How smart do smart meters need to be?

Nataliya Mogles¹, Ian Walker², Alfonso P. Ramallo-González³⁴, JeeHang Lee¹, Sukumar Natarajan⁴, Julian Padget¹, Elizabeth Gabe-Thomas², Tom Lovett¹, Gang Ren¹, Sylwia Hyniewska¹, Eamonn O'Neill¹, Rachid Hourizi¹, David Coley⁴,*

¹ Department of Computer Science, University of Bath, UK.
² Department of Psychology, University of Bath, UK,
³ Faculty of Computer Science, Universidad de Murcia, Spain.
⁴ Department of Architecture and Civil Engineering, University of Bath, UK,
d.a.coley@bath.ac.uk
* Corresponding author.

Graphical Abstract
Abstract

Governments across the world are investing in smart metering devices that report energy use to the user with the aim of reducing consumption. However, the effectiveness of such In-Home Displays (IHDs) has been questioned, since savings are small. This is possibly because informing the consumer of their consumption in kWh, or monetary units, fails to give context, or inform of possible actions to reduce consumption. We investigate, for the first time, the effect of replacing the simple statement of energy use an IHD gives, with a detailed array of information specifically...
designed to improve consumer energy literacy and suggest behaviour change through personalised actionable messages set against a series of psychological value systems for context, and which can identify potential profligacy. The results from a carefully controlled field experiment show: 1) value framing and action prompts have a significant effect on occupants' behaviour, with the mean temperature of homes being reduced from 22.4°C to 21.7 °C, and a marked reduction in gas consumption—22.0% overall and 27.2% in high consumers; 2) energy literacy increasing from 0.52 to 1.28 (on a 0-4 scale); 3) it is possible to target potentially profligate households, without inappropriately messaging others; 4) engagement is high, with 82% of the participants finding the system useful. These results emphasize the necessity of improving energy literacy when encouraging energy efficient behaviours and point to a new generation of smart meters with the potential to increase energy literacy, make much greater savings and impact climate change policy.

Keywords: energy feedback, In-Home-Displays (IHDs), internal values, action prompts, energy literacy, smart meter.
1 Introduction

The residential energy sector accounts for 23% of total energy consumption worldwide, placing it third after industry at 37% and transportation at 28% [1]. In developed countries the sector is even more important; in the US, for example, residential consumption represents 25% of total energy use [2] and in the UK 29% [3]. This translates to roughly 12% of UK greenhouse gas emissions [4]. In addition, it is projected that residential consumption worldwide will grow by an average of 1.4% a year from 2012 to 2040 [5]. Thus the sector is critical to both national energy policy and international climate change policy, and many attempts are being made to reduce residential consumption by influencing the behaviour of occupants.

Residential energy consumption is however a multidimensional phenomenon embedded within a socio-cultural and infrastructure context, and for this reason changing occupant behaviour might be expected to be complex. The focus of the research reported here lies in inducing behavioural change with the help of energy demand feedback via smart meters or ambient displays. Opinions regarding the effectiveness of such solutions to date are unfortunately not unanimous [6, 7, 8, 9, 10, 11, 12].

Approaches to energy feedback have so far tended to be one-size-fits-all solutions; however, with new developments in energy data management [13], more advanced feedback is now possible. The main contribution of the current paper to digital energy feedback research is twofold: i) in broad terms, it tests the effectiveness of a novel smarter, building-aware, and more user-personalised digital energy feedback in an experimental setting for the first time; ii) in more specific terms, it evaluates the effect of two new approaches to feedback – internal values and tailored action.
prompts, delivered via a computer tablet. It replaces the normal user- and building-blind smart meter concept with one that not only reports the energy used but recommends specific actions and works with the personal values of the user. For example, it might recommend turning the heating thermostat down one degree-centigrade, and explain this not just in terms of kWh or financial savings, but with respect to environmental gain or other personal values. The system tested here is a simplified version of a future building- and user-conscious approach that would report complex and tailored information to the user in written or spoken sentences. For example, “I note that the heating turns off at 9am each weekday, yet your home appears to be unoccupied from 7am; would you like me to change the heating timing? This might save you £47 per year and the stop the emission of 218 kg of carbon dioxide”. Or, “I note that your home might be over-ventilated when the heating is on, losing £98 of heat per year. You might like to keep most of your windows closed when the heating is on”. Such an approach, which we term intelligent smart metering (or ISM), requires knowledge of which values are most likely to prompt the user to act. These might be financial savings, the reduction of environmental damage, benefits to future generations, wastefulness, or some other concern specific to the individual.

ISM requires an energy (thermal) model of the building in order to predict accurate savings based on the building and its use, rather than on inaccurate typical values. To be cost effective, such a model would need to be automatically assembled from a minimum of sensor information so as not to overly increase the cost of the smart meter, for example a mix of utility meter data, room temperature, sub-circuit or high-frequency electricity data to infer occupation and maybe home CO2 concentration (as a proxy for ventilation rate). The sensors needed and the accuracy of such an
approach, which uses inverse modelling to obtain an accurate thermal model from a time series of data, have been reported by the authors in [14] and [15]. By creating such a model, the financial, and other, impact of any suggestions, for example turning down the thermostat by one degree, can be calculated for the specific home and reported to the user. In addition, inappropriate suggestions, for example suggesting a reduction in heated temperature in an already under-heated home, can be avoided.

One can imagine other advantages of the ISM approach, for example, the reporting of the presence of high U-Values in the fabric to utility companies for targeted intervention.

