



Department
of Energy &
Climate Change

Usability testing of smarter heating controls

December 2013 – Final Report

The views expressed in this report are those of the authors, not necessarily those of the Department of Energy and Climate Change (nor do they reflect Government policy).

Credits

This report was written by Dr. Steven Wall and Filip Healy of Amberlight.
The methodology for this usability study was designed by Filip Healy
The research team comprised of Steve Wall of Amberlight; Filip Healy of Amberlight; and Anthony Lau of Amberlight

Citation

Wall, S. and Healy, F. (2013). *Usability testing of smarter heating controls*. A report to the Department for Energy and Climate Change. Amberlight. DECC, London.

Acknowledgments

The authors of this report wish to thank the manufacturers of the smarter heating controls used in this study for their assistance, including the provision of devices for testing purposes, and their technical support. We would also like to thank the members of the public who participated in the usability sessions.

© Crown copyright 2013

URN: 13D/339

You may re-use this information (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence.

To view this licence, visit www.nationalarchives.gov.uk/doc/open-government-licence/ or write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or email: psi@nationalarchives.gsi.gov.uk.

Any enquiries regarding this publication should be sent to us at max.stanford@decc.gsi.gov.uk.

This document is available from our website at <https://www.gov.uk/government/publications/usability-testing-of-smarter-heating-controls>.

Usability testing of smarter heating controls

Final Report

Prepared by Amberlight

December 2013

Contents

Executive summary	5
1. Introduction	9
2. Research Design	11
2.1 Smarter heating controls	11
2.2 Sample	12
2.3 Desk research	13
2.4 Usability testing	14
3. Main findings: usability metrics and benchmarking	19
3.1 Correlation between usability measures	19
3.2 Effectiveness	19
3.3 Efficiency	20
3.4 Satisfaction	21
3.5 Usability systems scale	22
4. Main findings - usability findings	24
4.1 Complex setup of schedules	24
4.2 Lack of error prevention	24
4.3 Iconography, labelling and information architecture	25
4.4 Visual design and interaction metaphors	25
4.5 Difficulty identifying system state	25
4.6 Lack of confirmatory feedback	25
5. Main findings - emergent findings based on participant demographics	26
5.1 Age	27
5.2 Smart phone ownership	28
5.3 Education level	29
5.4 Users who have a mild dexterity or visual impairment	30
6. Installing Smarter Heating Controls	31
7. Conclusions and implications	34
7.1 Usability of smarter heating controls	34
7.2 Implications for manufacturers	35
7.3 Implications for consumers	37
7.4 Implications for DECC	38

Executive summary

Background

The Department of Energy and Climate Change (DECC) wished to gain insight into the usability of smarter heating controls to understand their suitability for future trials. DECC is interested in whether smarter heating controls have the capacity to support energy saving behaviours, but foresee a prerequisite of this is the ability of smarter heating controls to support easy use by consumers. Prior research suggested that people have difficulty using both standard and smarter heating controls.¹ DECC therefore commissioned Amberlight to measure the usability of five smarter heating controls.

Two previous pieces of work provided inputs into this research:

- Rubens, S., Knowles, J. (2013). *What people want from their heating controls: a qualitative study*. A report completed by New Experience for DECC to understand what users require from their heating controls
- An unpublished technology horizon scan of smarter heating controls and their characteristics

In particular the tasks evaluated in this research were based heavily on the outputs of Rubens, S., Knowles, J. (2013). Initial research involved an expert usability review of the controls and a manufacturers survey to gather contextual and background information on each control. A review of relevant literature in the heating control and usability fields was also conducted.

Outputs of these activities were used to inform the development of a suitable test protocol using fair and representative tasks, the sample of participants recruited to take part, and metrics of usability and benchmarks against which each control's performance could be compared.

72 participants (split into two matched samples of 36 each) attended one-to-one fieldwork sessions during which the smarter heating controls were tested for how well they supported task performance related to key user requirements. Usability metrics recorded the effectiveness, efficiency and satisfaction provided by each of the smarter heating controls. Industry standards for usability testing typically acknowledge 20 users as a standard sample size for gathering usability metrics.

In addition, observed participant behaviours and feedback were recorded using field notes and later analysed to identify any barriers to use that may account for the performance of the controls during the usability sessions.

The authors have anonymised the smarter heating controls evaluated in this investigation and all related findings, to reflect the terms of informed consent entered into by participating

¹ Rubens, S., J. Knowles (2013) *What people want from their heating controls – a qualitative study*. A report completed by New Experience for the Department of Energy and Climate Change. Available at - <https://www.gov.uk/government/publications/what-people-want-from-their-heating-controls-a-qualitative-study> (16/12/2013)

manufacturers and the researchers. This stipulated that all data, findings and reporting would be made anonymous in any published reports.

Findings: usability metrics and benchmarking

Overall performance of each smarter heating control was evaluated against a benchmark by collating the results across all participants and all tasks for each control.

Effectiveness

The effectiveness of each control was a measure of how well users were able to complete the set tasks. An overall task completion rate was calculated for each of the controls. A benchmark of 70% overall task completion rate was chosen as this would represent the majority of tasks being successful while allowing some leniency given it was the first time users had encountered the smarter heating control. None of the controls tested met this benchmark; in fact, performance for all controls was below 50%, indicating that the majority of tasks undertaken with the controls resulted in a failure to complete.

Efficiency

The efficiency of each control measured how long users spent attempting the tasks. Different benchmarks for performance were set at the individual task level in the range of 60 seconds (temporarily turning the heating on) to 240 seconds (setting up a weekly heating schedule for a 2 room house). Benchmarks were established based on the number of steps involved in the task and evaluated by a trained usability practitioner. To benchmark the performance of each control, the sum of total time spent on all tasks for each participant was compared to the sum of the individual benchmark times (900 seconds).

Three of the controls were able to meet the benchmark for efficiency, indicating that users were able to complete the tasks within a reasonable or desirable amount of time. These controls all had simpler wall mounted units and emphasised the use of web based or smart phone based interfaces for more involved tasks such as setting up the initial heating schedule, which may account for their improved efficiency.

Satisfaction

Satisfaction of each control was measured using a subjective assessment of the ease of completing each task. Participants were asked to rate the ease of completing each task on a scale of 1-5, where 1 was the most difficult and 5 was the easiest. A benchmark of 4 out of 5 was established as representing a high level of satisfaction with the control. None of the controls were able to attain the benchmark. This may be the result of participants' frustration regarding the difficulty of some of the tasks which they attempted.

System Usability Scale

Perceived usability of each of the controls was assessed by administering the System Usability Scale (SUS) questionnaire to each participant following completion of all tasks with each control. SUS is a widely recognised industry standard which uses 10 questions to provide a rating between 0 and 100 which represents the participant's subjective impression of how easy the control is to use. A benchmark score of 73 was set, as this represented the 75th percentile of scores based on meta-analysis of historical SUS data. None of the controls met this benchmark, which may be due to participants' frustrations at the difficulty they experienced completing tasks with the controls.

Findings: key usability findings

A number of barriers to use of the controls were inferred through observation of participant behaviours and their subjective comments regarding the controls. These were:

- **Complex setup of schedules.** Initial setup of the controls required multiple steps. The controls did not always clearly indicate necessary steps to participants.
- **Lack of error prevention.** Several controls allowed participants to enter non-valid times and dates when specifying heating schedules. With some controls it was possible for users to edit the schedule without being prompted to save changes.
- **Iconography, labelling and information architecture.** Icons and text labels used in the controls were frequently not clear to participants. The structure of menus and sub-menus often made it difficult for users to identify and find the correct action to take.
- **Affordances and interaction metaphors.** The visual design of some interface elements did not communicate their function to participants, or how participants were supposed to correctly interact with them.
- **Difficulty identifying system state.** Participants were often unable to identify the current state of the heating control, including disambiguating between whether the control was on, the timers were on or the heating itself was on. It was not clear with some controls how continuous heating mode interacted with timers and any thermostatic controls.
- **Lack of confirmatory feedback.** Several of the controls did not provide adequate feedback or reassurance to users that their commands had been registered or applied.

Findings: emergent findings based on participant demographics

Participant demographic data was analysed as a potential factor underlying the relative usability of the smarter heating controls. Due to the smaller sampling sizes involved when considering a sub-set of the sample, these emergent findings should be treated as illuminative case studies.

- **Age.** Higher performance was observed with younger participants compared to older participants. The youngest age group of users (18-35 years old) were able to perform significantly more tasks successfully, and with a greater efficiency, and were more satisfied. Additionally, the oldest age group (55 years old and above) were significantly less satisfied with devices that had more complex wall mounted units (Group B devices). This could be due to a higher aptitude or confidence with technology for younger participants. Visual acuity and dexterity tends to decrease with age, which may also account for the performance of the controls where older participants were considered.
- **Smart phone usage.** Participants who owned and had experience with a touch screen smart phone device generally performed substantively better with the controls than those who did not. This may be due to a higher aptitude and confidence with new technologies. Several of the wall mounted units also employed touch screen interfaces and displays that may be more familiar to users of smart phone technologies. Older smart phone users were under-represented in the sample. It was therefore difficult to conclude from this study whether smart phone usage or age has the largest impact on performance. Further research may be required to disambiguate this.

- **Education level.** Prior research that has shown a correlation between education level and cognitive and memory ability in the general population, suggests that the level of educational achievement could potentially affect performance on tasks that require interaction with an unfamiliar technology. Participants with degree level education or higher performed marginally better with the smarter heating controls than those with lower than degree level education. Users with higher education were able to complete significantly more tasks with controls that had simpler wall mounted units (Group A), and were more satisfied. No education level effect was observed with controls that had more complex wall mounted units (Group B).
- **Accessibility needs.** A small sample of participants (12) was recruited who reported a minor visual or dexterity impairment. Results showed that the usability of the smarter heating controls was substantively lower for participants with a minor visual or dexterity impairment. No statistical analysis was conducted for this effect due to the small sample size involved. The difference in performance may be due to the demands that some tasks placed on participants in terms of perceiving information on small digital displays or fine manipulation of user interface elements.

Findings: Installation requirements

A survey of participating manufacturers, undertaken to provide supplementary context relating to smarter heating control installation found that the smarter heating controls tested were not 'box ready' but required installation by suitably qualified technicians (e.g. heating engineers or electricians). This represents a potential additional cost to consumers wishing to purchase a smarter heating control, although it may also be that all costs are met if systems are installed by a landlord, utility or in new build properties. For the most part installation should be quick and fairly simple, although this is relative to the scope and the complexity of the system the consumer wishes to install.

