

Real-time feedback promotes energy conservation in the absence of volunteer selection bias and monetary incentives

Verena Tiefenbeck^{1,2*}, Anselma Wörner¹, Samuel Schöb³, Elgar Fleisch^{1,4} and Thorsten Staake³

Feedback interventions have proved to be effective at promoting energy conservation behaviour, and digital technologies have the potential to make interventions more powerful and scalable. In particular, real-time feedback on a specific, energy-intensive activity may induce considerable behaviour change and savings. Yet the majority of feedback studies that report large effects are conducted with opt-in samples of individuals who volunteer to participate. Here we show that real-time feedback on resource consumption during showering induces substantial energy conservation in an uninformed sample of guests at 6 hotels (265 rooms, $N = 19,596$ observations). The treatment effects are large (11.4% reduction in energy use), indicating that the real-time feedback induced substantial energy conservation among participants who did not opt in, and in a context where participants were not financially responsible for energy costs. We thus provide empirical evidence for real-time feedback as a scalable and cost-efficient policy instrument for fostering resource conservation among the broader public.

Individuals' choices and behaviour are a key lever influencing energy consumption, along with the technical energy efficiency of the products and infrastructure used¹. To tackle environmental challenges, it is important to put people at the centre of energy research, and to empirically validate how to promote sustainable decision-making among individual consumers. Energy consumption is a low-involvement topic for most people; many consumers are unaware of the energy efficiency of their homes and devices², or underestimate the long-term benefits of potential investments into energy efficiency upgrades³.

As digitization advances, it becomes increasingly feasible to monitor the energy consumption of households, specific appliances or activities in real time⁴. As a result, digitally enabled behavioural interventions can be deployed at population scale and become more powerful through personalization and context specificity. Beyond that, it becomes increasingly possible to systematically evaluate the impact of behavioural interventions with large and diverse samples of participants. The availability of high-resolution consumption data enables more and more personalized and flexible interventions⁴. These developments may open up new avenues towards more powerful digital strategies for behaviour change.

Yet, while early feedback intervention studies in which people were provided with information about their energy consumption reported large savings effects of 5–15%^{5,6} with small convenience samples, spurring the large-scale roll-out of smart meters in many countries, those savings have not materialized in larger field trials^{7–9}. The most widespread form of feedback intervention is 'home energy reports': periodic mailings that compare the electricity use of individual households with similar homes in the neighbourhood, thus tapping into social norms³. Deployed at population level (households can opt out, but few do), those programmes typically yield electricity savings of 2%^{3,10}. Other programmes use digital technologies, delivering feedback on electricity use via web portals or in-home displays; studies with large opt-in samples report electricity savings in the range of 1–5%^{8,11–13}—far less than the savings

reported by earlier studies with smaller samples and a higher degree of involvement from study administrators^{6,14}.

Early studies were subject to several methodological issues that compromised the internal and external validity of the results, overestimating the savings potential of these feedback interventions^{7,8,13}. For instance, a meta-analysis of 156 field trials on energy conservation found substantially smaller savings effects of 1.99% for high-quality studies with adequate controls, compared to studies without such controls (9.57%)¹³. Likewise, a meta-analysis of 33 field trials on in-home displays found weighted mean conservation effects of 2.61% for high-quality ('class A') studies using representative sampling techniques, compared to 8.21% for 'class C' studies characterized by small samples of volunteers and a high degree of involvement from study administrators⁸.

Although randomized controlled trials eliminate most threats to the internal validity of studies^{15–17}, the external validity of the results may still be compromised if the people who choose to participate in a study differ from the study's intended population⁷. The vast majority of feedback programmes on energy consumption use opt-in recruitment strategies, where participants actively register to take part in those programmes^{7,9}. There is increasing evidence that individuals who sign up for energy efficiency studies or demand-side management programmes are indeed different from the general population: participation rates are higher among households with high levels of education and income^{18,19}, among more altruistic and environmentally concerned individuals¹⁸ and among those with a higher interest and expertise in energy topics than the general population^{9,20}. Most behavioural programmes do not even provide information about the number of households initially contacted, and those that do, report participation rates in the range of 4–8%^{6,20–23}. These numbers have raised concerns that the results of opt-in studies may be largely biased by an already-motivated subgroup of the general population ('energy enthusiasts' or 'positive greens'), who represent only a small fraction of the population^{9,24}. The response of these volunteers to the treatments may not be very indicative for the response of the