This paper attempts to discover if this new approach can be applied to a group of homes and if it generates changes in behaviour.

2 Background

Some research findings [16] suggest that continuous energy feedback might be an effective driver of energy-related behaviour change. For example, Barbu et al. [17] suggest that energy feedback provided to users via smart meters could save 5-15% of total energy costs. Similarly, [18, 19, 20] suggest that energy feedback through advanced in-home displays (IHDs) could help to save up to 20% of energy costs, either for electricity or total energy bills. However, reality seems to fall short of such predictions. For example, in [21] a more modest average energy saving of 7% is reported across multiple utility pilot programs aimed at electrical energy conservation with the help of IHDs.

Current technological solutions for real time energy feedback suffer from multiple issues [22, 23], for example: unengaged users; failure to address users’ personal motivations and needs embedded in daily routines and social practices;
information comprehension issues caused by abstract numerical information in kWh or financial costs; and inattention to users’ personal characteristics [24]. It seems clear that users need something more to motivate and engage them than plain energy feedback in kWh or cost if we are to get energy reductions of 10% or more. Some research indicates that intelligent energy feedback that offers different feedback options might be effective [25].

Chiang et al. [26] have calculated that smart meters can pay back their installation costs in 4 years if the energy savings are 3% or more, although the observation that the utility company no longer need to visit the home to read the meter might currently be the main economic driving force for their installation. It is hence interesting to ask whether we might move the focus from a device that is mainly of use to the energy company to one that has equal utility to the occupant.

The question of what best motivates householders has received considerable attention from energy behaviour researchers, with varying success being noted. For example, attempts to stimulate savings through social comparison and competition have not achieved notable success [25, 27, 28, 29]. One possible reason for this is the so-called boomerang effect [30], whereby when households are told they are using less energy than average they start to use more when they see what is ‘normal’ or ‘permissible’. This suggests that if messages are to be sent to households, only those where there is some evidence of potential energy profligacy should be targeted with certain messages, and this is a key aspect of our study, and the one reason for its success.

The other key aspect is people’s values. Different disciplines, including economics, psychology, philosophy, sociology and anthropology, have all tried to
understand the role of personal values in conservation behaviour. From an economic point of view, people’s values correspond to long-term preferences, and can be explained by decision theory [31]. In anthropology, values are ‘cultural worldviews’ and their role in pro-environmental behaviour is studied within climate change risk perception [32]. In social psychology, Schwartz [33] identified a number of personal values universal for all cultures and nations. From this perspective, personal values are defined as superordinate goals that serve as enduring guiding principles in peoples’ lives [33]. Common to all these approaches is the idea that values are conceptually different from goals, opinions and attitudes: values reflect broad long-term preferences and so provide unity across a broad range of behaviours (a person who above all values their own self-interest will be self-interested in most settings); goals, opinions and attitudes, on the other hand, are much more situation-specific and changeable over time [34]. In the current study, we test personal values as motivators to energy-saving behaviour and focus on altruistic, egoistic and biospheric values, representing, respectively, concern for others, the self and the natural environment. These values were chosen from the various Schwartz’s personal value sets based on research by DeGroot and Steg [35, 36], who identified altruistic, egoistic, biospheric and hedonic (pleasure-seeking) values as the key value orientations highly correlated with pro-environmental behaviour. In this study, hedonic values were dropped from this set after pilot testing (with an opportunistic sample of 30 UK adults) found much lower engagement with such messages compared to the other three value orientations (which pilot participants found useful, and more engaging than energy messages expressed in standard kWh).

The final facet of this study concerned the nature of the action prompts. The
literature hints that energy feedback may be effective when it prompts concrete personalised actions (for example, “Please switch off unused appliances”, “Adjust your thermostat setting”), presumably because these do not require knowledge or problem-solving from the householder, as would a more general action prompt like “Reduce your energy consumption”. However, at the moment we lack the empirical quantitative data needed to substantiate these claims [24]. Research on the effect of general action prompts has shown that it has limited influence on behaviour, but with increased possibilities for specificity and context detection offered by modern technologies it might substantially improve the effect of feedback [37]. Some findings also indicate that for the energy information to be useful, it should be tailored to personal contexts [38]. Thus, specific action prompts, tailored to specific households, might have a more profound effect on energy related behaviour, and this idea is tested here.

Providing information on energy related actions and personal context is directly related to the educational component of energy feedback and the concept of energy literacy—which implies an in-depth understanding of energy consumption [39] and is frequently mentioned in pedagogical, educational and environmental policy literature. Energy literacy has been called “a broad term encompassing content knowledge as well as a citizenship understanding of energy that includes affective and behavioural aspects” [40], although arguably such a definition conflates knowledge about where energy is consumed with the motivation to reduce this. Cognitive (knowledge, in-depth understanding), affective (attitudes, values), and behavioural (social practices) modules can be distinguished within a complex concept of energy literacy. It is known that promoting energy literacy can foster a shift in knowledge and perception of energy and thus will facilitate responsible energy related choices and behaviour [40, 41, 42].
To the authors’ knowledge, there are no studies or commercially available energy feedback systems that use action prompts tailored to users’ behaviour in a systematic manner. The only work that provides some tailored heating action prompts is our previous study [43], however we did not aim to test different energy feedback approaches in an experimental setting. In the work of D’Oca et al. [44] tailored to users’ contexts, newsletters with action suggestions were sent via email to participants, though the study only addressed electricity consumption, and action prompts were not part of an energy feedback display. One work on energy consumption feedback in a non-residential sector reports the effectiveness of user-tailored context advice and action prompts [45], though this study again was focused only on electricity consumption and the tailored advice was sent by email not a smart meter.