Conclusions and implications

Results indicate that the smarter heating controls tested in this study were not able to support participants in completing the tasks set for them. A number of usability issues were identified through analysis of supporting data from the user sessions (observed participant behaviours and spontaneous comments by participants). The usability issues were inferred to have contributed to the overall difficulty that participants experienced trying to complete tasks with the controls.

These findings have implications for the potential users and manufacturers of smarter heating controls as well as for DECC. Difficulty in using current space heating controls may present a barrier to people who wish to save energy. The perceived difficulty in using smarter heating controls may also pose a substantive motivational barrier to the uptake of energy saving behaviours, even if controls are installed in a person's home. There is a need for innovation around the usability of controls in order to address the current barriers to use. A number of recommendations are provided to highlight opportunities for manufacturers and DECC to improve the usability of smarter heating controls.

1. Introduction

Domestic energy use accounts for 29% of UK energy consumption with space heating responsible for 66% of domestic energy consumption and water heating for an additional 17%.² Carbon budgets commit the UK to reduce overall CO₂ emissions by at least 80% by 2050.

Improving heating controls may be one way to improve the energy efficiency of domestic properties. Industry estimates that installing a standard set of three heating controls (timer, room thermostat and thermostatic radiator valves) in all homes could reduce the energy they use for heating and hot water.³ There are also claims that emerging heating control technologies could deliver additional savings. However, there is currently no robust evidence that standard or advanced heating controls actually save energy.

Recent studies have found no significant difference in internal temperatures for homes controlled manually compared to those controlled with a timer or programmable thermostat.⁴ The Department of Energy and Climate Change (DECC) wished to test the usability of smarter heating controls under the assumption that the capacity of heating controls to bring about savings rests upon the consumers ability to use them effectively .

A relatively recent literature review examined consumer use and understanding of their heating controls. It found that consumers find heating controls difficult to use and many do not use them effectively. It highlighted that older people and those in local authority housing were more likely to find their controls difficult to use.⁵

Building on earlier insights regarding consumer requirements for smarter heating controls,⁶ DECC commissioned Amberlight to conduct research to evaluate how well currently available smarter heating controls supported users towards the goal of achieving more energy efficient behaviours in the home.

The objective of this research was to ensure that difficulty of use could be eliminated as a potential reason for a failure to engage in energy efficient behaviours in any subsequent field trials of smarter heating controls. Specifically, the goals of this research were defined as follows:

² DECC (2013) Energy Consumption in the UK. Chapter 3 – Domestic sector energy consumption. Available at - https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/65954/chapter_3_domestic_factsheet.pdf (16/12/2013).

³ TACMA (2010) Heating and hot water pathways to 2020. Available at - http://www.beama.org.uk/en/news/index.cfm/hhwt_pathways_2020_report (16/12/2013).

⁴ Kelly, S., Shipworth, M., Shipworth, D., Gentry, M., Wright, A., Pollitt, M., Crawford-Brown, D. and Lomas, K. (2013) Predicting the diversity of internal temperatures from the English residential sector using panel methods, *Applied Energy*, 102, 601-621.

⁵ Consumer Focus (2012). Consumers and domestic heating controls: a literature review. London: Consumer Focus. Available at - <http://www.consumerfocus.org.uk/files/2012/01/Consumers-and-domestic-heating-controls-a-literature-review.pdf> (16/12/2013).

⁶ Rubens, S., J. Knowles (2013) *What people want from their heating controls: a qualitative study*. A report completed by New Experience for the Department of Energy and Climate Change. Available at - <https://www.gov.uk/government/publications/what-people-want-from-their-heating-controls-a-qualitative-study> (16/12/2013).

1. Support the goal to encourage energy efficient behaviours through good usability of heating controls
2. Objectively measure and compare a number of heating controls in terms of usability
3. Help identify smarter heating controls that are unsuitable for subsequent field trials

In order to address these goals, Amberlight planned, conducted and analysed a summative usability evaluation of the performance of 5 smarter heating controls. The usability sessions were designed to assess how well a range of smarter heating controls supported user requirements related to energy saving behaviours in a home environment.

Fieldwork for the research was conducted during July 2013. Section 2 of this report describes the methodologies used during the fieldwork and the rationale for selection of these methods. Section 3 describes the results of the fieldwork. Section 4 discusses the results in context of observed user behaviours during the fieldwork to identify potential barriers to use with the smarter heating controls that were tested. Section 5 highlights emergent results based on deeper analysis of the performance of heating controls based on specific demographics of the users who participated in the study. Section 6 discusses the conclusions and implications of the findings of this research. Detailed material (covering research materials, more detailed findings) is available in a separate volume of appendices, Wall, S. and Healy, F. (2013). *Usability testing of smarter heating controls - Appendices*. In addition, the raw user scores data is available in a separate volume of statistical tables, Wall, S. and Healy, F. (2013). *Usability testing of smarter heating controls – user scores data*.

2. Research design

2.1 Smarter heating controls

The definition of a smarter heating control flows directly from the user requirements gathered from Rubens, S., Knowles, J. (2013). *What people want from their heating controls: a qualitative study* a report commissioned by DECC's Smarter Heating Controls Research Programme. Smarter heating controls were seen as having the following characteristics:

- **Zonal:** independent monitoring and control of heating in separate rooms (or groups of rooms) in the house
- **Remote:** monitoring and control of the heating system via web and smart phone based channels from locations outside the house
- **Automation:** technologies that leverage sensors and other sources of data to provide energy efficient heating without the need for user intervention

From this DECC is currently evolving a working definition of smarter heating controls as a group of technologies that, installed into a domestic heating system, seek to facilitate the centralised control of heating. Active Smarter Heating Controls typically provide consumers with greater control by enabling the local and remote control and management of temperature, time and zone. Passive systems are characterised by technologies that seek to control heating by learning about a consumer's heating behaviour or by delivering the same or similar amount of comfort by use of less energy.

Five smarter heating controls were selected for evaluation in this research. Heating controls were selected based on zonal and remote capabilities. Based on a prior review of the marketplace, controls with automation were not fully ready and available in the UK for testing. A detailed scan of the marketplace that involved engagement with an array of manufacturers was conducted by DECC and Amberlight prior to the usability testing. The controls selected for participation were those that both met the functional requirements of customers that were outlined in Rubens, S., Knowles, J. (2013) and the availability of sample products for the study. Serious consideration was given to devices that were unavailable for this testing, however it was concluded that the dynamic nature of the smarter heating controls market meant that if the study were postponed to accommodate emerging devices, a further handful of devices might then reasonably be expected on the horizon. However, it has been ensured that the study design facilitates easy and accurate replication. Therefore, should a clear need to test further devices arise, this could be achieved 'measure' for 'measure'.

Further details of the smarter heating controls that were tested are not provided in this report to ensure the anonymity of the devices and manufacturers. The intent of this research is not to endorse or validate a particular manufacturer's approach as "better" in terms of usability, but to understand the current capability of the marketplace to support desired energy saving behaviours from the standpoint of "ease of use" for the consumer.

2.2 Sample

Based on prior research, a number of factors were identified as potential influences on the ease of using smarter heating controls. These factors were actively screened for, ensuring a diverse and purposive set of participants. This included:

1. **Age** – A recent literature review identified older users as a subgroup of users that are likely to find heating controls more difficult to use.⁷
2. **Education Level** – User interfaces that place a high demand on cognitive and memory ability of individuals can cause a barrier to use. Education level of individuals has been shown to correlate with cognitive and memory ability.⁸
3. **Household income** - Previous research has suggested people with the lowest income cut their energy usage in response to energy price inflation, whereas wealthiest homes simply spend more on energy.⁹
4. **Experience with smart phone technology** – All of the controls that were evaluated used a smart phone application as a means of controlling domestic heating; therefore it was important that the sample included and controlled for participants who had experience of using these technologies.
5. **Accessibility needs** – A small sample of participants with minor visual or dexterity impairments was included as these users may experience difficulty reading displays or interacting with physical controls.
6. **Attitude to energy saving behaviours** – Participants who did not report a motivation to manage and save domestic energy consumption were excluded from the sample.
7. **Prior experience with heating controls** – Participants with prior experience of smarter heating controls may introduce considerable bias and were therefore excluded from the sample. For similar reasons, heating engineers and plumbers were also excluded.

A purposive sample (sometimes referred to as a judgmental sample) is one where the participants are selected because of specific characteristics that are relevant to the purpose of the study. In this case our purpose was to ensure that the controls are usable, even for the less abled members of the population. We deliberately included these users (who were over-represented compared to a nationally representative sample), knowing that if they could carry out the key tasks, then more proficient and able users should be able to do so even more easily.

Two groups of 36 participants (giving a total sample of 72 participants) were recruited to participate in the fieldwork. Quotas for the above 7 factors were matched between the two groups of participants. The sample was split into two groups to maximise coverage of different

⁷ Consumer Focus (2012). Consumers and domestic heating controls: a literature review. London: Consumer Focus. Available at - <http://www.consumerfocus.org.uk/files/2012/01/Consumers-and-domestic-heating-controls-a-literature-review.pdf> (16/12/2013).

⁸ T. Falch, Sandgren, S. (2006), The effect of education on cognitive ability, Norwegian University of Science and Technology. *Economic Inquiry* Vol. 49, No. 3, July 2011, 838–85. Available at - <http://www.sv.ntnu.no/iso/Torberg.Falch/Articles/Article.pdf> (16/12/2013).

⁹ Consumer Focus (2011). Fuel price inflation and low income consumers. London: Consumer Focus. Available at - <http://bit.ly/PLS4QJ> (16/12/2013).

smarter heating controls and tasks, while ensuring validity of any comparisons between the relative performances of the different controls. For usability studies of this nature, the industry standard for an appropriate sample size is widely considered to be 20 participants. This estimate is based on the variances observed in usability metrics across a broad range of studies.¹⁰

A detailed breakdown of the recruited sample appears in Appendix A of the accompanying appendices, Wall, S. and Healy, F. (2013). *Usability testing of smarter heating controls - Appendices*. In addition, the raw user demographic scores data and graphs by task, usability metrics and device are available in separate statistical tables, Wall, S. and Healy, F. (2013). *Usability testing of smarter heating controls – user scores data*.

Participants were sourced through a third-party organisation that recruits participants for market research and usability trials.

