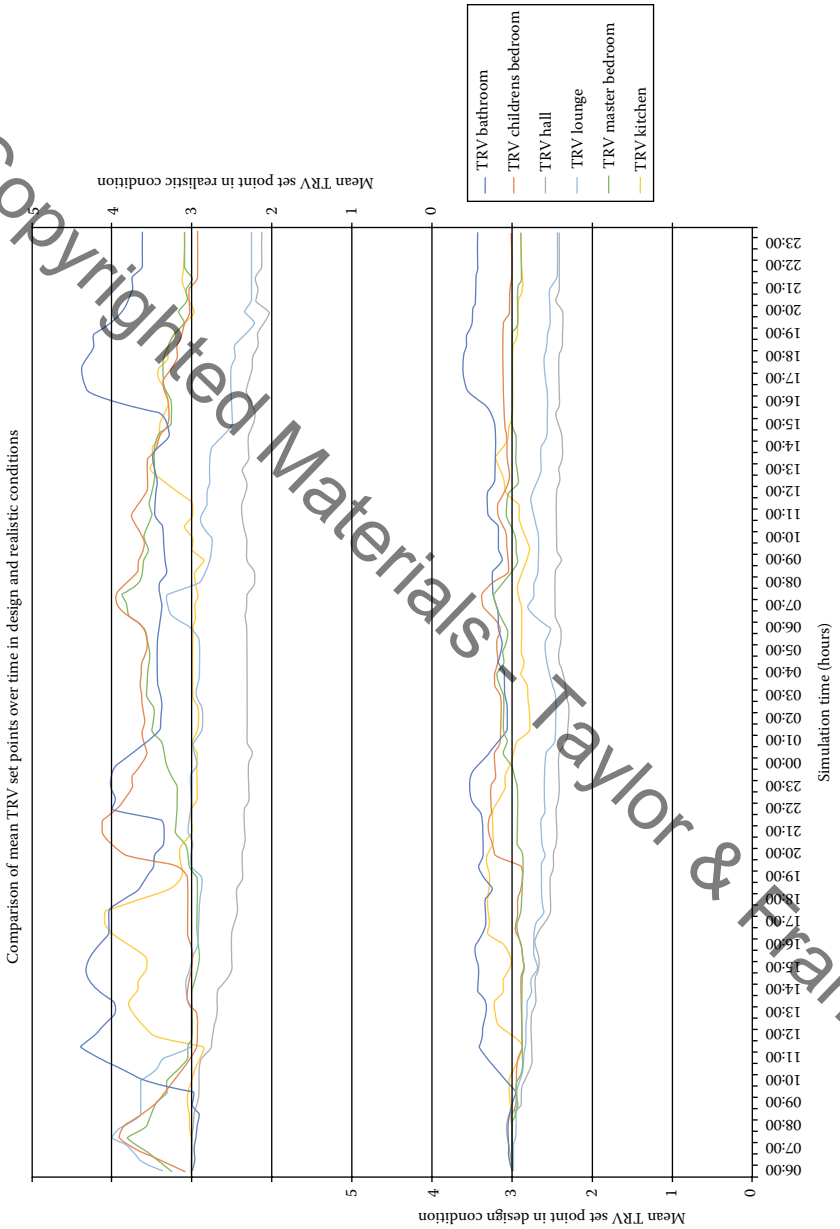


FIGURE 8.15 Frequency of use and set point choice over time of TRVs.