There are a lot of commercially available devices and applications that provide people with simply energy feedback and give general, not occupant- or building-specific tips and suggestions regarding energy saving issues. In other words, the existing IHDs are not smart enough to address building context or personal motivations.

3 Pilot study and research hypotheses

Very few studies test the effect of internal personal values embedded in a digital energy feedback within field experiments. One exception being the previously mentioned work of Schultz et al [29] which compared feedback based on normative comparison and financial costs, against a standard feedback showing consumption alone. We hypothesised that couching energy feedback messages in terms of several of the personal values cited in the environmental psychology literature would substantially improve users’ engagement and increase their motivation to save energy
since multiple values guide people’s lives [34] and focussing only on one value might be not effective to increase motivations.

In addition to this, we have included the facet of personalised feedback i.e. messages that refer, and are valid only for, specific households.

The main research questions that we wanted to answer were:

1. What is the overall effect of digital energy feedback interventions?
2. What is the additional benefit of internal values embedded in energy feedback messages?
3. What is the additional effect of personalised messages with action prompts?
4. What effect has a digital energy feedback on cognitive variables, such as energy literacy and users’ experience with a digital feedback?

To answer these questions, we formulated the following hypotheses:

- *Hypothesis 1*: Digital feedback in general will have an effect on energy related behaviour compared to a baseline period before the interventions.

- *Hypothesis 2*: Energy feedback translated according to internal values will have a significant effect on energy related behaviour.

- *Hypothesis 3*: Personalised messages with action prompts will have a significant effect on energy related behaviour.

- *Hypothesis 4*: No matter which behavioural effect of a digital energy feedback is observed, the digital interventions will influence energy related cognitive
variables, such as energy literacy.

4 Method

To answer the research questions defined in Section 3, we designed an energy feedback system with different types of feedback, we deployed the system in homes in the UK and performed a field experiment. The experiment had a 2 (Value framed vs. Non-value framed) by 2 (No action prompt vs. Tailored message with action prompt) within-subject factorial design resulting in four energy feedback experimental conditions (see Table 1).

4.1 Participants and procedure

Our participants were residents of social housing recruited in the first half of 2013 with the help of Exeter City Council. Initially, an information package was sent to the households and later recruiters went door-to-door to identify interested households. After a consent form was signed by the occupants, the sensors were installed during January-May 2014 in 73 homes. The first phase of the project established the baseline: sensor data were collected but no energy feedback was provided to the participants.

For the second phase of the project, an energy feedback experiment was conducted over three months (January 2016 - March 2016) with the help of a new energy feedback application (iBert) written by the authors and presented to participants via a tablet computer. The current paper describes the analysis of this second phase of the project. Due to the length of the study, and the social group involved, some households dropped out of the study, moved houses or changed contact details, which resulted in 43 homes participating for the full study period.
In this sample, each household had on average 2.8 residents. Each home had a main contact, and the mean age of these was 50.6 years (with a minimum age of 27 and a maximum of 81), 14 were male and 29 female. Seventy-five percent of households reported an income lower than the UK national median of £25,600 (identified at a national UK level for 2015, with the London area excluded [46]). Around 25% of households reported difficulties with paying energy bills. Fourteen percent of households already had a basic smart meter (could correspond to condition 1 in our study).

The focus of the study was in part the response of occupants to feedback from personalised messages. For this reason, individual basic thermal models of each of the homes were created. It had already been demonstrated by the authors that the sensors used can auto-generate a suitable dynamic thermal model of each building based only on the data received from the sensors (without the need for information on the size of the homes or their constructions). This model is a lumped parameter model and the work has been reported in another paper [15] and shown to produce accurate results. However, to make sure that no extra noise was added to the experiment and ensure the results reported were not affected by an auto-generated model which might well change during the experimental cycle, simple (steady-state) models were created by using floor plans and information about the materials and constructions from the city council and visual inspection. These models consisted of calculating the mean heat loss coefficient of each building. In future developments auto-generating models like those of [15] could easily be used.

The simple model obtained from the building geometry was used by iBert to investigate the energy savings of various strategies (turning the thermostat down,
closing windows, turning the heating off when the building was not occupied, turning
the lights and some other electrical items off when the building was not occupied) that
could be applied to each house if the household breached set action levels of: room
temperature, ventilation rate, having the heating on when building was unoccupied;
and high electricity use when the building was unoccupied. Depending on the
condition applied to the home at the time (see Table 1 and 2), this gave iBert the ability
to inform households of: their current energy consumption; the total consumption over
the last week; their consumption expressed in accordance to other values (such as an
environmental cost); the savings that might occur if they took action; and, most
importantly, the specific actions they might take.

Due to page limitations, full details of the electronics, sensors and coding that
lie behind iBert will be described in a later paper. But in outline, the system consisted
of sensors that measured real-time utility data, air temperature (in up to three rooms),
radiator temperature (to sense if heating was on), humidity (in up to three rooms),
light, motion and CO₂ levels, with this data collected approximately every 5 minutes
via Raspberry Pi and transmitted via the home’s broadband or 3G network to a secure
cloud store. These data were then used to see if action levels had been triggered,
calculate any savings that might be made if behaviour were to change in the future, and
report back to the occupant via the same connection to the Wi-Fi enabled tablet in the
home as a weekly digest. The frequency with which householders interacted with the
iBert app was recorded directly from each tablet and as well as a record of the tailored
textual messages sent to each house.