2.3 Desk Research

The following primary and secondary research activities were conducted prior to the fieldwork. This was to ensure that the methodology used took into account existing user requirements and the capabilities of the controls being assessed in the research.

2.3.1 Literature review

Prior outputs from DECC's Smarter Heating Control Research Program were reviewed to ensure that the fieldwork was representative of end user requirements and behaviours. The outputs of DECC's commissioned report, Rubens, S., Knowles, J. (2013), were reviewed to identify the most salient consumer needs for heating controls. These user needs were then used to define representative tasks for users to attempt with smarter heating controls, in order to assess how well the controls supported the underlying user need. The full list of user needs and tasks is discussed in section 2.4.2.

A wider literature review of relevant evidence was also conducted to inform development of a robust series of usability metrics to evaluate the performance of the controls.

2.3.2 Expert Review

The smarter heating controls selected for participation in the fieldwork were firstly independently evaluated by two Amberlight usability specialists. Evaluation of the five controls was performed using a heuristic evaluation checklist which rated the control's capability to support the identified user goals on a three point scale (does not support, partially supports, full support).

Based on the outputs of the expert review, one of the participating controls was identified as having wider capability to support the user goals than the other four controls. This included both the number of user goals that the control provided full support for, and the breadth of support across different platforms (web, smart phone, and wall mounted unit). This control was therefore tested with both groups of participants to ensure maximum coverage of the control's capabilities. The first group focussed on the web interface as the primary channel for completing tasks, while the second group mainly used the wall mounted unit.

¹⁰ Travis, D., Two kinds of usability testing. Available at - <http://www.userfocus.co.uk/articles/2-kinds-of-usability-test.html> (28/11/2013); Nielsen, J., 2006, Quantitative Studies: How Many Users to Test? Available at - <http://www.nngroup.com/articles/quantitative-studies-how-many-users> (16/12/2013).

2.3.3 Manufacturer Survey

Input was also solicited from manufacturers of the smarter heating controls that participated in the research.

Firstly, the manufacturers were asked to validate the correct sequence of instructions that participants should follow to complete the specified tasks.

Secondly, the manufacturers were asked to complete a survey to provide more background information and context on their control. This included information on any necessary installation process required, the platforms that the control supported (e.g. smart phone, tablet, web) and any minimum technical specifications necessary to use the control.

All the controls that participated in the research required a home with an existing broadband internet connection and a modern PC running a web-browser. In addition, several of the controls benefit substantially from smart phone technology in terms of access to available functionality. Several controls are targeted specifically at owners of smart phone and tablet devices. All manufacturers recommend a trained heating engineer to perform installation of the controls. An anonymised version of the full survey is available in the appendices for this report, Wall, S. and Healy, F. (2013). *Usability testing of smarter heating controls - Appendices*.

2.4 Usability Testing

The five participating smarter heating controls were evaluated for ease of use with a summative usability testing methodology. Each participant attended a 90 minute session in a controlled usability lab environment with an Amberlight usability consultant acting as the moderator at Amberlight's central London offices. During the session, the participant was required to attempt up to 8 compelled tasks with 3 smarter heating controls. The participant attempted all tasks with one control before moving on to the next one. The moderator introduced each control and task in a consistent and unbiased way, recorded performance metrics for each of the tasks and administered post-hoc questionnaires to gather the participant's subjective opinion of the controls.

2.4.1 Testing environment

Sessions took place at Amberlight's central London facilities. Two usability labs were setup with identical equipment specification and configuration to allow testing to take place concurrently in both labs while minimising any confounding factors or biases due to the testing environment. Testing utilised desktop PC (web browser), smart phone, tablet and wall-mounted unit platforms.

The instructions for each specific task indicated which platform the participant was to use to attempt the task. All necessary setup (for example, entering the website URL, installing the smart phone app, entering login credentials) was performed prior to the session by the moderator. The participant was instructed to start each task from the home page or start screen of the website, app or wall mounted unit display (where applicable).

Wall mounted units were mounted 140cm off the floor to ensure they were in a realistic and comfortable position for participants to use. Only artificial lighting was used in the labs so that variations in natural lighting conditions did not affect the readability of any displays.

2.4.2 Tasks

Tasks for the usability sessions were defined based on user requirements identified in prior research commissioned by DECC. Research identified the following high priority requirements for users of smarter heating controls:

Requirement	
1	See how much heating is costing you
2	Ability to set different temperatures at different times for different rooms – all from one central panel
3	Understand savings that can be made by making adjustments to settings on the heating controls
4	Ability to remotely turn on heating before returning home
5	Clear and permanent display of times and temperatures set
6	Simple way to switch heating on/ off
7	Rapid warming when turning heating on

A set of tasks were defined to address the high priority requirements. Requirement 3 was not considered as none of the controls supported real-time feedback on costs and savings incurred as a result of using different settings. Requirement 7 was not considered as the requirement for rapid warming did not pertain to the usability of the controls. The tasks were as described below:

Task	Corresponding requirements
1a Setup the weekly heating schedule for a two room house (bedroom and living room)	2,5
1b Edit the heating schedule for the bedroom to come on earlier one day per week	2,5
2 Edit the heating schedule remotely (using mobile or desktop) to come on earlier one day per week	2,4,5
3 Temporarily switch the heating on when returning home	4, 6
4 Temporarily stop the heating schedule for 1 week while on holiday, ensuring the system is protected in the event of very cold weather	6
5 Temporarily switch the heating off, without affecting the schedule	6
6 Turn the heating on remotely (using mobile or desktop)	4, 6
7 Find information about your energy usage	1

Order of task presentation was the same for each of the smart heating controls. The tasks were ordered so as to present a logical journey with each control from initial setup of the schedules through to checking energy usage after the control had been in use for a certain period (using simulated data provided by the manufacturers to represent 1 months use).

Each task was presented to the user in the form of a scenario to provide additional context and motivation for the task. The verbatim descriptions that were given to users are included in the appendices for this report.

For each control, the most appropriate platform (web, smart phone, tablet, and wall mounted unit) was selected for each task, based on the inferred context of use (whether the user would normally be attempting the task inside or outside the home environment) and the platform that provided the best support for the task (based on the expert review and input from manufacturers).

2.4.3 Usability Metrics

The international standard for guidance on usability (ISO 9241-11) defines usability as “the extent to which a specified product can be used to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use”.

In the specific case of this research, the aim of the usability testing is to assess whether the different smarter heating controls support the participants in successfully completing tasks outlined in Section 2.4.2 with effectiveness, efficiency and satisfaction.

Three metrics were therefore selected to evaluate the relative usability of each control for each task:

- **Effectiveness** – *Did the control support the participant to achieve the specified goals?*

The moderator assigned a score between 0.0 and 1.0 depending on the user’s performance. A success (1.0) was awarded if the user completed all the necessary steps of the task. A partial success (0.25) was awarded if the user had completed the majority of the steps involved in the task, but not all. A failure (0.0) was awarded if the user had completed less than half of the steps.

- **Efficiency** – *Did the control support the participant in completing the task within a reasonable amount of time?*

The moderator recorded the amount of time in seconds that the user spent attempting the task.

- **Satisfaction** – *How much did the participant agree that the control supported them in completing the task?*

Following each task the user was asked to give a subjective rating from 1-5 to indicate how much they agreed with the statement that “the task was easy to complete”, where 1 was strongly disagree and 5 was strongly agree.

In addition, following completion of all tasks with a particular control, the participant was asked to complete a System Usability Scale (SUS) questionnaire to measure their perception of how usable the control was overall (SUS is widely recognised as an industry standard with

references in over 600 publications). SUS provides a score between 0 and 100 for each system where a higher number indicates a better perception of system usability.¹¹ The complete SUS questionnaire is included in the appendices of this report.

2.4.4 Analysis

Meier et al.'s "M-statistic" method was proposed as a single, consolidated metric that combines measures of effectiveness, efficiency and satisfaction in to a normalised score between 0 and 1.¹² The M-statistic was developed as a means of quick and easy comparison between heating control devices.

However, analysis of historical data from usability studies suggests that measures of effectiveness, efficiency and satisfaction may not correlate (for example, see Hornbaek and Law¹³). If the measures do not correlate, a single consolidated measure of usability such as the M-statistic could potentially result in a loss of granularity in the data.

As an alternative, Amberlight also assessed measures of effectiveness, efficiency and satisfaction as independent metrics for purposes of analysis. This is recommended by ISO 9241-11, the international standard for usability, and represents best practice in the field.

2.4.5 Benchmarks

To assess the relative usability of each smarter heating control, benchmarks were established for each of the usability metrics. Controls that exceeded the benchmarks would therefore be considered to provide adequate support for users to achieve the specified goal of energy saving behaviours within the home. Controls that failed to meet the benchmarks could be inferred to have usability issues that may impact a user's ability or motivation to engage in energy saving behaviours.

The following benchmarks were assigned for each metric:

Usability Metric	Benchmark
Effectiveness	Overall task completion rate of 70%. This was selected to ensure that the control was able to support the completion of the majority of tasks. This was deemed to make reasonable allowances for performance given that participants were using the controls for the first time.
Efficiency	Benchmarks were set for each task based on the perceived complexity of the task as estimated during the expert review phase. Task benchmarks for efficiency ranged from 60 seconds (Task 3,5 and 6 – turning the heating on or off) to 240 seconds (Task 1a – setup the weekly heating schedule)

¹¹ Brooke, J. "SUS – a quick and dirty usability scale", Redhatch Consulting Ltd. Available at - <http://hell.meiert.org/core/pdf/sus.pdf> (16/12/2013).

¹² Meier, A., Aragon, C., Pefer, T., Perry, D., Pritoni, M., 2011, "Usability of residential thermostats: preliminary investigations", *Building and Environment*, 46, 1891-1898.

¹³ Hornbaek, K., Law, E.L-C., 2007, "Meta-analysis of correlations amongst usability measures", In *Proceedings of SIGCHI Conference on Human Factors in Computing Systems*, pp. 617-626

Usability Metric	Benchmark
Satisfaction	Overall satisfaction rating of 4 out of 5. This was selected to ensure an overall high level of satisfaction with the controls.
System Usability Score	SUS score of 73, which corresponds to the 75 th percentile of performance based on meta-analysis of historical SUS data from previous studies. ¹⁴

2.4.6 Smarter heating controls

Each group of 36 participants attempted tasks with 3 different heating controls. The controls were assigned to one of two groups based on outputs from the expert review and manufacturer feedback. Controls that primarily supported tasks through web and smart phone based platforms, and had correspondingly simpler wall mounted units, were assigned to the first group (Group A). Controls with more complex wall mounted units that supported completion of more of the tasks were assigned to Group B.