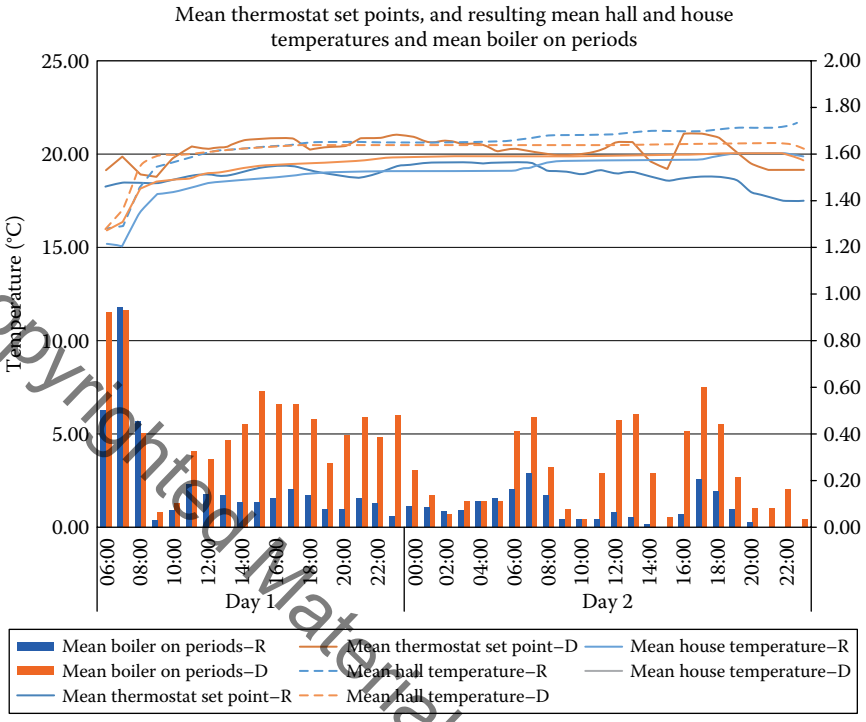


FIGURE 8.16 Control of boiler activation by thermostat adjustments.

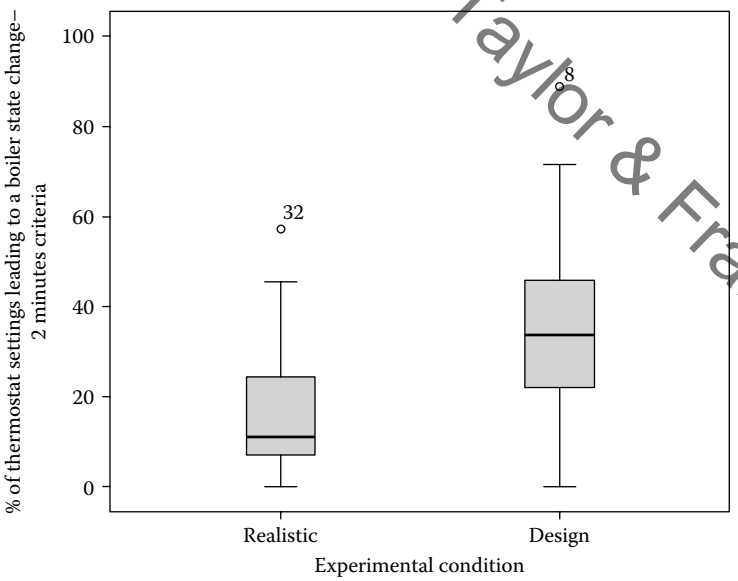


FIGURE 8.17 Percentage of thermostat set point choices leading to boiler state change.

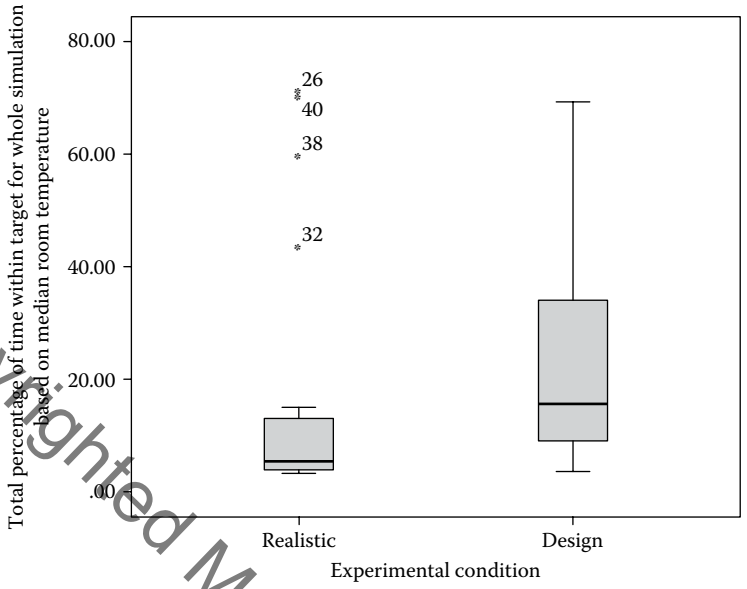


FIGURE 8.18 Total proportion of time within goal temperature range.

