



Department
of Energy &
Climate Change

Usability testing of smarter heating controls

December 2013 – Appendices

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 – Appendix*. 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

Appendices

Prepared by Amberlight

December 2013

Contents

Introduction	5
Appendices	6
Appendix A - Participant sample details.....	6
Appendix B - Manufacturers survey	8
Appendix C - Scenario descriptions provided to participants	12
Appendix D - System Usability Scale Survey.....	15
Appendix E - Notes on methodological limitations of usability testing.....	17
Appendix F - Results of correlation analysis for usability metrics	19
Appendix G - Statistical comparison of age and education levels.....	20

Introduction



This document contains the appendices to Wall, S. and Healy, F. (2013). *Usability testing of smarter heating controls* main report. Amberlight conducted the research and prepared the report for the Department of Energy and Climate Change (DECC) between July and August 2013.

The overall purpose of the research was to assess the usability of smarter heating controls for suitability for future research trials. Amberlight conducted summative usability testing of 5 smarter heating controls with a sample of 72 participants. Participants were divided in to two matched groups of 36 participants each. Each group evaluated 3 heating controls (one of the heating controls was tested with both groups, with one group focussing on the web portal as a platform for that particular service, and the other group focussing on the wall mounted unit). Each participant attempted 8 compelled tasks with the 3 smarter heating controls assigned to their group. Metrics were recorded for effectiveness, efficiency and satisfaction for each controller. The overall metrics for each device were compared to a benchmark level of performance to determine whether difficulty using smarter heating controls may potentially pose a barrier to people engaging in energy saving behaviours.

The following appendices contain supporting materials used in the study but not included in the main report.

Appendix A – Participant sample details

Each device was tested by a sample of 36 users according to the following profile:

Gender	 male N = 18			 female N = 18		
Age	18 to 25 N = 5	26 to 35 N = 7	36 to 45 N = 7	46 to 55 N = 7	56 to 65 N = 5	66 to 75 N = 5
Education Level	Below or completed GCSE N = 16			Above GCSE N = 20		
Household Income	Below £12,000 N = 8		Above £12,000 N = 28			
Experience with technology	Smart phone user N = 18			Non-smart phone user N = 18		
Accessibility Needs	Low dexterity N=3	Low vision N=3	No accessibility needs N = 30			
Attitude	Intention to save energy or save household costs through cutting energy use N = 36					
Previous Experience	Have a central heating system and boiler installed at their current home N = 36					

There were 72 participants in total, from 2 groups. The 2 groups were matched in terms of key criteria.

User ID	Age range	Education level	Category of phone	How much is your total pre-tax household income?	Accessibility Screener
1	46 to 55	Completed A Level	Non-SMART	£30,001 - £50,000	No
2	66 to 75	Below or completed GCSE	Non-SMART	£30,001 - £50,000	No
3	46 to 55	Completed A Level	Android	Above £50,000	Mild visual impairment
4	46 to 55	Completed A Level	Non-SMART	£30,001 - £50,000	No
5	36 to 45	Completed A Level	iPhone	£12,001 - £30,000	No
6	36 to 45	Graduate (Masters or above)	Non-SMART	£30,001 - £50,000	No
7	18 to 25	Graduate (Masters or above)	Android	£12,001 - £30,000	No
8	26 to 35	Undergraduate	iPhone	£30,001 - £50,000	No
9	26 to 35	Completed A Level	Android	Below £12,000	No
10	36 to 45	Below or completed GCSE	Non-SMART	£12,001 - £30,000	No