2.4.7 Usability session structure

The following steps were taken to ensure that no bias was introduced into the results due to the format of the session:

- The order of presentation of the three controls was counterbalanced within each group to control for order effects.
- Printed instructions were provided for each task. The participant was instructed to read aloud these instructions prior to beginning each task. This was to control for any biases in the way the moderator introduced the task.
- The participant was given the opportunity to clarify the definition of each task prior to attempting it. However, the moderator did not provide any assistance to the participant or answer any questions once the task had begun, with the exception of the participant causing a potential critical failure of the control (for example, selecting a “factory reset” option).
- No feedback was provided to participants at any time regarding how successful they had been at any of the tasks.

It should be noted that in this study participants attempted the tasks on each smarter heating control with no specific training or guidance. For more information on this and a description of the task users undertook, see Appendix C of Wall, S. and Healy, F. (2013). *Usability testing of smarter heating controls - Appendices*.

A user manual for each device was available for reference throughout the session, but there was no time specifically allocated for reading and reviewing the manual before attempting the tasks. In real world settings, where devices are installed into consumer’s homes, installers would be expected to provide some additional support following installation such as demonstrating the key features of the heating control.

¹⁴ Sauro, J. “Measuring Usability with the System Usability Scale”, Available at - <http://www.measuringusability.com/sus.php>. (28/11/2013).

3. Main findings – usability metrics and benchmarking

This section of the report considers the performance of the smarter heating controls based on the usability metrics outlined in Section 2 (see 2.4.3 Usability Metrics).

Recordings for the three usability metrics outlined in Section 2 were aggregated for each control by considering performance of all participants across all tasks. As highlighted in section 2.3.2 Expert Review, one control was identified as having wider capabilities to support user goals than the other four controls and was therefore tested with both groups of participants to ensure maximum coverage of the control's capabilities. This resulted in six systems being tested. For each system, a mean score for each of the usability metrics was calculated using this aggregated data.

Where percentages are quoted these refer to the sample of participants that took part in the usability sessions rather than the wider population of users of smarter heating controls.

We first considered the correlation between the different measures for usability, as this dictated whether Meier et al.'s M-statistic or the method of considering independent measures would be used for further analysis.

The raw user score data by task and device testing the three usability metrics can be found in statistical tables, Wall, S. and Healy, F. (2013). *Usability testing of smarter heating controls – user scores data*, which accompanies this report.

3.1 Correlation between usability measures

For each smarter heating control the Pearson co-efficient for correlation was calculated between each of the measures of usability. For all controls there was a low correlation observed between effectiveness and efficiency ($r < 0.3$).

Incomplete tasks tended to have taken longer but there was no overall clear trend found in the data. Partially completed tasks tended to show one of two main trends: in some cases partial completes took as long as incomplete tasks, but in other cases they were as quick as successful instances of the task. These two trends contributed to the lack of a correlation between effectiveness and efficiency measures.

There was a consistent correlation between the SUS score and the satisfaction metric for each control, where the r co-efficient was never less than 0.67. It would therefore be reasonable to only consider one of these two measures in any future analysis conducted with this data. The consistent lack of correlation between effectiveness and efficiency measures favoured independent analysis of each of the metrics instead of combining them into a single M-statistic.

3.2 Effectiveness

Overall, none of the controls met the benchmark of 70% task completion rate when performance of all participants on all tasks was considered. The overall effectiveness of each control is shown in Figure 1. Performance ranged from 23% task completion rate for System D through to 48% task completion rate for System F.

These results suggest that all the smarter heating controls may have substantial barriers to use for participants trying to achieve energy saving behaviours.

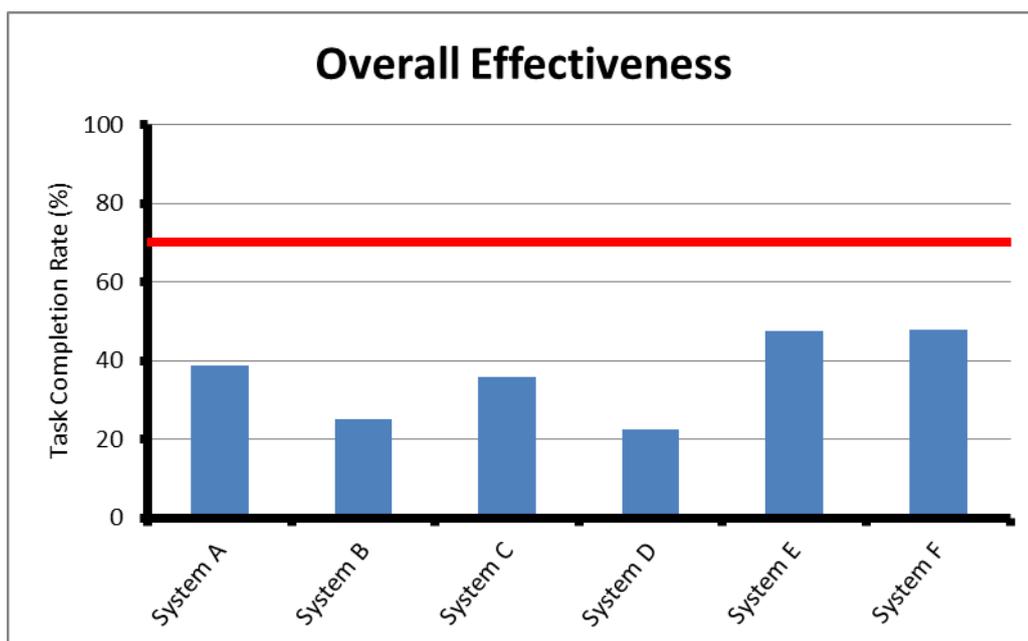


Figure 1: Overall effectiveness of each control as measured by task completion rate when considering performance of all participants over all tasks, compared to a benchmark of 70% task completion rate.

3.3 Efficiency

The overall efficiency of each control was evaluated by summing the total time on task across all tasks for each participant. It is important to note that only partially successful or successful tasks were considered in this analysis. Time taken by unsuccessful participants may not be indicative of the actual time required to complete tasks if, for example, participants abandoned tasks early. The total time on task was averaged across all participants to generate a mean total time on task for each control. This value was compared to the sum of the benchmark times for each individual task (900 seconds). The performance of each of the controls for overall efficiency is shown in Figure 2.

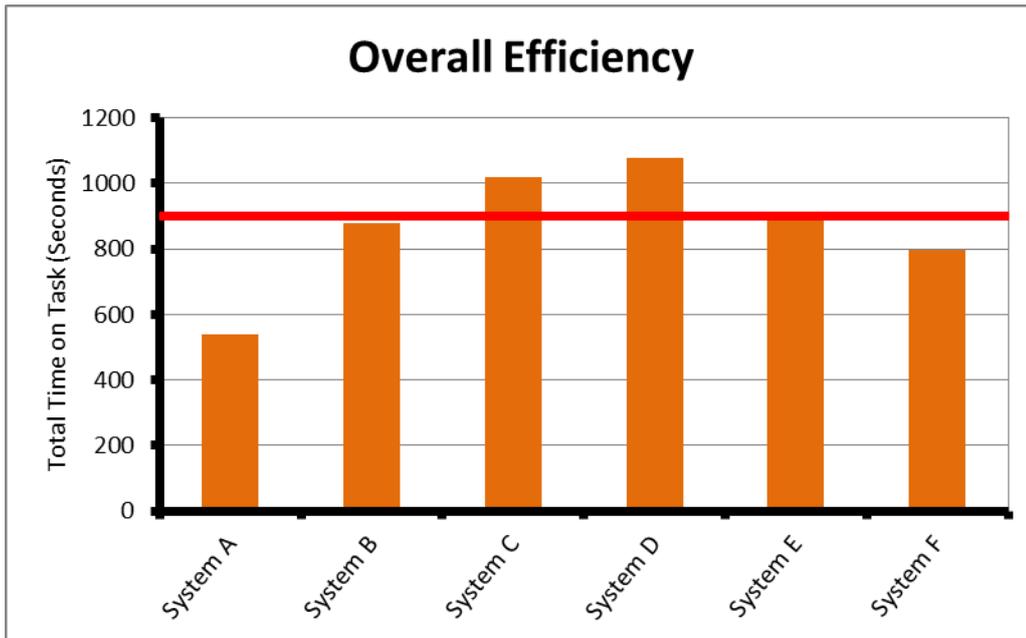


Figure 2: Overall efficiency of each control as measured by total time to complete all tasks compared to a benchmark of 900 seconds.

In order for a control to meet the required benchmark for efficiency the average total time on tasks should be below the benchmark of 900 seconds.

The total time on task for all the controls ranged from 540 seconds for System A through to 1077 seconds for System D. Three controls were able to meet the benchmark for efficiency: System A, System B and System F. The remaining three controls had average total times that exceeded the benchmark of 900 seconds.

The results suggest that the three controls meeting the benchmark (A, B and F) may support users in completing tasks within a reasonable and desirable period of time.

The three controls that met the benchmark all had relatively simple wall mounted units that emphasised the use of a small number of physical controls (e.g. push buttons) over a dynamic display screen. More complex tasks such as setting up a weekly heating schedule were handled through a web portal or smart phone interface. This difference in design may account for the more efficient performance of these particular devices.

3.4 Satisfaction

Overall, none of the controls met the benchmark for an average score of 4 out of 5 based on aggregated scores for all tasks and all participants. The performance of each of the controls for overall satisfaction is shown in Figure 3. Satisfaction scores ranged from 2.20 for System D through to 3.31 for System A.