8.4 DISCUSSION

The underlying assumption for this study derived from Kempton (1986), and extended by Revell and Stanton (2014) in Chapter 3, is that householders' UMMs of their home-heating system affect their behaviour with that system. This underlying assumption was supported by comparing the presence of heating controls in participants' UMMs together with at least one instance of adjustment during the experiment. The statistically significant difference between the use of controls based on presence in UMM descriptions was not reliant on the type of interface presented, as comparable results were found for both conditions (see Figure 8.11). Identification of a link between UMMS and behaviour complements previous literature from other domains (Gentner and Stephens 1983, Kieras and Bovair 1984, Wickens 1984, Hanisch et al. 1991). The QuACk method for capturing UMMs is proposed as a useful tool in understanding and predicting whether a device is likely to be used. The level and appropriateness of use is subject to more detailed analysis, however, and the subsequent hypotheses go some way in exploring the value of the method in that respect.

8.4.1 Improved Discoverability of Home-Heating Controls

The greater range of heating controls present in UMMs of participants in the design condition (supporting hypothesis A_1) suggests that the changes in the home-heating interface improved the discoverability of controls overall (see Figure 8.4). The difference in discoverability of home-heating controls in the realistic condition supports findings by Shipworth et al. (2009), and Revell and Stanton (2014) in Chapter 3.

When polling 427 English homes, Shipworth et al. (2009) found many people did not recognize many of their home-heating controls. In Chapter 3, Revell and Stanton (2014) found key heating controls were missing from UMMs of occupants new to heating control. Similarly, Brown and Cole (2009), who compared heating controls in standard and 'green' office buildings, found that the highest reason occupants had for not using controls was 'Controls don't exist' followed by 'I don't know where they are'. The implications of a greater number of controls described in UMM descriptions in the design condition are that the appropriate control(s) are available in the UMM to draw upon when determining an action specification to fulfil a given goal (Norman 1986). However, it should be noted that since both conditions had a high prevalence of the key heating controls (see Figure 8.5), the differences in the inclusion of controls in behaviour strategies, attributable to discoverability alone, would be most marked when attempting to fulfil goals met by 'advanced controls' (in this case, the frost protection and holiday buttons). Improvements in discoverability of less familiar controls by representing physically distributed devices in a single 'control panel' could hold the key to fulfilling the potential of energy-saving systems and monitoring technology (e.g. smart meters). Glad's (2012) study on the benefits of new energy systems in Swedish housing cited that lack of discoverability of devices thwarted expected improvements in CO_2 saving.

8.4.2 More Appropriate Mental Models

Participants from the design condition were found to have more appropriate functional models of key heating controls, and key system elements supporting hypothesis A_2 (users were more likely to hold an appropriate functional model of key devices) and hypothesis A_3 (users would describe a greater number of key mental model elements), respectively (see Figures 8.7 and 8.9). These results are particularly encouraging, given the lack of formal 'training' to promote specific model types in comparison to studies like Kieras and Bovair (1984) who provided extensive training, followed by a test for comprehension and a re-test a week later to ensure knowledge retention, and Halasz and Moran (1983) who provided 30 minutes training in advance of their experiment. Both of those studies used novice participants who did not have to 'overcome' an existing mental model of the test device. This study, in comparison, comprised of participants with between 5 and 40 years' experience of home-heating controls, so existing knowledge structures would need amendment (Johnson-Laird 1983). The largest difference between conditions was found with the functional model of the TRV as a feedback device and the conditional rule for the boiler, which requires both the thermostat and the programmer (or boost control) to be 'calling for heat' to trigger boiler activation. The function of the TRV was misunderstood by most participants, supporting the findings found in Chapter 6. Improvements seen in the appropriateness of the TRV functional model backs the view by Kempton (1986) that when the operation of controls is 'visible', the correct model is adopted. Improvements in the presence of the 'conditional rule' element reflect advice from Kieras and Bovair (1984:271) that the most useful information to provide users is 'specific items of system topology that relate the controls to the component's and possible paths of power flow'. Contrary to the work by Kempton (1986), Norman (2002) and Peffer (2011),

but in support of Revell and Stanton (2014) described in Chapter 3, most participants provided an appropriate functional model for the thermostat device (see Figure 8.8) suggesting present-day UK householders *do* understand this device. The programmer control was not only included in all but one of participants' UMMs, but its purpose was well understood. No inappropriate functional models were given by participants for this device, so issues with operation are likely to result from inappropriate models at a system level, or the known usability issues with typical device designs seen in the literature (e.g. Combe et al. 2011, Peffer et al. 2011). That amendments to existing UMMs of home-heating systems can be achieved within a very short period of time (25 minutes of accelerated interaction) without 'formal instruction' has favourable implications for using UMM-based design for home-heating systems. This is likely to be generalizable to other systems where inappropriate UMMs have been shown to result in inappropriate behaviour.