11	36 to 45	Undergraduate	Android	Above £50,000	No
12	26 to 35	Undergraduate	iPhone	£12,001 - £30,000	No
13	26 to 35	Undergraduate	Android	£30,001 - £50,000	Mild visual impairment
14	66 to 75	Below or completed GCSE	iPhone	£30,001 - £50,000	No
15	18 to 25	Completed A Level	iPhone	£12,001 - £30,000	No
16	56 to 65	Completed A Level	Non-SMART	£30,001 - £50,000	Low dexterity
17	56 to 65	Below or completed GCSE	Non-SMART	Below £12,000	Mild visual impairment
18	56 to 65	Below or completed GCSE	iPhone	£30,001 - £50,000	No
19	66 to 75	Below or completed GCSE	Android	Above £50,000	No
20	18 to 25	Completed A Level	iPhone	£12,001 - £30,000	No
21	36 to 45	Undergraduate	Non-SMART	Above £50,000	No
22	46 to 55	Completed A Level	Non-SMART	Below £12,000	Low dexterity
23	56 to 65	Completed A Level	Android	£12,001 - £30,000	No
24	18 to 25	Graduate (Masters or above)	Android	Below £12,000	No
25	56 to 65	Completed A Level	Non-SMART	£12,001 - £30,000	No
26	18 to 25	Completed A Level	iPhone	£12,001 - £30,000	No
27	46 to 55	Completed A Level	Non-SMART	Below £12,000	No
28	66 to 75	Completed A Level	Non-SMART	£12,001 - £30,000	No
29	56 to 65	Below or completed GCSE	Android	£12,001 - £30,000	No
30	26 to 35	Undergraduate	iPhone	Below £12,000	No
31	46 to 55	Below or completed GCSE	Android	Above £50,000	No
32	66 to 75	Below or completed GCSE	Non-SMART	£30,001 - £50,000	No
33	36 to 45	Below or completed GCSE	iPhone	£12,001 - £30,000	Mild visual impairment
34	46 to 55	Undergraduate	iPhone	£30,001 - £50,000	No
35	36 to 45	Undergraduate	Android	£12,001 - £30,000	No
36	46 to 55	Undergraduate	Non-SMART	£30,001 - £50,000	No

Appendix B – Manufacturer survey

Manufacturer Survey: Questionnaire

More About Your Device

- 1) Name of Manufacturer:
- 2) Name of device:
- 3) What is the recommended retail price of your product/system:
- 4) What steps are required to install the device/system in a consumer's home?

Steps	Level of technical knowledge required	Who is likely to meet this step?
	High/Medium/Low	Consumer/Technician/ Heating engineer
1.		
2.		
3.		
4.		
5.		
6.		

- 5) How long do you estimate an average domestic installation will take (in hours)?
- 6) Do you plan/envisage that installation will be offered at point of sale?
- 7) If so what do you project the cost of this service will be to the average domestic customer?

8) What platforms can be used to access/control the device/system?

Platforms	Level of functionality accessible via platform	
	Yes/No	Full/Semi
1. Manual control panel		
2. Website		
3. Smartphone app		
4. Tablet app		
5. Other (please state)		
6. Other (please state)		

9) Please list/detail any minimum technical specifications and/or device requirements the customer must meet in order to be able to install and use your system (e.g. If they must have their own smartphone or tablet in order to use the system or if they need home wifi)?

10) What is the target audience for your product?

Target	Yes/No (and any relevant details)
a) General population (i.e. no specific target, targeting as broad a base as possible)	
b) Specific audience or demographic (please specify which audience)	

Manufacturer Survey: Results

The table below represents a summary of survey responses. The data has been presented in this manner to support the anonymity of respondents (manufacturers).

Questions	Range of responses
What is the recommended retail price of your product/system	<p>Typical costs are between £108 and £250 per controller, however to achieve zonal control consumers may need to purchase multiple control units (in one case £108 each and in another £250 each).</p> <p>This cost does not include any modifications required to electrical wiring or TRVs.</p>
What steps are required to install the device/system in a consumer's home?	<p>Manufacturers reported a 4 – 6 step process, where all but one system required a qualified technician to install hardware connected to the boiler system or wall mounted control units. One system could be installed by a competent DIY enthusiast, if they were the homeowner.</p> <p>All but one system appears to require electrical rewiring if the household does not already have room thermostats installed. The system that doesn't need pre-existing wiring, works on a wireless connection and can be powered via an external mains adapter. If the user wanted to hide the power cables, they would then need to run the cables into their walls.</p>
How long do you estimate an average domestic installation will take (in hours)?	In most cases installation by a technician would take 2 hours or under and in one case 6 hours or over.
Do you plan/envisage that installation will be offered at point of sale?	<p>3 manufacturers plan to offer an installation service through official retailers and utility providers.</p> <p>1 system is aimed at new builds and property developers</p> <p>1 system has no confirmed strategy for this</p>
If so what do you project the cost of this service (installation) will be to the average domestic customer?	Between £60 and £130 in most cases although this will be dependent upon whether the consumer wishes to deploy additional components and whether these require installation by a qualified electrician. A system that offered comprehensive control across the whole house could be substantially more costly.
What platforms can be used to access/control the device/system?	<p>All systems offer a web interface and a smartphone app, so remote control is possible on all.</p> <p>2 systems offer advanced proprietary wall mounted units that allow access to virtually all the functions. This may be an important consideration as users can still control the system even without a web connection.</p>