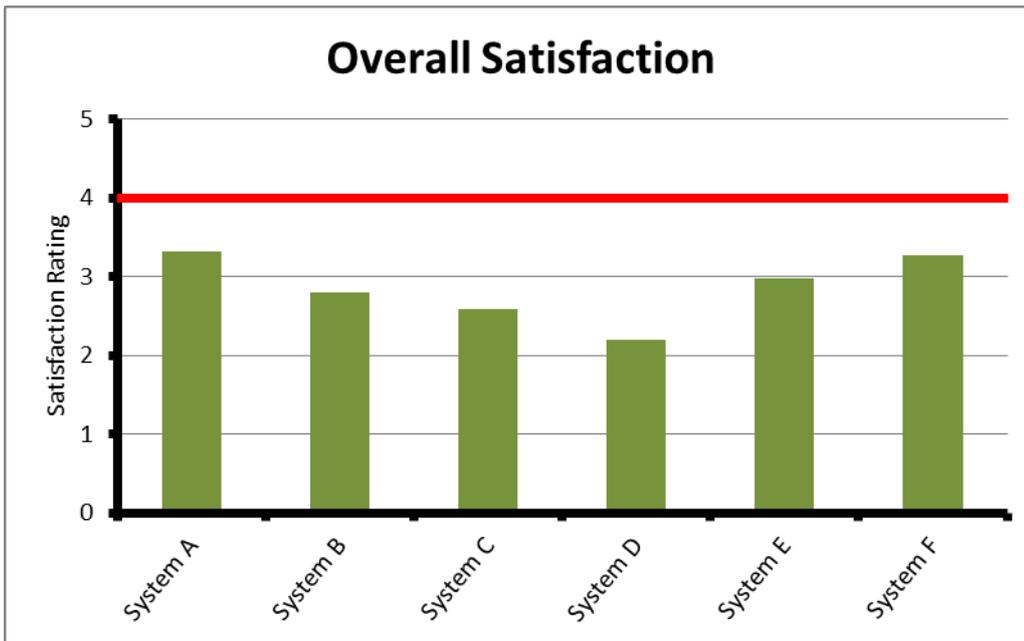


Figure 3: Overall satisfaction for each control as measured by participant’s subjective ratings from 1 to 5 for perceived ease of completing each task. A higher overall score indicates a higher overall satisfaction with the control. Scores are compared to a benchmark of 4 out of 5.

This result suggests that participants may have been dissatisfied with the controls due to the difficulty they experienced in completing the majority of tasks.

3.5 System Usability Scale

Individual SUS scores for each control were averaged across all participants to calculate a mean score. The mean score for each control was then compared to the benchmark score of 73. Overall, none of the controls met the benchmark set for the SUS score. Average SUS scores ranged from 31.1 for System D through to 56.6 for System A. The results are shown in Figure 4.

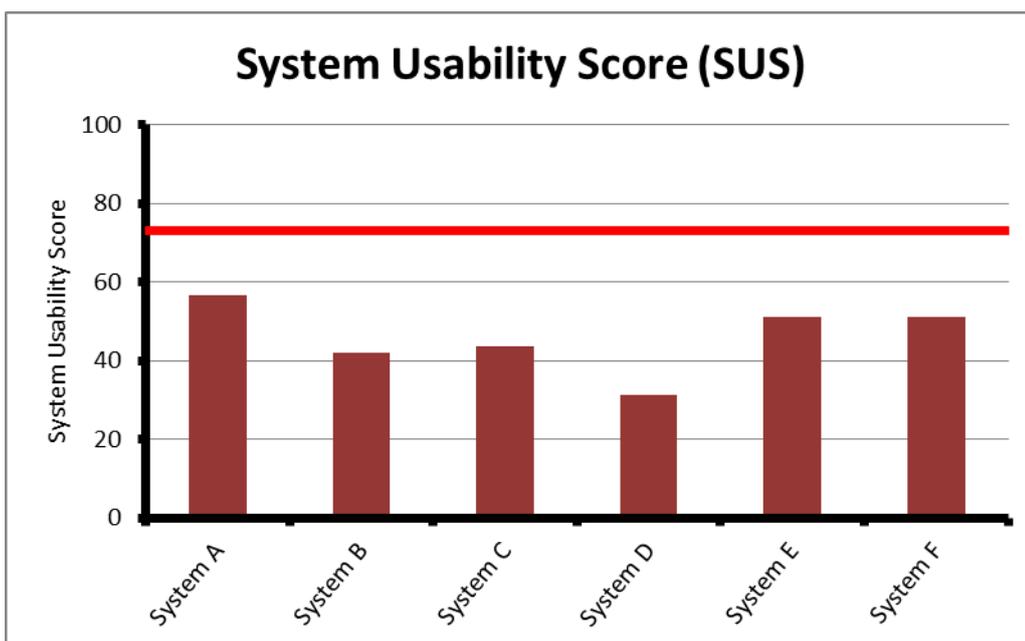


Figure 4: Overall perception of usability for each control as measured by average SUS score across all participants. A higher overall score indicates a higher overall perception of control usability. Scores are compared to a benchmark of 73.

The results suggest that participants had a negative perception of the relative usability of the controls that were tested. This may be due to the difficulties that participants experienced completing the majority of tasks.

4. Main findings – usability findings

The usability metrics showed that controls largely failed to meet the benchmarks established for effectiveness, efficiency and satisfaction. This section considers the potential usability issues of smarter heating controls that may account for the performance of these controls. Usability issues are conceptual or physical barriers to use that may prevent users from achieving specific goals with a particular control.

Potential usability issues underlying the performance metrics were inferred through considering a range of evidence:

- Observations made by the moderator during the usability sessions
- Spontaneous comments made by participants during the usability sessions
- Rationale provided by the participants for their subjective ratings of controls
- Findings of the expert review conducted by a usability expert prior to the testing

These sources of information were collated as field notes during the user sessions and cross-referenced during a period of analysis in order to arrive at an understanding of the possible reasons for the relative performance of the controls.

4.1 Complex setup of schedules

Setting up an initial heating schedule using the smarter heating controls involved multiple steps. These typically included configuring the timing for at least one room or day, ensuring these settings were copied across to multiple rooms in the house over multiple days, and specifying a target temperature for the room. Some controls involved additional steps, such as configuring rooms of the house into zones.

Several of the controls did not make participants aware of all of the necessary steps to the setup process. This often led to a situation where participants believed they had successfully configured the heating system but had perhaps omitted one or two essential steps, earning the control at best a partial success for that task. For System E no single participant discovered all the necessary steps required to setup a heating schedule successfully.

4.2 Lack of error prevention

Several of the controls permitted the participant to enter incorrect details in input fields. This included, for example, entering invalid alpha-numeric or punctuation characters when typing in a time or date. In some cases the control did not make the participant aware of this error, leading to incorrect scheduling.

Another common usability issue encountered with several of the controls was allowing the participant to edit a schedule without prompting them to save the changes. In many cases participants did not realise that the changes they had made to schedule had not been saved, and believed they had setup the schedule correctly.

4.3 Iconography, labelling and information architecture

Participants often found it difficult to locate or identify necessary actions while using the smarter heating controls due to the language and icons used to represent them, and how the control organised the functions and features into sub-menus (the information architecture).

Ambiguous or generic labels or headings for menus such as “settings” or “details” caused a great deal of confusion for participants trying to locate specific features of the control.

4.4 Visual design and interaction metaphors

The visual design and layout of some user interfaces did not implicitly communicate to participants how they were supposed to use them. This occurred with both digital user interfaces and the physical user interfaces on the wall-mounted units. This included, for example, the visual design of buttons on a smart phone interface not clearly communicating to users that the buttons were part of the user interface and could be tapped on.

4.5 Difficulty identifying system state

Identifying and modifying the state of the heating was difficult with some controls. Several different paradigms were employed between the different controls for how temporary changes to the heating would interact with any saved schedules, and in many cases the controls did not communicate this to the participant. For example, some controls had a setting that would override the schedule to turn the heating on regardless of the thermostatic temperature. Whereas other controls required the participant to also modify the target temperature of a thermostat but did not communicate that this step was necessary. Displays of the system state were frequently ambiguous as to whether the control system, the heating or the timers were “on”, leading to difficulty for users identifying whether they had succeeded at a task or not.

4.6 Lack of confirmatory feedback

Several of the controls tested did not provide adequate reassurance or closure for participants upon successfully completing a task. This left many participants unsure of whether they had actually completed a task, which may have contributed to increased time on a task and lower satisfaction ratings. In some cases participants would repeat all the necessary steps of the task to verify the control had acknowledged the commands.

This is particularly salient for platforms where the user can control the heating system remotely, such as via a web portal or smart phone, where implicit environmental feedback would not be available, such as the sound of the boiler turning on or off, or a perceived change of radiator temperature.

5. Main findings – emergent findings based on participant demographics

This section considers the results of the usability testing from the perspective of different demographic factors that were used in the sample. Prior research has indicated that participants from older age groups or who live in social housing may experience more difficulty using standard heating control systems.¹⁵

The results discussed in this section should be considered as illuminative case-studies in that they may suggest particular differences between demographic groups that manufacturers should be aware of when considering future innovations in smarter heating controls, or that DECC should be aware of when planning future research or policy decisions. Therefore, results should be treated as preliminary findings due to the smaller sample sizes used for each of the individual demographic groups.

Several of the demographic factors that were actively screened for in the sample (as described in Section 2) were considered as an independent variable for purposes of this analysis. The four measures of usability were evaluated over all tasks with all devices in order to understand how easy the smarter heating controls were to use for each of the demographic groups.

Statistical analysis for the effects of age and education level was conducted using an ANOVA. Analysis was performed separately for controls in Group A and Group B. Statistically significant results are reported in the relevant section below.

No statistical analysis was performed for the effect of smart phone ownership as there was a substantive skew in the participants recruited for the testing. Smart phone owners tended to be younger in both groups of participants and older smart phone users were under-represented. Given the strong effect of age on performance of the controls, it would therefore be difficult to disambiguate whether any effect of smart phone ownership was due to the younger demographic of smart phone owners who participated in the study.

No analysis was performed for users with accessibility needs due to the small sample recruited (6 participants in each group) which would preclude meaningful analysis of the results due to the variances that are typically observed in usability studies of this kind.

Full results of the statistical analysis performed for the effects of age and education level are included in the appendices of this report. In addition, the raw user demographic scores data and graphs by task, usability metrics and device can be found in statistical tables, Wall, S. and Healy, F. (2013). *Usability testing of smarter heating controls – user scores data*, which accompanies this report.

¹⁵ Consumer Focus (2012). Consumers and domestic heating controls: a literature review. London: Consumer Focus. Available at - <http://www.consumerfocus.org.uk/files/2012/01/Consumers-and-domestic-heating-controls-a-literature-review.pdf> (16/12/2013).