8.4.3 Increased Use of Frost Protection and Holiday Buttons

Hypothesis B_1 represented differences in behaviour with heating controls following from statistically significant differences in UMM content. At the device level, this focused on difference by condition in the use of the frost protection and holiday buttons. As expected, the design condition encouraged significantly more interaction with these controls than the realistic condition. Whilst the former device is concerned with safety, increased use of the holiday button (that replaces the existing programmer schedule with minimal on periods to reflect lack of occupancy) is clearly relevant in terms of reducing energy consumption. How significant this is in terms of other recommended behaviour requires further research. That lesser known controls are used when present in UMMs, and accessible on an interface, lends support to the view that increasing the discoverability of energy-saving technology could indeed increase inclusion in behaviour strategies, potentially realizing the energy-saving potential promised by technology. This would, however, still rely on the adoption of corresponding energy-saving goals (Norman 1983, Moray 1990, Bainbridge 1992, Chapter 6).

8.4.4 More Appropriate Behaviour with TRV Controls

Following from analysis of the appropriateness of functional models of key control devices, corresponding differences by condition in the behaviour patterns relating to TRVs were expected (hypothesis B_2(i)). Figure 8.15 showed TRV adjustments over time for both conditions and it was clear that expert advice promoting static set points (Revell and Stanton 2016) were not evident in either condition. However, this can be explained by the style of the prescribed goals for 'comfortable' conditions in specific rooms, which may have convinced participants that this level of custom control was possible, as well as the lower effort levels necessary to make changes to distributed controls when presented on a single interface. The results in this study showed increased interaction with TRVs in the realistic condition, but this was not statistically significant (see Figure 8.13). A statistically significant difference was found, though, in the mean range of set point adjustments (see Figure 8.14).

The more extreme set point choices found in the realistic condition are reminiscent of the behaviour pattern described by Kempton (1986) for 'valve' model holders of a central thermostat, whilst the moderate adjustments made in the design condition are more in line with Kempton's (1986) behaviour pattern corresponding to a 'feedback' model. Further examination of the data revealed that a 'valve' functional model for the TRV was held by the majority of participants in the realistic condition. This supports the view given in Chapter 4 and by Revell and Stanton (2014) that Kempton's insight can be applied to alternate control devices, and suggests that differences in UMMs in terms of function can explain behaviour patterns. Whilst TRV is not cited in the literature as a key player affecting domestic energy consumption, issues of unnecessary adjustment may not be considered a problem at the device level. However, when viewing heating controls as an integrated system, Chapter 5 (Revell and Stanton 2015) showed that inappropriate UMMs of other controls, such as the TRV, can have a significant negative impact on both comfort and energy consumption.

8.4.5 Greater Control of Boiler Activation

Hypothesis B_2(ii) examined differences in boiler activation predicted from significant differences in the presence of the 'conditional rule' in UMMs. To intentionally fulfil heating goals and manage energy consumption, it is necessary for the participant to have an understanding of the link between the set points of the thermostat, its relationship to sensed temperature in order to 'call for heat' and its dependency on the setting of the programmer and boost for boiler activation. The graph in Figure 8.16 shows the mean number of thermostat adjustments over time for both conditions. Whilst a static thermostat set point is encouraged in manuals and by expert advice (Revell and Stanton 2016), the appropriate set point choice requires an appreciation of the thermodynamics of the house structure. Crossman and Cooke (1974) emphasized that operator control of dynamic systems requires sufficient time for experiment and observation, which was not provided to participants in this experiment. In addition, the (unintended) non-typical thermodynamic model for the simulation resulted in particularly high temperatures in the hall where the central thermostat was located. This meant that far higher set point values would be necessary to activate the boiler than participants would be used to selecting at home. Expectations for a static thermostat pattern were unreasonable in these circumstances. The mean data represented in Figure 8.15, suggest that participants in the design condition were able to influence boiler activation by varying the thermostat set point (thick orange line) above and below the hall temperature value (dashed orange line). Participants in the realistic condition display very little influence over boiler activation as after 0800, day 1, the mean thermostat set point (thick blue line) remains below the mean hall temperature (dashed blue line). For the majority of participants in the realistic condition, therefore, the boiler would be inactive for a substantial part of the simulation, despite repeated set point adjustments in both directions. Results from a t-test supported the hypothesis B_2(ii) that participants in the design condition operated a greater level of intentional control over the boiler. This result supports the findings of Kieras and Bovair (1984) who found that participants with an appropriate UMM

of the system engaged in very few 'nonsense' actions, favouring behaviours that were consistent with the device model. This result further supports what Revell and Stanton (2014) described in Chapter 3, that UMMs of home heating must be considered at the system level, since the majority of participants in the realistic condition had an appropriate functional UMM for the thermostat device in isolation. Usability or mental model promoting design initiatives focused solely at the single device level (e.g. programmer, thermostat, TRV) are unlikely to maximize desired benefits for comfort or consumption goals.