Table 1: Energy feedback study design (examples of information presented within
each condition in italics – other non-monetary values, for example environmental
impact, were also used.

<table>
<thead>
<tr>
<th>No Action No personalised</th>
<th>No Value</th>
<th>Value:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C1: Standard display (kWh and financial costs)</td>
<td>C2: Standard display + values</td>
</tr>
<tr>
<td></td>
<td><em>Your home has used 95 kWh this week.</em></td>
<td><em>Your home has used 95 kWh this week; this is equivalent to £12.</em></td>
</tr>
<tr>
<td></td>
<td>C3: Standard display + + tailored action prompts</td>
<td>C4: Standard display + + tailored action prompts + values</td>
</tr>
<tr>
<td></td>
<td><em>If you reduced the thermostat temperature in your house one degree you would save 11 kWh.</em></td>
<td><em>If you reduced the thermostat temperature in your house one degree you would save 11 kWh; this is equivalent to £1.43.</em></td>
</tr>
</tbody>
</table>

The households were cycled through the four conditions shown in Table 1 with each condition lasting three weeks. Households were randomly assigned to one of the four conditions according to the Latin-Square counterbalancing design of Table 2. For example, during the first three weeks, one-quarter of the homes were under condition C1 of Table 1 and in the following three weeks they had condition C2 applied to them.

In conditions C3 and C4, tailored energy feedback messages are generated once per week, based on energy-related information aggregated over the previous week, and were displayed during the subsequent week. If the algorithm for identifying energy-wasting events and variables did not detect any potentially profligate energy use in a particular household, that household did not receive any tailored messages that week.

Table 2: Latin square counterbalancing design for the four experimental conditions: C1 to C4 across the four temporal phases of the experiment.

<table>
<thead>
<tr>
<th>Baseline</th>
<th>Phase 1</th>
<th>Phase 2</th>
<th>Phase 3</th>
<th>Phase 4</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>14/12/15</td>
<td>04/01/16</td>
<td>25/01/16</td>
<td>15/02/16</td>
<td>07/03/16</td>
<td>28/03/16</td>
</tr>
</tbody>
</table>
This results in sixteen textual messages in total: four for personal value neutral condition C3 and twelve for value-framed condition C4 (egoistic, altruistic and biospheric values). A household could receive up to four different messages since a value message is framed according to one of the three personal values and a value is selected at random. Messages in conditions C3 and C4 are sent only if there are energy-wasting events identified by the intelligent algorithms.

Participants received an incentive of £150 in the form of supermarket vouchers, sent by post in three instalments of £50 (at the beginning of the project, at the end of year one and at the end of year two after the completion of the project). In addition, they were allowed to keep the tablet once the study has ended. All participants who completed a post-study phone survey received an additional incentive with a chance to win a £20 supermarket voucher during a prize draw. After the end of the experiment the energy feedback app was deactivated and sensor equipment was collected from the participating homes four months after the end of the energy feedback interventions.

Prior to both phases of the study participants were given paper information sheets, consent forms and surveys that contained among others demographics and the first seven items from the energy literacy survey described in [40]. This is an updated version of a validated questionnaire and energy literacy assessment framework developed in the by DeWaters and Powers [40]. To the authors’ knowledge, it is the only validated questionnaire that evaluates energy literacy according to a broader
definition. The seven items used from this questionnaire addressed only the general knowledge of energy consumption. (The full energy literacy and user experience survey used for this work can be found in Appendix A).

As previously mentioned the interactions with iBert were also monitored—Figure 1. The reader can see that almost all houses had substantial level of interactions with the tablet. Each interaction with the app consisted of a minimum of three recorded app related events, such as ‘app started’, ‘app resumed’, ‘app closed’ etc.

Figure 1: Interactions heat map—the colour indicates the number of interactions. The House ID is the ID of the house within the database.

4.2 The four experimental conditions

**C1: Standard energy feedback.** For the standard display condition (C1: No value – No action; see Table 1), we followed the recommendations of the minimal
requirements for an In-Home Display (IHD), i.e. we replicated a typical smart meter in the UK [47]. This displayed: current gas and electricity consumption; cumulative consumption for the past week; information in kWh; and a visual (i.e. non-numerical) presentation alongside a numeric one.

**C2: Standard feedback with personal values.** In a personal value only condition (C2: Value - No action) the display was the same as in C1, but complemented the weekly summary information translated into one of the three internal values: *egoistic* – expressed in money; *biospheric* – expressed in trees destroyed; or *altruistic* – expressed as the cost to society, couched as the number of minutes with a family doctor that could be provided at the same cost. So, for example, a household might have the kWh display replaced with information that their energy use in the last week was equivalent in carbon terms as the destruction of 1.7 trees or 20 minutes with a doctor. Egoistic, biospheric or altruistic personal values for each household were chosen randomly.