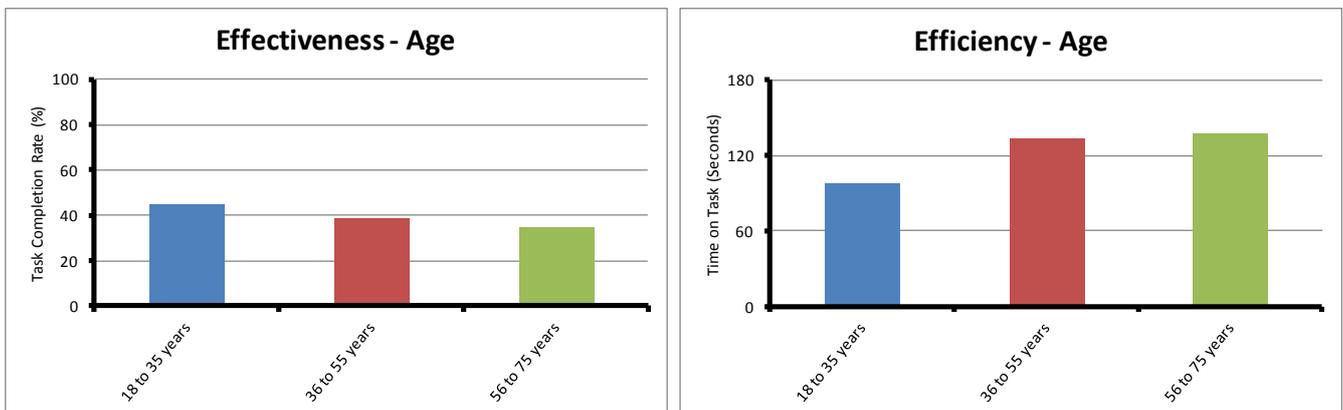
5.1 Age

Participants were split into three age groups, 18-35 years, 36 – 55 years and 56 – 75 years. The results suggest that users from older age groups may experience more barriers to use when attempting to employ smarter heating controls compared to younger users. Performance of the smarter heating controls was seen to decrease with increasing age group across all four of the usability metrics. Figure 5 shows the performance of the smarter heating controls on each of the usability metrics.

For controls in group A, participants in the age range 18-35 years were able to complete significantly more tasks ($P < 0.01$), were significantly quicker at doing so ($P < 0.01$) and scored the tasks with a significantly higher satisfaction ($P < 0.01$) than the other age groups.

For controls in group B, again the youngest age groups of participants were able to complete the tasks with significantly higher completion rates ($P < 0.001$) and significantly lower completion time ($P < 0.001$). In addition the oldest age range of participants 56-75 years rated the devices in Group B as significantly lower satisfaction ($P < 0.001$).

Older users may have experienced more difficulty with the controls due to lower visual acuity or manual dexterity compared to the youngest age group. Younger users also tended to have a greater experience or aptitude for digital technology and connected devices, which may have influenced their confidence and performance with the smarter heating controls. However, the older two age groups considered in the analysis are more likely to own their own home rather than be a tenant. Prior research has indicated that tenants are less likely to use their heating controls or engage in energy saving behaviours than home owners.¹⁶



¹⁶ Consumer Focus (2012). Consumers and domestic heating controls: a literature review. London: Consumer Focus. Available at - <http://www.consumerfocus.org.uk/files/2012/01/Consumers-and-domestic-heating-controls-a-literature-review.pdf> (16/12/2013).

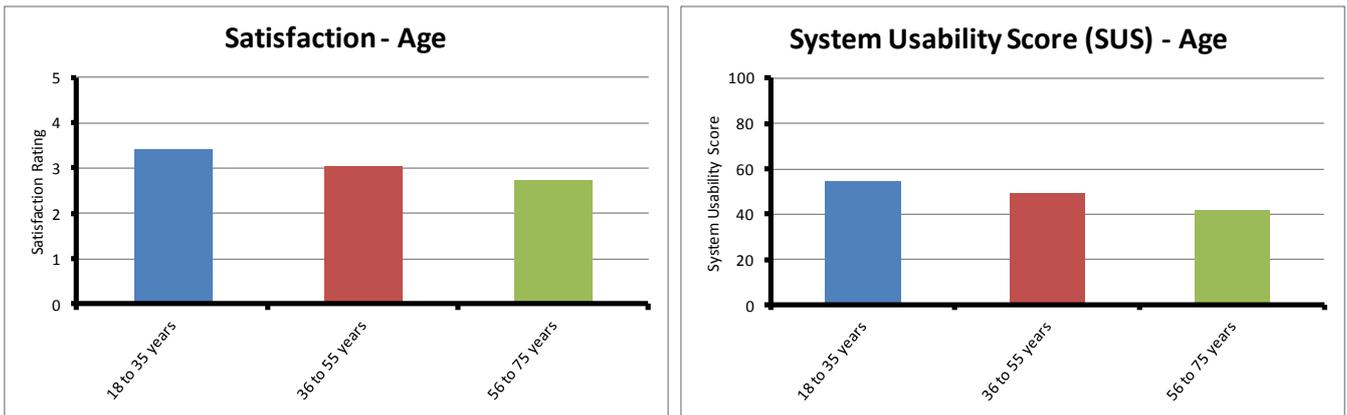


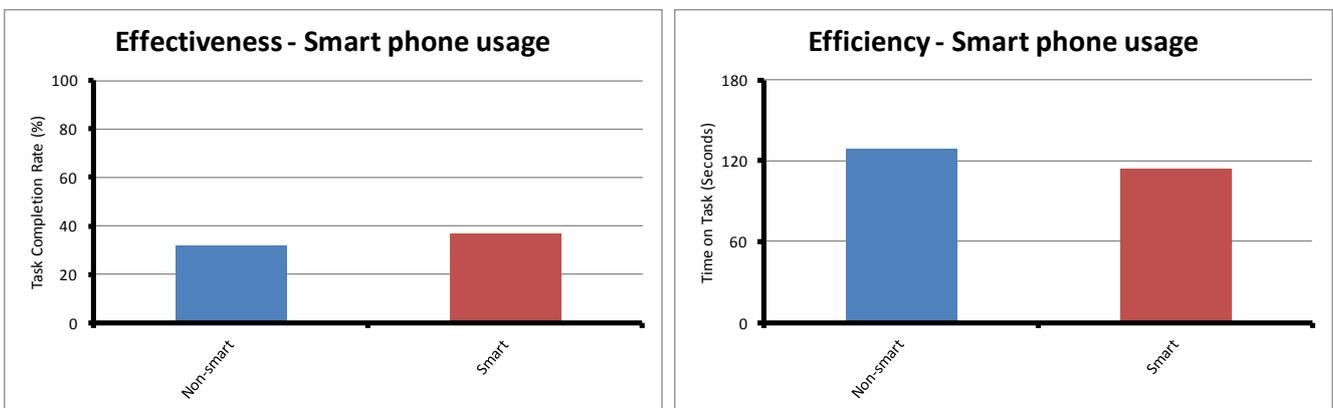
Figure 5: Usability metrics for smarter heating controls when participant age is considered as an independent variable.

5.2 Smart phone ownership

Smart phone ownership was considered as an independent variable as all the controls required interaction with a smart phone app for at least one of the tasks on which they were evaluated. Users who already own and use smart phones may therefore benefit from prior experience with this technology. In addition, several of the heating controls tested employed digital displays and touch screen interfaces which may be more familiar to participants who have experience with smart phone technology.

The smarter heating controls were shown to perform better on all four usability metrics for smart phone owners versus non-smart phone owners. This held true for the majority of the individual tasks across all devices, not only tasks that involved the use of a smart phone. Results are illustrated in Figure 6.

In the sample tested for this research, there was a skew for users from the younger age groups to own smart phones. It is therefore difficult to conclude whether the difference in performance is strictly due to smart phone ownership, or due to the smart phone owners group being made of a larger proportion of younger users who may have more general aptitude for technology due to their exposure to connected devices from a younger age.



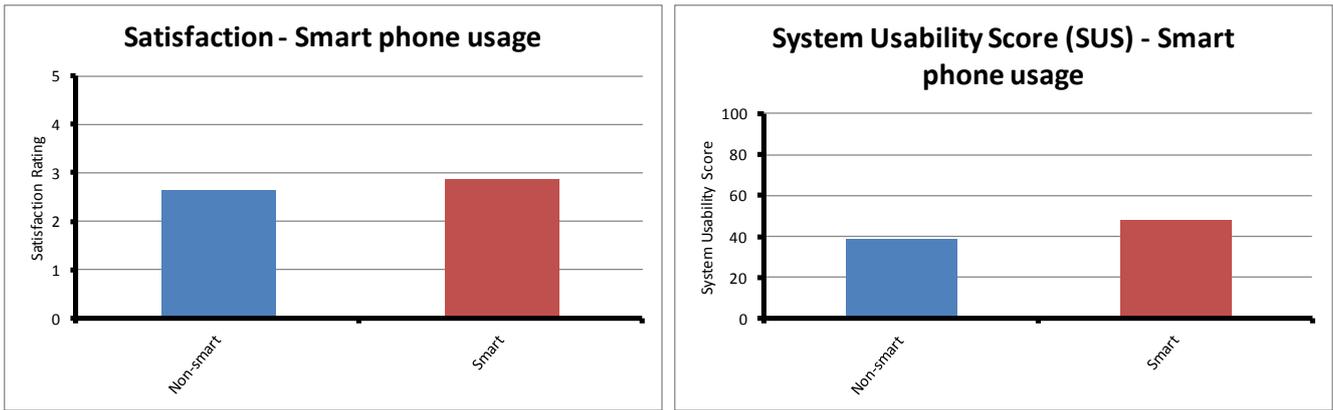


Figure 6: Usability metrics for smarter heating controls when smart phone ownership is considered as an independent variable.