	<p>One system does not offer and user interface hardware. The user interacts with the system via a web app that can be accessed on a tablet, smartphone or through a web browser.</p> <p>2 systems had very basic wall mounted units that served mainly as thermostats with only basic controls.</p> <p>One system did not support zonal control at all.</p>
<p>Please list/detail any minimum technical specifications and/or device requirements the customer must meet in order to be able to install and use your system</p>	<p>Varies between systems, but in general an internet connection is required and in some cases a wi-fi network to connect to.</p> <p>All systems require a suitably modern central heating system.</p> <p>3 systems require a hard wired power supply or main unit connected. One could be powered by internal and changeable battery.</p> <p>One system states that it can offer full control and functionality with no internet connectivity at all, just a connection to the heating system. This system could be considered the most stand alone.</p>
<p>What is the target audience for your product?</p>	<p>4 systems are targeting the general population, any home with a suitable central heating system. Of these one is focus more on fuel poverty areas and the cost of the control unit reflects this.</p> <p>Another is targeting more tech savvy customers and the advanced nature of the control unit reflects this. 1-2 are targeting customers of specific utility companies and these companies will subsidise the cost.</p> <p>1 system appears to be targeting a very different market. It focuses on new builds and redevelopment projects and sees developers and builders as potential specialist customers/resellers.</p>

Appendix C – Scenario descriptions provided to participants

Task Number	Task purpose	Scenario provided to users
1a	Setup the weekly heating schedule for a two room house (bedroom and living room)	<p>Imagine that you have installed a new heating control system at your home. Your heating schedule has not been setup.</p> <p>Using the <platform>, can you please set up a heating schedule for the whole house based on the information below?</p> <p>Monday to Friday Turn the heating on at 7am when you get up, and turn the heating off at 8am when you leave for work Turn the heating on again at 7pm when you arrive home, and turn the heating off at 11pm when you go to bed</p> <p>Temperatures should be set at 20°C</p>
1b	Edit the heating schedule for the bedroom to come on earlier one day per week	<p>Imagine that you need to wake up a little earlier every Wednesdays.</p> <p>Using the <platform> can you please set up a heating schedule for the bedroom based on the information below?</p> <p>Wednesday You want to make sure your upstairs bedroom will start being heated from 6:30am when you wake up to 7:30am when you leave the house. Temperature should be set at 20°C for the bedroom only. Please edit the bedroom schedule without affecting any other schedules.</p>
2	Edit the heating schedule remotely (using mobile or	<p>Imagine that you are at work at the moment. You've just changed your working times so that</p>

Task Number	Task purpose	Scenario provided to users
	desktop) to come on earlier one day per week	<p>you will be home early on Wednesdays starting from today. You want to change your heating schedule so that you return to a warm house later today.</p> <p>Using your <platform>, can you please change your heating schedule so that the heating will come on at 20°C between 4pm to 10pm for the whole house every Wednesday?</p>
3	Temporarily switch the heating on when returning home	<p>Imagine that you just came home at 4pm and the heating is not yet switched on because it normally comes on at 7pm.</p> <p>Using the <platform>, can you temporarily switch the heating on for the whole house, without cancelling the saved schedule, so that the program will return back to normal later automatically?</p>
4	Temporarily stop the heating schedule for 1 week while on holiday, ensuring the system is protected in the event of very cold weather	<p>It is a cold December. Imagine that you are going to France for 1 week. You don't want to waste energy by leaving your regular heating schedule running, but you also don't want your system to freeze over when it's cold.</p> <p>Using the <platform>, can you temporarily stop the heating schedule for 1 week, ensuring you are protected in the event of very cold weather?</p>
5	Temporarily switch the heating off, without affecting the schedule	<p>Imagine that you are leaving home for 2 to 5 hours, and you want to temporarily switch the heating off while you are away today.</p> <p>Using your <platform>, can you temporarily switch off the heating, without cancelling the saved schedule, so that the program can be returned back to normal later automatically?</p>
6	Turn the heating on remotely (using mobile or desktop)	<p>Imagine that you are outside your home at the moment, and you are returning home earlier than expected. Your heating is off now according to the schedule.</p>

Task Number	Task purpose	Scenario provided to users
		Using your <platform> , can you switch on the heating for the whole house temporarily without affecting the weekly heating schedule?
7	Find information about your energy usage	<p>Imagine that you have installed the new heating control system last month.</p> <p>Using the <platform>, can you tell me where you would find information about your gas usage of the whole house in the last month?</p>

Appendix D – System Usability Scale Survey

Please indicate the extent that you agree with the following statements:

1. I think that I would like to use this system frequently.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

2. I found the system unnecessarily complex.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

3. I thought the system was easy to use.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

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

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

5. I found the various functions in this system were well integrated.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

6. I thought there was too much inconsistency in this system.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

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

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

8. I found the system very cumbersome to use.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

9. I felt very confident using the system.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

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

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

Appendix E – Notes on methodological limitations of usability testing

While Amberlight have attempted to ensure as much as possible that the results of this study are purposive and repeatable, lab-based usability testing has inherent limitations due to the controlled and artificial nature of the environment in which it must take place. This section considers steps that were taken to mitigate limitations of the methodology and testing environment.