¹ETH Zurich, Zurich, Switzerland. ²University of Bonn, Bonn, Germany. ³University of Bamberg, Bamberg, Germany. ⁴University of St. Gallen, St. Gallen, Switzerland. *e-mail: vtiefenbeck@ethz.ch

Table 1 | Overview of the participating hotels

Hotel	Category	No. of participating rooms	No. of observations
Hotel 1	Business, four-star	96	7,923
Hotel 2	Business, four-star	67	6,123
Hotel 3	Business, four-star	43	2,789
Hotel 4	Business, three-star	11	1,494
Hotel 5	Tourism, four-star	42	814
Hotel 6	Tourism	10	453
Total		269	19,596

general population (volunteer selection bias). On the one hand, it is conceivable that those individuals are already more aware of effective energy conservation measures and have already taken action before the intervention, making it more difficult for them to realize additional savings in those studies²⁵. On the other hand, it is likely that they are particularly open and receptive to these interventions, thus inflating estimates of intervention effectiveness²⁶.

One commonality that the majority of larger energy feedback trials share is that they provide aggregate consumption information at the household level. This makes it difficult for the individual to establish a link between the current action and its impact on energy consumption²⁷. The results of a recent randomized controlled field trial²⁵ suggest that real-time feedback on a specific, energy-intensive activity may induce much larger savings. In a two-month study with an opt-in sample of 697 Swiss households, the treatment group received real-time feedback on the environmental impact of specific, energy-intensive activity (showering), while they could directly take action. The intervention yielded large and stable energy savings of 22% on the target behaviour over the duration of the study. At the household level, this reduction led to much larger conservation gains—also in absolute terms—than aggregate feedback on energy use among the same pool of households. From a technology and cost perspective, the large-scale roll-out of focused real-time interventions is increasingly feasible²⁵. Yet, given the decline in the effect size and the resulting wave of disillusionment once smart metering trials with aggregate feedback moved from small convenience samples to a broader population, the key question is whether the promising large savings effects of activity-specific real-time feedback will also materialize among individuals who do not self-select into a research study.

Another controversial issue is that the communication strategies of most energy-conservation programmes focus on the financial benefits for the consumer as incentives for behaviour change^{28–31}. From a standard economics perspective, this approach makes sense, as rational consumers should respond to monetary incentives in their resource-consumption decisions^{32–34}. Consequently, monetary incentives play a key role in demand-side management^{35,36}; they have the potential to break established consumer patterns and to initiate the development of new patterns of behaviour by making an alternative behaviour more attractive^{37,38}. However, in many contexts, financial motives are not a viable strategy to promote energy conservation: employees, tenants whose rents include utilities or hotel guests who do not pay the marginal cost of their energy consumption. While the provision of large, persuasive monetary benefits does not scale well to the wider population, monetary incentives may also crowd out the intrinsic motivation for pro-social behaviour^{28,39,40} and generate adverse effects⁴¹: as individuals tend to internalize the logic of reward systems easily, monetary incentives can lead to the deterioration of morals and reduce intrinsic motivation^{40,42}.

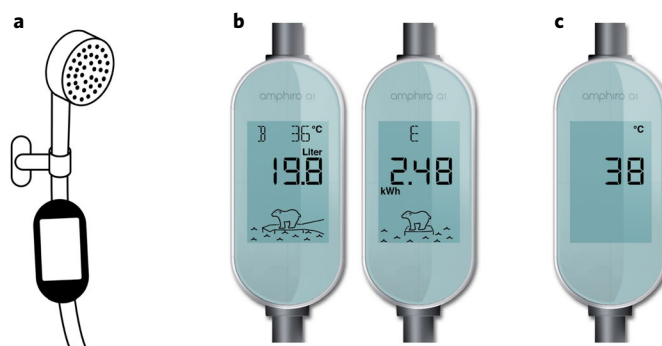


Fig. 1 | Smart shower meter. **a**, The smart shower meter for displaying real-time feedback on resource consumption to hotel guests was installed between the shower head and the shower hose. **b**, Two snapshots of the treatment group's display. **c**, The control group's display.

Here we evaluate whether the large savings effects from digital activity-specific feedback²⁵ are also realistic in settings where a volunteer selection bias can be ruled out, and where study subjects have no financial incentives for resource conservation. We provide activity-specific feedback on resource consumption to uninformed hotel guests during a habitual resource-intensive activity: showering. We find that even in this setting, the digital behavioural intervention creates large and significant conservation effects of 11.4% or 0.215 kWh per shower. Given that most people take a daily shower, scaling up this kind of intervention could produce substantial energy (and water) savings. More importantly, the results suggest that activity-specific real-time feedback—and possibly other digital interventions—has the potential to transform behavioural interventions into a highly relevant policy instrument for fostering energy conservation and behaviour change at the population level^{3,13,43}.