**C3: Standard feedback with tailored actions.** This condition was the same as C1 except households also saw a tailored message box with action prompts in text form about any potentially energy profligate behaviour (room temperatures > 21°C, occupied CO$_2$ levels of <800 ppm (possibly indicating excessive ventilation), mean unoccupied electricity consumption in excess of 600W, or heating on when the house was unoccupied) during the previous week. Action prompts described concrete actions with low personal costs, e.g. lowering a thermostat setting or leaving windows open less frequently. A participant could receive from zero to a maximum of four messages. Messages consisted of four parts: (1) communication of factual information regarding
energy related variables [for example, we have noticed that your thermostat was set to 23°C]; (2) evaluation of the factual information [e.g. this value is unnecessarily high]; (3) a motivational component which expressed wasted energy information aggregated over a longer time period (over a whole winter, over a year) [e.g. in your house this may imply 520 kWh use more per winter]; and (4) an advice part which contained an action prompt [e.g. we recommend that you lower the thermostat to 21 degrees centigrade and that you check if you are comfortable at that temperature].

The trigger of 21°C was based on national recommendations [48]. The 800 ppm was chosen by considering EN 15251. This suggests IDA 2 air quality (medium IAQ) is achieved below 1000 ppm. In an effort to be precautionary, a lower value of 800 ppm was used - this was the median value found during the baseline phase. The 600 W was also derived from data gathered during the baseline phase. This showed that in all houses consumption was bi-modal: low values during unoccupied periods or the night; and high values during occupied periods. The mean figure that separated the two being 600W.

**C4: Standard feedback with tailored actions and values.** In experimental condition C4, households received the same information as in C3, but the kWh part of the message was translated into one of the three internal personal values just like in C2. An example of an energy feedback message in condition C4 where energy information is translated into biospheric value concepts might be: “I think a lot of heat might be escaping from open windows. The escaping heat results in wasted energy. This requires roughly 12 more trees to compensate for the extra pollution caused by your home over a winter. This may be because there are too many windows open, they are open too wide or for too long. Try changing how many windows you open and for
how long and you may save energy that over a whole winter is equivalent to planting 12 trees. Try it, and check if this message disappears next week”.

Appendix B shows a screen shot of ibert as seen by a household.

5 Results

In this section, we show the results from the different analyses that were performed to evaluate the effectiveness of the iBert system.

Analysis 1: Overall effect of iBert

This analysis tests hypothesis 1, asking if digital feedback in general has an effect on energy-related behaviour compared to a baseline period before the interventions. The purpose of this is to examine if our findings are in line with those from other EHD trials. A repeated-measures ANOVA was used to evaluate the effect of iBert on internal temperature, electricity consumption and ventilation rate (CO₂ levels).

We found a significant effect of Time (i.e. study phase) on home internal temperature, $F(2,58) = 3.78, p = 0.029, \eta^2_p = 0.12$. Post-hoc pairwise comparisons using Holm-Bonferroni $p$ value correction reveal that home internal temperature before iBert was active is significantly higher than the internal temperature during the iBert experimental phase ($d = 0.24, p < 0.001$) and the difference between the internal temperature during the experiment and after the experiment is not significant either ($p = 0.97$)—indicating the intervention had lasting impact (Table 3). We understand that lowering the temperature can lead to larger issues such as comfort problems or even health risks in vulnerable households, it is for this reason that this analysis is expanded later to check if this temperature reduction was from all households, or only from ones
with initially high temperatures..

Table 3: Mean internal temperature (in °C) for households before, during and after the digital iBert energy feedback interventions.

<table>
<thead>
<tr>
<th>Temperature (°C)</th>
<th>Standard deviation (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>20.4</td>
</tr>
<tr>
<td>iBert Phase</td>
<td>19.6</td>
</tr>
<tr>
<td>After</td>
<td>19.4</td>
</tr>
</tbody>
</table>

In addition, we conducted two analyses of covariance to see how often occupants interacted with iBert and if occupants receiving tailored textual feedback influenced these results. Frequency of interaction with iBert proved not to be statistically significant. However, change in internal temperature was predicted by whether householders had received at least one tailored textual feedback message, $F(2,56) = 3.45, p = 0.04, \eta^2_p = 0.11$. This interaction shows that the difference in internal temperature between homes with and without a tailored textual feedback is not visible during the digital interventions phase, but it is after the interventions. In the light of these results, we conducted a follow up $t$-test on the differences in internal temperature between with-iBert and after-iBert phases for two groups of households: those who did not receive any temperature-related tailored textual message and those who did.

This difference is significant, $t(29) = 2.29, p = 0.03$, which indicates that homes that received tailored textual messages with action prompts had lowered their internal temperature after the digital interventions; while homes which did not receive messages, did not decrease their internal temperature.
The overall effect of iBert on electricity consumption and CO\textsubscript{2} concentration, with frequency of app usage and received textual messages as covariates, would appear not to be statistically significant.

**Analysis 2: Post-intervention effect of messages**

Temperature-related action prompts were triggered in homes over 21°C. To assess whether people responded to these, we conducted four follow-up one-sample t-tests to compare internal temperatures to 21°C, before and after the intervention, for the group who received temperature-related action prompts and the group who did not. For the non-message group (the subgroup whose home temperature was lower than 21°C and who therefore were never told to change the temperature even under experimental conditions C3 and C4), internal temperature was significantly lower than 21°C both before (mean temperature = 18.7°C, SD = 1.6, \( t(24) = -6.99, p < 0.001 \)) and after the experiment (mean temperature = 18.7, SD = 1.9, \( t(16) = -5.09, p < 0.001 \)) – as would be expected. The group that received the temperature message because their baseline temperature was over 21°C (mean temperature = 22.3°C, SD = 1.1, \( t(14) = 4.38, p <0.001 \)), however, had internal temperatures that essentially settled at 21°C after the messages (mean temperature = 21.1°C, SD = 2.00, \( t(11) = 0.20, p = 0.80 \)), Figure 3. In summary, people who received temperature-related action prompts saw a reduction in internal temperatures that put their home temperatures almost exactly at the level that was suggested; a similar change was not seen in those who received no prompts, suggesting the effect is likely not the result of weather-related changes. Note that both groups could still see the standard IHD module of the app after the experiment, although the tailored messages function was deactivated. No messages were triggered
based directly on gas consumption (but rather on internal temperature, heating times and excessive ventilation), however the use of iBert reduced the mean household gas consumption by 14 to 29% with a mean of 22.0%.