Participants in the usability testing were not given the opportunity to read the manuals for the smarter heating controls prior to attempting tasks and did not receive a briefing from a trained installer or heating engineer regarding operation of the controls. While both of these were considered for inclusion in the testing, they were rejected on the grounds that it would be problematic to control for the quality of the briefing, user manuals, or how much attention participants would devote to them. Anecdotally, many users do not read manuals for home appliances prior to operation, and several participants in the study spontaneously mentioned this behaviour. Ultimately, users not having to refer to manuals, help lines, or call outs for heating engineers could be considered of benefit to manufacturers in terms of the impact on on-going costs incurred for these services. As a compromise, manuals and quick start guides were provided for participants during the sessions, and they were free to look at them or not during the time allocated for tasks.

The usability laboratory also did not provide environmental feedback based on user actions (e.g. noise from the boiler turning on, pipes heating up), which could affect user performance through reinforcement of actions. However, in practice this may not be available for several of the tasks if the user was in a different room from the boiler, and due to the lag between the boiler turning on and pipes heating up. It would also not be applicable for any of the “remote” tasks (task 2 and task 6) and tasks that did not involve the heating turning on or off immediately as part of the success criteria (task 1a, 1b, 5, and 7).

Finally, it is inadvisable to estimate the impact of learning effects through continued and frequent use of controls, as this was not explicitly tested as part of the methodology. It is possible that with repeated exposure to controls, users may find using the controls to be more effective, efficient and satisfying.

Appendix F – Results of correlation analysis for usability metrics

Pearson coefficients of correlation were calculated for each pair of usability metrics, for each smart heating control. Low co-efficients of correlation that would not support the argument for a single, consolidated measure, such as the M-statistic, are highlighted below in yellow using 0.2 as a cut-off value. This is quite a generously low value and low cut-off of 0.3 is more typical.

Note that in all devices, average task completion score and average task time do not correlate well. There seemed to be no simple overall picture. Incomplete tasks tended to have taken longer. Partial completes could go two ways – partially complete tasks were either done about as quickly as completed tasks or about as equally as incomplete tasks. This suggests that in the minds of the users, if they could evaluate that it was not complete, they carried on longer as they would for a failed task. If they wrongly evaluated that they had done the task, they stopped sooner like those who had completed the task.

Overall then, the lack of a systematic pattern of correlations and the failure of completion to provide a consistent picture, it was not recommend to combine the three separate scores as a single metric. The only very consistent result is the agreement between SUS scores and average satisfaction scores for each task where the correlation is never less than 0.67. Only one of these measures need be used in future.

Below is the Pearson co-efficient for each pair of metrics for each heating control:

System A:

	Effectiveness	Efficiency	Satisfaction	SUS
Effectiveness		-0.20	0.45	0.46
Efficiency			-0.54	-0.52
Satisfaction				0.69
SUS				

System B:

	Effectiveness	Efficiency	Satisfaction	SUS
Effectiveness		-0.16	0.39	0.26
Efficiency			-0.34	-0.14
Satisfaction				0.70
SUS				

System C:

	Effectiveness	Efficiency	Satisfaction	SUS
Effectiveness		0.00	0.72	0.37
Efficiency			-0.16	-0.26
Satisfaction				0.67
SUS				

System D:

	Effectiveness	Efficiency	Satisfaction	SUS
Effectiveness		-0.04	0.44	0.25
Efficiency			0.04	0.19
Satisfaction				0.70
SUS				

System E:

	Effectiveness	Efficiency	Satisfaction	SUS
Effectiveness		-0.16	0.44	0.40
Efficiency			-0.36	-0.22
Satisfaction				0.83
SUS				

System F:

	Effectiveness	Efficiency	Satisfaction	SUS
Effectiveness		-0.18	0.29	0.32
Efficiency			-0.53	-0.49
Satisfaction				0.73
SUS				

Appendix G – Statistical comparison of age and education levels

The two factors of age and education were believed to be important for people's ability to use the different devices. With only 36 people in each group, there is a risk of slicing the data too finely to be useful. To mitigate the danger of considering small sample sizes, education was grouped by 'A' level and below or undergraduate and above. Age was grouped by 35 or below, 36 to 55, 56 and above. This gives group sizes of:

Group A:

	18-35	36-55	56 above	Total
Lower Education	4	9	11	24
Higher Education	6	6	0	12
Total	10	15	11	36

Group B:

	18-35	36-55	56 above	Total
Lower Education	5	9	10	24
Higher Education	5	6	1	12
Total	10	15	11	36

Whilst it would ordinarily be desirable to treat these as two factors in an ANOVA of the different measures, the group of 56 year olds and above who have higher education was under-represented. This uneven-ness undermines the effectiveness of ANOVA. The factors are therefore considered separately.