Effect of feedback among uninformed hotel guests

We conducted a natural field experiment in the context of an energy-intensive habitual activity: showering. In a randomized controlled trial, guests at six Swiss hotels (see Table 1) encountered a smart shower meter fitted to the shower in the bathroom of their hotel room. The devices measured the energy and water consumption of every shower taken, and displayed feedback on each ongoing shower in real time. We equipped a total of 265 rooms with these devices and collected a data set with 19,596 observations (water extractions, after pre-processing) from February to April 2016.

The shower meters were installed between the shower hose and the shower head, and included a small screen that activated to display feedback as soon as the water was turned on (Fig. 1). In the treatment condition (60% of rooms, randomly assigned), smart shower meters displayed real-time feedback on the resource consumption of the ongoing shower: total water consumption in litres (one decimal), total energy use in (kilo)watt hours, a dynamic rating of the current energy-efficiency class (A–G) and a four-stage animation of a polar bear standing on a gradually melting ice floe with stage transitions at predefined energy-use thresholds (see Methods for details). This is the same intervention with the same device and display elements as the treatment group of the opt-in household sample in Tiefenbeck et al.²⁵. The remaining 40% of rooms were randomly assigned to the control condition: while guests in the treatment rooms received real-time feedback on their resource use from the beginning of the shower (along with the current water temperature), control-group devices displayed only water temperature.

Hotel guests exposed to real-time feedback consumed significantly less energy per shower than the control group (Fig. 2b). The treatment effect of our intervention is large and significant: guests

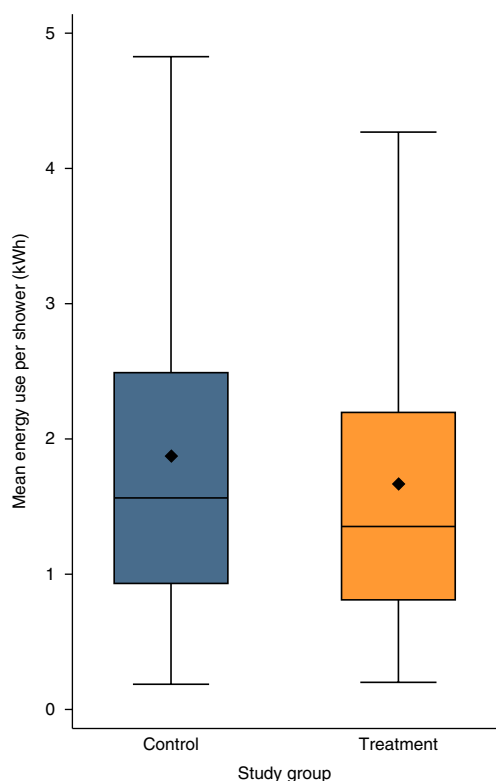


Fig. 2 | Effect of consumption feedback. Group-wise distribution of energy use per shower in hotel rooms with energy consumption feedback (treatment group) and the control group, shown as boxplots ($N=19,596$). The line in the middle of the box represents the median, and the diamond represents the mean energy use. The box spans the first quartile to the third quartile, and the whiskers extend up to 1.5 times the interquartile range from the top or bottom of the box.

in the treatment group used on average 0.215 kWh less energy per shower than the control-group mean of 1.883 kWh (Table 2, column 1). This represents a reduction of 11.4% ($t(19,596) = -4.88$, $P < 0.001$). Controlling for flow rate (column 2), the effect is still highly significant, with a reduction of 0.188 kWh ($t(8,882) = 3.59$, $P < 0.001$), or 10.0%. To determine whether subsampling for observations in which flow rate is available biases this results, we included a third model specification for this subsample but without controlling for flow rate (column 3). The treatment effect is significant in all three models (see Methods for details on the regression analyses) and large (ranging between 10.0% and 13.2%).

The results illustrate that activity-specific real-time feedback can be an efficient measure to foster energy conservation, not only

among a volunteer sample, but also among a random, uninformed sample of individuals.

Additional analyses

As an alternative functional form, we estimated a log-linear regression model. The results (reported in Table 3) are consistent with the results of the non-transformed version reported above and show a strong and significant treatment effect of the real-time feedback.