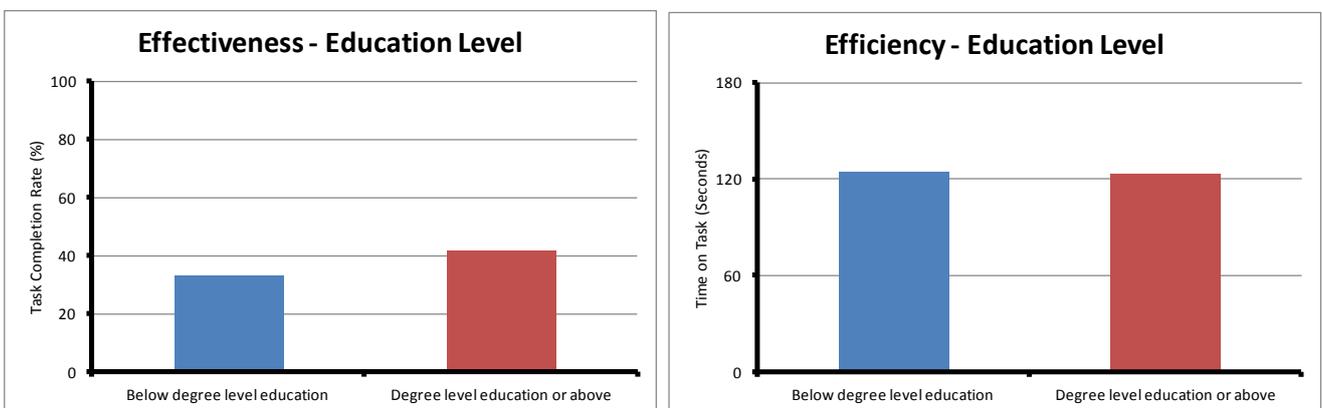
5.3 Education Level

Particularly complex user interfaces may place a high cognitive or memory load on users. Previous research has shown that education level tends to correlate well with cognitive and memory ability across the population. The education level of users was therefore considered as a factor that may influence the usability of smarter heating controls.

Participants were divided into two groups, those with degree level education or higher, and those who did not have a degree or a higher qualification. The results showed that the smarter heating controls were easier to use for participants who had degree level education or higher.

For controls in Group A, users with degree level education or higher were able to complete significantly more tasks ($P < 0.05$), with higher reported satisfaction levels ($P < 0.01$). Controls in Group B showed no significant effect of education level on performance.

This may indicate that the heating controls placed significant cognitive demands on participants due to the need to interpret, evaluate and remember specific commands within the interface. The results are shown in Figure 7.



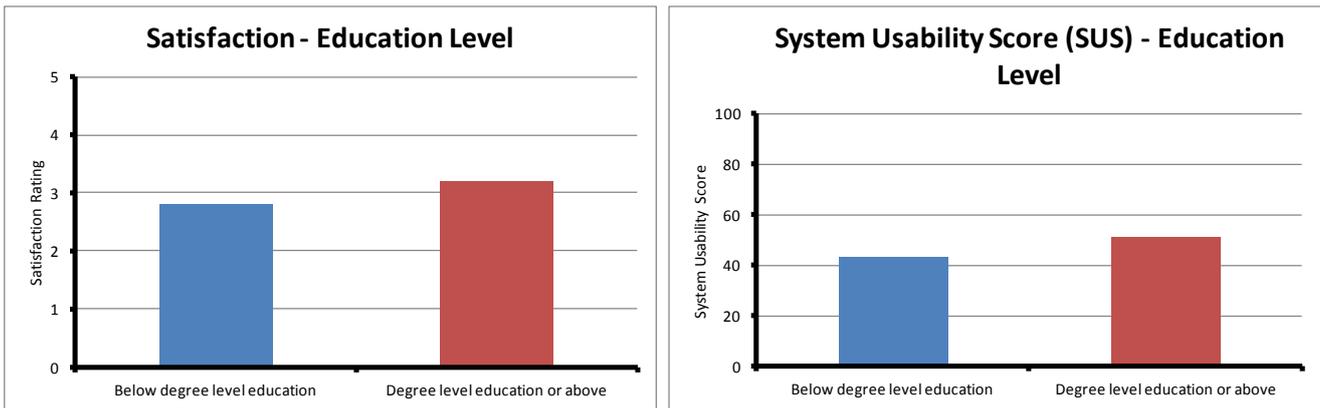
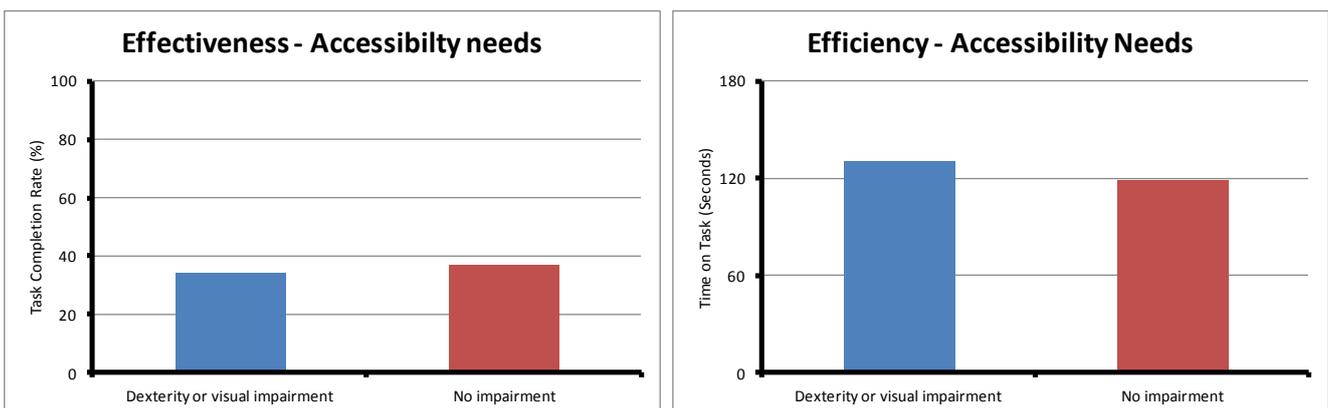


Figure 7: Usability metrics for smarter heating controls when education level of participants is considered as an independent variable.

5.4 Users who have a mild dexterity or visual impairment

Use of the smarter heating controls required users to perceive information shown on PC monitors, smart phone screens or wall mounted displays. The tasks also required users to manipulate controls using a touch screen, mouse, or the user interface of the wall mounted displays. Users who experience either visual or dexterity related impairments may have additional accessibility needs when attempting to engage in energy saving behaviours with smarter heating controls.

The results showed that the usability of the smarter heating controls was rated as lower when the presence or absence of a mild dexterity or visual impairment was considered as a factor. Satisfaction was rated slightly higher for users who had a mild dexterity or visual impairment. These results were informed by the data of 12 participants and therefore must be considered as an illuminative case study. The results are illustrated in Figure 8.



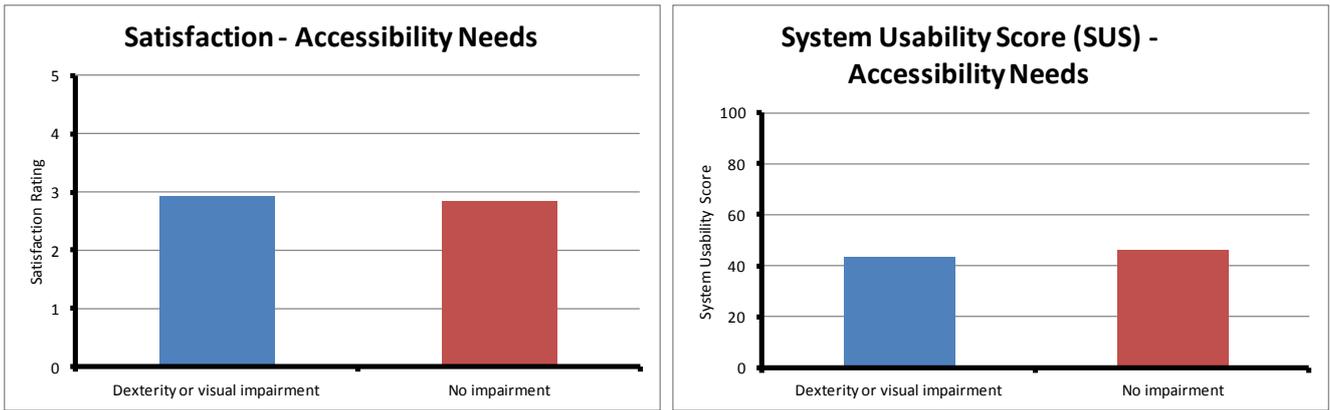


Figure 8: Usability metrics for smarter heating controls when accessibility needs of participants is considered as an independent variable.

6. Installing Smarter Heating Controls

Each participating manufacturer completed a short survey that asked them to provide information relating to the estimated resources required for initial installation and set-up of their device. Users in this study did not experience the installation process and therefore the survey findings do not represent a measure of consumer satisfaction, but rather a supplementary level of context relating to home installation.

To support anonymity the systems names have been replaced with characters, findings are presented thematically and the discussion of findings is restricted to a generic overview rather than a case by case examination of devices.

System	Can homeowners install the device themselves?	What is the level of skill required to install the device? ¹⁷	Estimated time for installation
A	No	A heating engineer is required to connect a wireless device to the boiler. The remaining steps could be done by the homeowner.	2 hours or under
B	No	A technician is required to correctly install this system. A qualified electrician may also be required to install wiring between the boiler and desired rooms.	2 hours or under
C	No	A technician is required to correctly install this system. A qualified electrician may also be required to install wiring between the boiler and desired rooms.	2 hours or under
D	No	A control unit needs to be connected to the boiler and this is done by a qualified technician A qualified electrician may be required to wire cables	2 hours or under

¹⁷ As opposed to the replacement of an existing Smarter Heating Control, which is estimated to be less resource intensive, as existing wiring will already be in situ

System	Can homeowners install the device themselves?	What is the level of skill required to install the device? ¹⁷	Estimated time for installation
		through the home.	
E	No	A qualified engineer should install this system.	6 hours or over
F	No	A qualified electrician may be required to install wiring between the boiler and desired rooms.	2 hours or under

Figure 9: Findings from the manufacturer’s survey.

The manufacturer’s survey indicates that:

- These Smarter Heating Controls (SHCs) are not ‘box ready’; consumers cannot merely purchase a system, take it home and plug it in.
- Not all SHCs are the same; some work by a single device that replaces the existing thermostat, if this exists, whereas most have a series of components that are deployed across the house.
- A high level of technical skill is will be required to install the device into the existing heating system. For most devices installation will require a qualified Heating Engineer, and all devices will require electrical cables to be installed between rooms where these do not already exist.
- Installation is estimated to be reasonably quick, with most devices requiring 2 or less hours of time by a qualified/skilled technician (assuming electrical wiring is already in place), although more time may be required, depending on the number of components (sensors etc.) dispersed throughout a dwelling.
- A key feature that will influence installation is whether system components are mains powered, requiring wiring, or battery powered, allowing relatively cost free installation.