8.4.6 Increased Goal Achievement

Participants in the design condition were also significantly more successful at achieving the goals provided (hypothesis C). This result supports the work of Kieras and Bovair (1984) and Hanisch et al. (1991). To be as realistic as possible, comfort goals made up a large part of goals, and energy conservation was the focus only when the house was to be unoccupied. The proportion of goal achievement was relatively low in both conditions compared to the results of Sauer et al. (2009), who found the proportion to be between 73% and 94%. Sauer et al.'s (2009) study tasked participants with choosing setting in advance to achieve a specific daily profile, allowing greater opportunity for planning and amendment. In contrast, this study presented a succession of changing goals that incorporated not only typical planned changes in comfort goals, but a more realistic 'ad hoc' adjustment of goals throughout the day, which were not necessarily achievable. This study also placed participants with greater time pressure to make their heating adjustments. Previous work in Chapters 3 and 5 has led to the conclusion that a 'gap' between heating expectations (in terms of speed of achieving comfort goals, consistency through the house in space heating temperature and effectiveness of 'custom' room heating) and what typical UK heating systems can deliver has a large influence on encouraging inappropriate behaviour. Greater generalizability of the results to everyday behaviour is therefore possible by providing 'realistic' goals in the simulation. The measure for goal achievement was based on temperature values for rooms rather than consumption or deployment of appropriate action sequence with controls. This meant that appropriate choices in behaviour (e.g. programmer settings, or deployment of the frost protection and holiday buttons) did not get recognized unless there was an impact on room temperature values within the target time period. Further analysis that matches behaviour strategies to specific goals will be the focus of further work. The most important aspect of this study is that performance improvements can be explained by better control of boiler activation following from increased understanding of the conditional rule, resulting from design changes in the interface and instruction manual. This result is highly encouraging as it demonstrates that home-heating performance can be affected through design for a given set of goals.

8.4.6.1 Limitations of Study

There were limitations of the thermodynamic model used for the simulation. The mean house temperature over time for each condition shows a gradual increase in value throughout the simulation (see Figure 8.16). Given the variations in the

controls used and boiler on periods, this suggests the insulation parameters were unusually high, preventing heat loss from occurring. This has implications in terms of energy-saving goal achievement based on room temperature, since a deliberate drop in room temperature cannot be contrived by behaviour strategy. The thermodynamic characteristics of the simulation were therefore atypical and limit the generalizability of goal achievement findings, particularly relating to comparison in energy conservation. To reduce the variables under analysis, the simulation did not allow the participant to control ventilation or heat flow within the house through opening and closing of doors and windows. Comments from participants indicated that this would have made up part of their strategy for controlling room temperatures, suggesting that restricting behaviour to heating controls only may not be representative for how people manage space heating in the home.

8.5 CONCLUSIONS

Differences in the design of an interface have been shown to change the content of mental model descriptions. The hypotheses that correspond with the design condition to improve the functional mental model of heating controls, the discoverability of heating controls and the number of key system elements were supported. Key differences in users' mental model descriptions focused on the TRV control (appropriate model and feedback link), an awareness of the conditional rule for the boiler and the presence of the frost control and holiday buttons. Differences in the content of mental model descriptions were found to correspond with differences in operation of the heating simulation. Inclusion of controls in a mental model was a highly significant predictor of whether that control would form part of a behaviour strategy. The action specification with controls was also seen to vary between conditions in line with the key differences found in users' mental models, significantly improving TRV set point choices and control of boiler activation by the central thermostat in the design condition. Participants in the design condition achieved significantly greater proportion of the experimental goals than those in the realistic condition. It is concluded that a control panel interface that promotes a UMM integrating heating controls with energy monitoring and predictive technology will help users to have more control over heating consumption. This simulation focused on communicating control device models and interdependency of devices. The promotion of appropriate UMMs of home space heating within the broader context of thermodynamics, weather conditions and house structure would be the natural next step to continue this research.