**Analysis 3: Effect of values and action prompts**

To evaluate our second and third hypotheses, regarding the effect of internal values framing and tailored action prompts during the energy feedback iBert phase, we conducted a two-way repeated measures ANOVA on home internal temperature, electricity consumption and CO₂ level data with the factors Personal Values (present or absent) and Action Prompts (present or absent). Baseline measurements of these dependent variables, taken before iBert deployment, were used as covariates. After controlling for the baseline temperature, homes in the value framing conditions (C3 and C4) have a lower internal temperature (19.23°C) compared to homes in the conditions without value framing (C1 and C2) (19.32°C).

An ANOVA on these internal temperature readings during the iBert phase, with baseline temperature as a covariate, revealed a significant effect for Value Framing, $F(1,33) = 5.29$, $p = 0.028$, $\eta^2_p = 0.14$. The baseline temperature covariate was also significant, $F(1,33) = 4.94$, $p = 0.033$, $\eta^2_p = 0.13$, suggesting the warmer homes at the start of the study still tended be the warmer homes at the end.

As shown later, iBert had a marked impact on energy literacy. Hence it is unlikely that action prompts were forgotten between experimental phases. In light of this, the data were separated into two groups, before the issuing of action prompts and post issuing. Action prompts were found to have a statistically significant impact under a pairwise t-test ($p = 0.021$) on the heating degree day adjusted mean gas consumption
(mean before action prompts = 1239W, SD = 379; after 1120W, SD = 348), contributing 9.6 percentage points to the overall reduction in gas consumption.

In contrast to internal temperature and gas consumption, there were no effects of Actions or Values on electricity consumption and CO₂ concentration.

**Analysis 4: Case studies of subgroups with high and low energy consumptions**

To get more insight into the changes in internal temperature, gas and electricity consumption, and CO₂ concentration, across all the stages of the study, we examined the box plots of the data for two subgroups: one being the households that had high energy use and the other those homes with low energy use.

Figure 2 shows that the median temperature of the houses that had a high thermostat temperature (upper graph) went down during the iBert interventions and stayed low after the study.
Figure 2: Home internal temperatures (in °C) before, across the iBert digital intervention phase and after the feedback. The graph has been separated as houses likely to receive a personalised message (upper graph) and those that did not (lower). The blue box represents 50% of the data, the median is represented as a horizontal (red) line within the box and the blue whiskers represent the 97.5% of the data. The thick green line is at 21°C—the trigger for receiving a message. The individual dots, and dashed lines, represent individual homes.

The temperatures before the intervention were shown to be statistically different to 21°C in the group with temperatures higher than 21°C in the baseline: with a p-value rejecting the Null hypothesis (they are equal) of p=0.0039 (<0.05); whereas the temperatures before and after in the group in which the temperatures were already
below 21ºC cannot be considered different (p-value of 0.46). Moreover, a t-test was used to prove if the hypothesis that the temperatures in the baseline in the houses in which the starting temperature was larger than 21ºC was statistically different to 21ºC after the study. The test proved that the temperatures of the group which behaviour needed to be modified were not statistically different to 21ºC after the study with a p-value of 0.17 (the null hypothesis mean = 21ºC, can not be rejected) meaning that the behaviour of the occupants had been modified.

A similar evaluation was done with the CO₂ concentration. Houses in which the CO₂ levels were less than 800 ppm were considered as potentially over ventilated, and therefore it was assumed that some energy savings could be achieved by reducing ventilation.

To evaluate this, we again separated the homes into two groups: the ones that showed high levels of ventilation in the baseline (CO₂<800ppm) and those that did not (Figure 3). In this case, all the tests show that the CO₂ concentrations were not changed in the intervention; neither in the homes that showed high levels of CO₂ nor in those that showed low levels. This is possibly because changing the ventilation regime in a house is not a trivial task. Also, it should be noted that CO₂ is not only an indicator of ventilation, but also an indicator of infiltration occurring through cracks and other defects, and that is not under the control of the occupants.
Figure 3: Home CO₂ concentrations before, across the iBert digital intervention phase and after the feedback. The graph has been separated into houses likely to receive a personalised message (upper graph) and those that were not (lower). The thick green line is at 800ppm—the trigger for receiving a message.

With respect to electricity, it also seems that the intervention failed to change the behavioural habits of the occupants, with no statistically significant change seen in the behaviour of big consumers shown in the top part of Figure 4.

The gas consumption however shows a clear indicator of the effect of iBert. Although iBert contained no direct messaging with respect to gas consumption, three statements within iBert were designed to reduce gas consumption: change in the thermostat temperature, change in the operating scheme (turning heating off when house unoccupied) and change in the ventilation regime. As there was no direct
messaging with respect to gas consumption, we cannot separate high and low consumers based on whether they received a message. A reasonable alternative is to simply classify those in the top quartile of consumption as high consumers. Calculating the pre/post iBert household gas consumption for just these high consumers shows a mean household saving of 27.2%, compared to 22.0% for all homes.