7. Conclusions and implications

7.1 Usability of smarter heating controls

The main objective of this research was to evaluate the usability of a range of smarter heating controls to assess their suitability for supporting energy saving behaviours in the home. Controls were assessed for their performance against a number of tasks derived from user needs identified in prior research conducted on behalf of DECC in the Smarter Heating Controls Research Programme. This research was unable to address all the user requirements flowing from the research programme's report, Rubens, S., Knowles, J. (2013). *What people want from their heating controls: a qualitative study*. This was due to a lack of support for feedback on potential cost savings (the number one rated user need from that emerged from the qualitative study), nor could this research evaluate the perceived rapidity of heating, which was also a key user requirement. These findings should therefore be couched in these known limitations of the research.

The results of the usability testing suggest that there are significant barriers to use with all of the smarter heating controls tested as part of this study. It is strongly recommended that difficulty in using controls cannot be ruled out as a potential factor underlying any failure to engage in energy saving behaviours in any subsequent field trials involving these controls.

The smarter heating controls that were tested did not meet the benchmarks for usability that were established prior to conducting the fieldwork. Three of the controls were able to meet the benchmark that was set for efficiency, which suggests that they may support users in achieving tasks within a reasonable and desirable amount of time. However, these results need to be subject to the caveat that these controls did not meet the minimum benchmarks set for effectiveness and satisfaction, suggesting that users experienced barriers to completing tasks, and that they perceived the controls to be difficult to use overall.

Observation of participants' behaviour while attempting to complete tasks and their subjective comments regarding the controls (both spontaneous and solicited) revealed the following potential usability issues underlying performance of the controls:

- Complex setup of schedules
- Lack of error prevention
- Iconography, labelling and information architecture
- Visual design and interaction metaphors
- Difficulty identifying system state
- Lack of confirmatory feedback

7.2 Implications for manufacturers

The insights provided by this research can be used as a basis for manufacturers of smarter heating controls when considering future innovations of their products. In particular, the barriers to use observed with the smarter heating controls suggest fruitful areas for innovation where manufacturers may significantly increase the ease of use. The impact of the usability issues observed during the research could potentially be mitigated by manufacturers of smarter heating controls taking steps to address the barriers to use. The remainder of this section outlines these opportunities when considered against the main findings of the usability research.

7.2.1 Provide a more linear first time user experience

Participants struggled to setup and modify heating schedules due to the complex, multi-stage process required with the majority of the controls tested in this study. Participants were unable to complete some scheduling tasks as a number of controls gave no visibility of the necessary steps involved.

Providing a step-by-step, linear process (a “wizard”) to guide first-time or unconfident users through the activity of setting up a heating schedule has the potential to reduce the ambiguity surrounding the necessary steps that the user needs to complete. One of the controls used a wizard for initial setup. While performance for this task failed to meet the benchmarks for usability, it was observed that the wizard did support participants towards successful completion of this task.

The wizard should walk users through all necessary steps to setup a heating schedule, including editing the timers, setting the target temperatures, and configuring any necessary zonal settings to allow independent control of different rooms in the household.

7.2.2 Improved capacity for error detection and recovery

Several controls failed tasks because they allowed the user to make errors such as entering an invalid time or date, or exiting the setup process without saving changes.

A control with good usability should support the user in identifying potential errors and then helping them to recover. For example, using appropriate fields to constrain user inputs only to relevant values, or notifying the user if they enter a value that potentially causes a conflict with another parameter (for example, two time periods which overlap each other).

Controls should prompt the user to confirm or save any changes they make to the schedule instead of allowing them to exit the schedule editor without saving changes.

An undo option may help users who make changes to the heating schedule and want to quickly return to its previous state.

7.2.3 Use visual design that communicate the function of interface elements

Standard interface elements of user interfaces such as buttons and menus should communicate their function to users by employing appropriate visual design. This will help users identify active elements of the user interface, and help set their expectations about possible actions supported with the interface.

Where possible, smart phone apps should strive to follow the visual conventions and metaphors established by the device’s operating system. This will help to ensure learnability of the app and the transfer of experience from the use of other apps.

7.2.4 Strive for a simpler information architecture

Users found it difficult to locate and identify the correct actions needed to complete tasks. The structure of menus on the devices did not follow logical groupings in some cases. For example, users incorrectly looked under “settings” menus when trying to make changes to the heating schedule.

The information architecture (IA), how the various elements of the control are grouped into a hierarchy of menus and sub-menus, should closely follow the target user group’s mental model. In other words, the different options should be presented where users expect them to be. Further user centred research is required to understand this mental model.

Limiting the number of options presented to the user will reduce the cognitive load of the interface (fewer decisions for the user to make), but the IA should also be as “flat” as possible, avoiding multiple levels of sub-menus which users will find more difficult to recall and more time consuming to navigate. Similar or related commands should be grouped together within this structure.

Having a flatter IA may promote the speed with which users can learn to use the control as they will become more familiar with a flatter menu structure through repeated exposure than they would with a larger, more complex IA.

Similarly, the controls tested as part of this research that had simpler wall mounted units with fewer options were able to meet the usability benchmark for efficiency. This suggests that presenting a more limited set of core options may reduce the time users take to perform key tasks via the wall mounted units. This may increase the likelihood of users adopting energy saving behaviours. These controls circumvented the need for a complex wall mounted unit by providing more detailed controls through smart phone or web based user interfaces. The wall mounted units were designed to support quick, frequent tasks such as turning the heating on or off temporarily, while more involved, infrequent tasks such as configuring an overall heating schedule were handled via a web portal. Therefore, these devices did not have a lack of support in any of the core tasks tested despite simpler wall mounted units.

7.2.5 Adopt the user’s language when labelling interface elements

Participants had difficulty identifying which actions to take with controls due to a lack of understanding of the labels used for certain steps.

The control should adopt the user’s language and terminology for the labels of actions and options within the system. Technical terms should be avoided where possible. Identifying specific language and labels that would be more easily understood by users would require more user centred research.

The labelling of actions and menu options should also strive to avoid ambiguity or duplication where possible.

7.2.6 Provide persistent feedback on the system state

It was difficult for participants to identify whether systems were on timer mode, constant heating, or switched off. It was also not clear to participants how these different states interacted with each other. For example, in some controls “continuous” mode could be switched on but would still operate based on a target thermostatic temperature, resulting in the heating not being correctly switched on.

A control should provide clear, unambiguous information that allows the user to quickly identify which mode or state it is currently in. When the system is “on” this should clearly distinguish between whether the timers are active, continuous heating mode is on, and whether the heating is currently activated based on any thermostatic controls.

The control should also indicate whether any thermostatic controls are currently active and, if so, whether modifying these controls changes the current state of the heating (analogous to the audible “click” that can be heard with some standard thermostatic heating controls).

7.2.7 Provide feedback and closure on user actions

Participants were often not confident that they had successfully completed tasks due a lack of confirmation provided by the controls. This could lead to users surmising that they were unsuccessful at the task, looking for additional steps to complete, or repeating the steps involved in the task.

Any actions undertaken by users that modify the state of the control should clearly communicate the outcomes of those modifications, for example, unambiguously confirming that the heating was turned on or off temporarily. This gives users reassurance and closure on the task, and may improve their confidence or satisfaction in the control’s ability to support that task.

When a user is in the same location as the heating system, there may be a degree of intrinsic, environmental feedback such as the sound of the boiler or radiators turning on or off, or a perceivable increase in the temperature of the room or radiator. However, in the case where users are interacting with the heating remotely, such as via a smart phone or via a web interface, it is imperative that the user interface clearly communicates any changes that have been made to the heating system.

7.2.8 The process of installation

These Smarter Heating Controls are not a ‘box ready’ solution for heating control, as they will involve installation into a consumer’s existing heating system. For most consumers this is likely to entail additional resources:

- A Qualified Heating Engineer, to connect the device with the boiler and existing electrical connections (e.g. left from pre-existing thermostats)
- An electrician for rewiring electrical connections to power new control units deployed in all or particular rooms
- Additional wall mounted units
- Additional modifications to each or selected radiators

Smarter Heating Control systems that offer wireless connections between the room thermostats, controller, boiler and radiator valves arguably carry the lowest overall installation cost (in terms of money and inconvenience).

7.3 Implications for consumers

Rubens and Knowles (2013) outlined key user requirements for future heating controls based on an understanding and analysis of existing user behaviours and needs. The usability tests conducted for this research were designed to test the ability of the current set of smarter heating controls to address these needs. However, the results suggest that without considerable

development, difficulty of use is likely to present a barrier to consumers wishing to engage in energy saving behaviours.

It is possible that with repeated use of controls, consumers may find operating the controls to be more effective, efficient and satisfying, although it is unclear whether such learning would lead to consumers using the controls to their fullest capacity. What is clearer is that any initial interaction, by a new tenant or home owner for example, is likely to be inhibited by usability issues that mirror the findings of this study.

From a consumer standpoint, the potential for saving energy and ultimately, spend on energy used for space heating, is likely a significant driver for adoption of smarter heating controls. However, the results of this testing cast some doubt on whether smarter heating controls of the type tested are usable enough now to allow users to reap the benefits of energy saving. In addition, data from the manufacturers survey suggests that there may be substantial costs and expertise involved in installing smarter heating control systems that consumers will have to consider against any potential longer term savings made possible by the controls.

6.4 Implications for DECC

There are three key implications for DECC from this research with regards to policy and future research undertaken as part of the Smarter Heating Controls Research Programme.

The first implication is that the results of the usability testing strongly suggest that difficulty using these smarter heating controls cannot be ruled out as a potential reason for participants of future trials failing to engage in energy saving behaviours. This raises potential concerns over the utility of future field trials to deliver deeper insights into how and why smarter heating controls can impact on domestic energy consumption. There is a risk that trials will result in a lack of energy saving behaviours due to controls being difficult to use. Controls that are overly time consuming to use or perceived as overly technical or frustrating to use pose a potential barrier to motivation of users to engage in energy saving behaviours.

The second implication is regarding future engagement between DECC and the manufacturers of smarter heating controls. The results of the usability trials constitute part of a growing body of evidence from this and previous work packages, firstly, that current heating controls (both standard and smarter versions) tend to be difficult for many people to use, and secondly, that smarter heating controls may not support people's requirements as a result of this difficulty of use. This presents an opportunity for DECC to work closely with manufacturers to help ensure future innovations in smarter heating control to address barriers to understanding and effective use.

Finally, illuminative findings from this research suggest that smarter heating controls of the type tested are unlikely to constitute a universal solution for energy saving behaviours. Significant differences in performance were seen for the controls when different demographic user groups were considered. Results suggest that older users and users of a lower cognitive or memory ability may not be adequately supported by the current marketplace of devices.