To further corroborate the reported results, we ran the same models with varying filter thresholds, reducing the data pre-processing to an absolute minimum, with very similar results: if we remove only observations deviating over five standard deviations from mean energy or water consumption, and mean average temperature, we obtain a sample of 25,490 out of the initial 25,647 observations. Running model (1) on this sample yields a slightly smaller, but still highly significant treatment effect of -0.182 kWh (s.e.m. 0.044, $P < 0.001$).

To get an understanding of the effects of the six hotels with their different infrastructure and setting, we also computed a fixed-effects model with dummy variables for the individual hotels. The results are presented in Table 4 and show that the treatment effect is highly significant, albeit slightly smaller than in models (1)–(3). Only in hotel 5, the energy use per shower differs significantly from the other hotels, which may be due to different infrastructure (for example, more low-flow shower heads) or guest characteristics. Otherwise, the impact on energy use per shower is very similar between the different hotels. Regardless of the model specification, the treatment effect is large and significant; thus, non-self-selected participants also respond to real-time feedback in the complete absence of monetary incentives.

Furthermore, we conducted a cost–benefit analysis for installing the metering device in the hotels' showers based on the treatment effect estimated in model (1). We assumed a retail price of 40 CHF for the smart shower meter and fuel cost for water heating of 0.128 CHF kWh⁻¹ and water cost of 3.8 CHF m⁻³, as in Tiefenbeck et al.²⁵. If we extrapolate from the treatment effect of 0.21 kWh and 3.56 l per shower and assume on average 1.2 showers per day per room, as observed during the period of our study, this results in an amortization time of 2.2 years.

Comparisons with a volunteer-household sample

In line with the earlier findings on volunteer selection bias^{7–9}, the treatment effect in the previous study with a volunteer-household sample¹⁸ (0.592 kWh, or 22%) was larger than the effect we observe in the hotel setting (0.215 kWh, or 11.4%). However, it is important to note that we do not seek to quantify the self-selection effect. Our hotel setting and the study context in Tiefenbeck et al.²⁵ differ in multiple aspects other than the two key variables of interest. First, our study examines the behavioural response to feedback in the short term (one or a few nights spent per guest at the hotel). Consequently, any attempt to compare the two studies would need to focus on the short-term behaviour of the household volunteer

Table 2 | Main treatment effect

	Energy use per shower (kWh)		
Consumption feedback (treatment = 1, control = 0)	-0.215^{***} (0.044)	-0.188^{***} (0.050)	-0.252^{***} (0.056)
Flow rate (mean-centred, min ⁻¹)	-	0.098 ^{***} (0.010)	-
Constant	1.884 ^{***} (0.032)	1.881 ^{***} (0.036)	1.902 ^{***} (0.039)
Observations	19,596	8,824	8,824
R^2	0.008	0.047	0.011

Standard errors are in parentheses, adjusted for clustering at the room level; *** indicates significance at the 0.1% level.

Table 3 | Main treatment effect with log transformation of the dependent variable

	Energy use per shower (kWh)		
Consumption feedback (treatment = 1, control = 0)	-0.123*** (0.024)	-0.112*** (0.028)	-0.148*** (0.032)
Flow rate (mean-centred, l min ⁻¹)	-	0.056*** (0.005)	-
Constant	0.412*** (0.017)	0.410*** (0.020)	0.422*** (0.022)
Observations	19,596	8,826	8,826
R ²	0.008	0.047	0.011

Standard errors are in parentheses, adjusted for clustering at the room level; *** indicates significance at the 0.1% level.

sample. Indeed, an analysis of the first three showers only in the volunteer sample also yielded a smaller treatment effect (17.8% or 0.46 kWh) than in the full two-month evaluation. Robustness checks using the first two, four or five (instead of three) showers were conducted, with very similar results. Second, individuals may react differently to feedback in their familiar environment at home versus in a hotel room, or guests may perceive the mere presence of the shower meter as a signal that the hotel management cares about environmental issues and pays attention to how much energy and water their guests are using.