It is unknown why there was no reduction in electricity use, but the following are possibilities: (1) for an occupant, reducing electricity consumption is more complex than simply turning a thermostat down, or ensuring heating is off when leaving the building; (2) rational electricity consumption is more individualistic than heating consumption, with a multitude of different devices in different homes, this could lead to high consumers not necessarily being profligate; (3) it is known that people are unsure of which electrical items use the most electricity, and hence which to manage more effectively.
Figure 4: Box plots of daily electricity usage after separation of big consumers (top) and small consumers (bottom). The thick green line is at 600 W—the trigger for receiving a message.

**Analysis 5: Energy Literacy.**

A paired sample t-test was conducted to evaluate the effect of digital energy feedback interventions on the cognitive component of energy literacy. The results indicated that the mean energy literacy score before the interventions (M= 0.52, SD= 0.71) was significantly lower than the mean energy literacy score after the interventions (M = 1.28, SD =1.06, t(24) = 3.17, p = 0.004). This confirms our fourth hypothesis that the energy literacy of participants will improve after digital feedback that is framed in an educational way.
Analysis 6: User experience and engagement

Three months after the end of the energy feedback study, participants were contacted by phone to arrange for the equipment to be collected and were asked to complete the same items from the energy literacy survey described in [50,51], again along with nine questions related to system usability and user experience: four items adopted from the System Usability Scale (SUS)[52]) which is the most frequently cited scale for system usability evaluation, one item addressing participants’ preferences for different components of iBert and four items on the effect of iBert on participants’ behaviour. In total, 65% of the participants completed the post-study survey. The answers were rated on a 3-point ordinal Likert scale. The usability questions were:

“I think that I would like to use iBert frequently”

“I found the system very difficult to use”

“I needed to learn a lot of things before I could get going with the system”

“How useful did you find iBert overall?”

The results show that the majority of the participants liked to use iBert frequently, found the system easy to use and in general found iBert useful. Eighty-two percent of the respondents found live energy consumption data useful, 77% liked the weekly energy summary and 78% the tailored messages; ANOVA analysis showed the three components of feedback were found roughly equally useful, $F(2, 32) = 1.27$, $p = 0.30$. Some 50% of the respondents took energy saving actions suggested by iBert and roughly 36% of them took other energy saving actions not suggested by the system. User experience with iBert is summarised in Table 4.
Table 4: Self-reported user experience with iBert energy feedback.

<table>
<thead>
<tr>
<th>Interaction</th>
<th>Percentage of households</th>
</tr>
</thead>
<tbody>
<tr>
<td>Took energy saving actions suggested by iBert</td>
<td>50%</td>
</tr>
<tr>
<td>Took additional energy saving actions NOT suggested by iBert</td>
<td>36%</td>
</tr>
<tr>
<td>Paid attention to iBert text messages</td>
<td>78%</td>
</tr>
<tr>
<td>Other households members were aware of the iBert information</td>
<td>86%</td>
</tr>
<tr>
<td>Found iBert (partially) useful</td>
<td>82%</td>
</tr>
</tbody>
</table>

6 Discussion

In this section we discuss the findings in the light of the four hypotheses formulated prior to the study.

**Hypothesis 1:** Digital feedback in general will have an effect on energy related behaviour compared to a baseline period before the interventions.

This hypothesis was confirmed. Homes with high internal temperatures reduced their temperatures and gas consumption was reduced.

**Hypothesis 2:** Energy feedback translated according to internal personal values will have a significant effect on energy related behaviour.

This hypothesis was confirmed for home internal temperature measurements. Framing messages in value-based terms led to greater changes in behaviour.

**Hypothesis 3:** Personalised messages with action prompts will have a significant effect on energy related behaviour.

This hypothesis was confirmed, with an additional 9.6 percentage point reduction in gas consumption from the use of action prompts.
**Hypothesis 4:** No matter which behavioural effect of a digital energy feedback is observed, the digital interventions will influence energy related cognitive variables, such as energy literacy.

This hypothesis was confirmed. We found a statistically significant positive effect of iBert interventions on energy literacy. One may argue that the effect of energy literacy could be attributed to other factors and not the tailored educative messages *per se*, e.g. by the Hawthorne effect [53]. This is a phenomenon where people behave differently when they know they are being observed. Given the fact that the sensors sets were installed in these homes two years earlier and they had been monitored for a long period of time prior to the energy feedback intervention study, we believe that the Hawthorne effect was minimal in comparison to other studies as many of them did not control for this effect.

7 **Conclusions**

The main innovative aspect of this study is that it tests for the first time a combination of two energy feedback strategies, personal values and tailored action prompts, applied to energy related behaviour in an experimental setting. Value-framed motivational messages and action prompts tailored to user behaviour and building characteristics were embedded into an intelligent energy feedback system called iBert. The system gives advice according to individual building context and problematic energy-related behaviour detection, such as high internal temperature, excessive ventilation, and heating and electricity usage while the home is not occupied.

Our findings suggest a positive effect of digital feedback on home internal temperature, and a specific positive effect of internal values and action prompts incorporated in energy feedback, from digital energy feedback delivered by iBert. This
is the first time a smart metering approach has been shown to directly influence room temperature. Unprecedented gas consumption saving were achieved with much of this directly attributable to the use of action prompts.