Whilst it might be useful to consider each control separately, this would be to ignore relationships between the measures on the different controls. For instance, in both groups, average time to complete the tasks correlates well across all three controls. There are similar good correlations with satisfaction (though less so with SUS). Completion does not correlate across controls within Group A but does correlate well within Group B. However, it is still meaningful to look at overall completion rates across all controls and with Group B, this is in fact a wise thing to do because of the correlations.

The following measures of performance were therefore used:

1. Total of the average completion scores across all three controls in the group
2. Total of the average completion times across all three controls in the group
3. Total of the average satisfaction rating across all three controls in the group

SUS results are not reported as they give essentially the same picture as the Satisfaction ratings due to the high correlation observed between these two metrics.

For group A, the mean (sd) of the three aggregate measures by age and education are:

Age	Total Average Completion	Total Average Task Time	Total of Average Satisfaction
18-35	1.41 (0.39)	261.1 (76.4)	11.08 (1.73)
36-55	1.02 (0.32)	357.2 (82.2)	8.92 (1.52)
55 and above	0.97 (0.38)	374.6 (79.2)	8.51 (1.91)
Rankings * = $p < 0.05$ ** = $p < 0.01$ *** = $p < 0.001$	18-35 greatest **	18-35 quickest **	18-35 most satisfied **

Education	Total Average Completion	Total Average Task Time	Total of Average Satisfaction
Lower Education	1.00 (0.34)	337.4 (89.9)	8.81 (1.89)
Higher Education	1.34 (0.41)	332.6 (96.8)	10.57 (1.64)
Rankings * = $p < 0.05$ ** = $p < 0.01$ *** = $p < 0.001$	Higher > Lower *	No diff	Higher > Lower **

For group B, the mean (sd) of the three aggregate measures by age and education are:

Age	Total Average Completion	Total Average Task Time	Total of Average Satisfaction
18-35	1.45 (0.46)	320.4 (75.79)	8.7 (1.05)
36-55	1.01 (0.42)	389.69 (90.0)	7.92 (1.12)
55 and above	0.77 (0.42)	442.2 (64.9)	6.72 (1.32)
Rankings * = $p < 0.05$ ** = $p < 0.01$ *** = $p < 0.001$	18-35 greatest***	18-35 quickest**	55-above least satisfied***

Education	Total Average Completion	Total Average Task Time	Total of Average Satisfaction
Lower Education	0.96 (0.5)	391.0 (93.9)	7.53 (1.51)
Higher Education	1.26 (0.44)	377.5 (85.9)	8.25 (0.93)
Rankings * = $p < 0.05$ ** = $p < 0.01$ *** = $p < 0.001$	No diff	No diff	No diff

Overall, education has only a modest influence at best and only with the Group A devices. Age has substantial effect with younger people being able to complete more of the tasks, to do so quicker and to be more satisfied. Additionally older people are less satisfied with the group B devices.

Methodological Note: The perils of testing multiple factors

Whilst it could be interesting to break the sample down by further demographic factors and characteristics and perform further statistical analysis to explore where there might be significant differences in the results, this would not be an advisable or valid way to use statistical analysis.

With any set up where there is a level of uncertainty, usually reflected by the use of probabilities, there is always a degree of concern that any result may really be just a chance occurrence. The use of statistics in this context is not merely a matter of applying mathematics to numbers but an argument form couched not only in the mathematical analysis but in the experimental design that gathers the data and the reasons for doing the experiment in the first place.

Age and education were a priori concerns going into this study and this makes it legitimate to consider these in the analysis. However, the study was not solely designed with these in mind. Rather the dominant independent variable was the device being used. One way to see this is that an experiment purely to examine age and education would have looked quite different. Thus, whilst these factors were considerations the current data is not targeted to address them. This means that interpretation of the above tests must be cautious. They are indicative but not definitive.

There may be an argument that tests of this sort are widely done. This is true but they are also susceptible to over-confidence in the interpretations.

© Crown copyright 2013
Department of Energy & Climate Change
3 Whitehall Place
London SW1A 2AW
www.gov.uk/decc

URN 13D/339