The 22.0% reduction in gas consumption (27.2% in high consumers) found is substantially greater than the reductions seen in the literature; this provides strong evidence of the power of personalised feedback, personal values, action prompts and an active improvement of energy literacy, and could be industry-transformative.

The digital feedback system we have demonstrated in this paper had a marked positive effect on the knowledge component of participants’ energy literacy. This result is very promising as it indicates that digital feedback (a standard IHD accompanied by tailored educational messages as in the current study) can improve general energy understanding. A better understanding of energy use would be a natural first step towards forming new attitudes and foundations for energy-saving behaviour amongst occupants.

In this paper we have evaluated a ground-breaking way of informing occupants about their energy consumption with the purpose of educating users and lowering their demand. By being building- and occupant-aware, and only reporting messages when consumption is potentially high, many of the criticisms of smart metering have been addressed. For the first time a smart metering system has been shown to reduce room temperatures directly, with the temperature chosen by the occupants identical to that suggested by the algorithm. The gas consumption savings found in this two-year study have the potential to be transformative for both the occupants and for climate change policy.
Acknowledgements

This work was done under the auspices of the ENLITEN project, funded by EPSRC (grant reference EP/K002724/1). We thank Exeter City Council for their support provided during the project. The internal temperature and energy consumption data used for the analysis together with the energy literacy and usability surveys data can be found here: http://doi.org/10.15125/BATH-00256. Ramallo-González would like to thank the program Saavedra Fajardo (20035/SF/16) funded by Consejería de Educación y Universidades of CARM, via Fundación Séneca-Agencia de Ciencia y Tecnología de la Región de Murcia.

References


Appendix A   Energy literacy and user experience survey

Part 1

1. How much do you feel you know about energy? (Please circle one)
A lot (expert)

Quite a bit (informed)

Not much (novice)

Nothing

2. Which of the following sources of information has contributed most to your understanding of energy issues? (Please choose one option)

   Further or higher education

   School

   Books, newspapers or magazines

   Friends or family members (including parents)

   Internet

   Television iBert system

   Other (please specify)

3. The term renewable energy resources means (please circle one answer)

   Resources that are free and convenient to use

   Resources that can be converted directly into heat and electricity

   Resources that do not produce air pollution

   Resources that are very efficient to use for producing energy

   Resources that can be replenished by nature in a short period of time

4. Most of the renewable energy in the UK comes from which of the
following sources? (Please circle one)

Solar

Water (hydro/tidal/wave) power

Wind

Landfill gas

Geothermal

Don’t know

5. Which of the following actions, if everyone did this all the time, would save the most energy in the UK? (Please circle the most important)

Turn off lights when they are not in use

Turn down the heat in rooms

Reduce water consumption

Walk or cycle short distances instead of going by car

Turn appliances off at the plug

6. Which kind of lighting uses the least amount of energy? (Please circle one answer)

Standard light bulbs

Low energy light bulbs

Fluorescent lights

LED lights

Don’t know
Part 2

1. I think that I would like to use iBert frequently
   
   Agree
   
   Somewhat agree
   
   Disagree

2. I found the system very difficult to use
   
   Agree
   
   Somewhat agree
   
   Disagree

3. I needed to learn a lot of things before I could get going with the system
   
   Agree
   
   Somewhat agree
   
   Disagree

4. In which extent did you find the following parts of the iBert system useful:
   
   Live status panel
   
   Useless
   
   Moderately useful
   
   Very useful
   
   Weekly energy consumption summary
   
   Useless
Moderately useful

Very useful

Text messages with tips and advice/suggestions

Useless

Moderately useful Very useful

5. How useful did you find iBert overall?

Useless

Moderately useful

Very useful

6. Did you pay attention to the text messages that were displayed by iBert?

Yes

No

Partially

7. Did you take any energy saving actions suggested by iBert?

Yes

No

8. Did you take any additional energy saving actions NOT suggested by iBert during or after the study?

Yes

No
Partially

9. If there are other adults living in your home, did they attend/were aware of the information provided by iBert?

Not aware

Aware

Partially aware

Appendix B    iBert running on an Android tablet
1. We have noticed that the temperature in your house is usually 25. This is very warm and uses more of our limited energy resources. Lowering the temperature by 1 degree could help to keep energy prices lower for everyone in future. If you still feel cold, would you consider wearing warmer clothes too?

2. We think the heating was on when there was no one in the house some time last week. This uses more of our limited energy resources. Remember to turn the heating off when you leave the house – it could help to keep energy prices lower for everyone in future.

3. We have noticed that your electricity use was higher than expected when there was no activity in the house. This uses more of our limited energy resources. Make sure you turn off thing like lights, TVs, and other electronic devices when you leave the house - it could help to keep energy prices lower for everyone in future.

---

Live Status

<table>
<thead>
<tr>
<th>Energy Use</th>
<th>Living Room</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electricity</td>
<td>Temperature</td>
</tr>
<tr>
<td>£0.01</td>
<td>22°C</td>
</tr>
<tr>
<td>Gas</td>
<td></td>
</tr>
<tr>
<td>£0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CO₂</th>
<th>Kitchen</th>
</tr>
</thead>
<tbody>
<tr>
<td>672PPM</td>
<td>Temperature</td>
</tr>
<tr>
<td></td>
<td>24°C</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bedroom</th>
<th>Radiator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>Humidity</td>
</tr>
<tr>
<td>19°C</td>
<td>36%</td>
</tr>
</tbody>
</table>