Another remarkable difference relates to participants' energy use per shower in the absence of feedback, measured in the respective control groups. Hotel guests in the control group consumed 28% less energy per shower, namely $M = 1.88$ kWh, s.d. = 1.25 kWh, than the control group in the household setting, with a mean of $M = 2.62$ kWh, s.d. = 1.67 kWh (ref.²⁵), $t(20,236) = -11.1$, $P < 0.001$. This difference is noteworthy for two reasons. First, according to standard economic theory, one would expect individuals to take longer showers at a hotel than at home, as they do pay a marginal cost for every kilowatt hour of consumed energy. Yet, the results suggest that the hotel guests did not exploit the zero marginal cost of consumption. We attribute lower consumption in the control group largely to differences in the technical infrastructure between the hotels and households: we observed lower flow rates in the hotel rooms, probably caused by a higher share of water-saving shower heads installed in the hotel rooms. Second, this difference may also partially explain the smaller treatment effect in the hotels. The study in the household setting had revealed a strong positive interaction between the treatment effect and baseline consumption: in the household study, a 1-kWh increase in baseline consumption led to a 0.32 kWh increase in the savings effect²⁵. To put it simply, it is far easier to cut a 20-min shower short by a few minutes (and kilowatt hours) than to realize substantial reductions on a 1-min shower. The control-group mean of energy consumption per shower in our hotel sample is 1.88 kWh, compared to 2.62 kWh in the household sample (0.74 kWh difference). Interpreting the control-group mean as a proxy for baseline consumption, an increase in baseline consumption by 0.74 kWh would increase the savings effect by 0.24 kWh, which almost exactly matches the difference in the observed savings effect. Thus, if we control for the lower consumption at the hotels in the absence of feedback, the savings effects among the hotel guests and among the volunteer household sample are in fact comparably large.

Discussion

In the case of aggregate feedback, most energy efficiency studies yielded much smaller savings effects once those interventions were evaluated with large, non-self-selected samples⁷⁻⁹. In other words, those interventions resonated much less with broader, non-self-selected audiences than with those individuals who had opted to participate. By contrast, with a highly significant treatment effect of 11.4% among uninformed hotel guests, this study provides empirical

evidence that activity-specific real-time feedback can induce substantial behaviour change among a broader population, even in a setting without monetary incentives for resource conservation. Thus, providing real-time feedback on a specific energy-intensive activity may generate large and persistent savings effects²⁵ not only among the small percentage of the population who tend to opt into energy efficiency studies^{11,20,21,23}; the results indicate that the intervention successfully induces substantial behaviour change and resource conservation among broader audiences.

Regarding the cost-effectiveness of the intervention in the hotel context studied, on the basis of the savings effects observed, the device pays itself off in a hotel within 2.2 years on average, which is a very low amortization time as compared to other energy efficiency investments^{2,44}. Thus, the results suggest that even in settings where third parties pay for the marginal cost of resource consumption, activity-specific feedback can be a cost-effective and scalable strategy to foster energy conservation.

Despite our best efforts, there are limitations to our study. While we were able to measure effects for 100% of the hotel guests, they may not be representative of the general population. Although we diversified our sample by including different types of hotel, with different comfort categories, room rates and primary target customers (business versus tourism), and in different locations, additional studies in other settings and other countries would be valuable. Furthermore, despite our efforts to limit differences in potential Hawthorne effects by displaying real-time water temperature on the control-group devices, it is conceivable that the treatment with real-time feedback on resource consumption draws more attention to the fact that the smart shower meter measures data than does real-time information on water temperature, conveying a stronger feeling of being monitored among its users. Moreover, due to the short duration of guests' hotel stays, we are not able to examine effects over

Table 4 | Treatment effect of consumption feedback and fixed effects for different hotels

	Energy use per shower (kWh)
Consumption feedback (treatment = 1, control = 0)	-0.196*** (0.041)
Hotel 2	0.038 (0.051)
Hotel 3	-0.071 (0.091)
Hotel 4	0.117 (0.090)
Hotel 5	-0.361*** (0.058)
Hotel 6	-0.191 (0.127)
Constant (Hotel 1)	1.917*** (0.046)
Observations	19,602
R ²	0.021

Standard errors are in parentheses, adjusted for clustering per room; *** indicates significance at the 0.1% level.

time. While several studies (lasting between 2 and 16 months^{25,45}) document the mid-term effect stability of activity-specific real-time feedback with opt-in samples, further research needs to investigate whether the large savings effects also persist over time among non-self-selected participants.

Here, we provide robust empirical evidence that activity-specific real-time feedback can induce substantial behaviour change and resource conservation—even for a sample of individuals who neither volunteered to participate in an environmental study, nor reaped financial benefits from energy conservation. Given the debate on volunteer self-selection⁷ and the dwindling treatment effects of other feedback interventions once they are deployed among broader samples^{7–9,12}, this empirical validation is critical to provide solid recommendations for the design of future energy conservation programmes⁴⁶. Information technology increasingly makes it possible to monitor behaviour in real time, to provide individuals with feedback on their ongoing activities and to collect granular data on the real-world impact of interventions from millions of individuals in the field⁴ at rapidly declining costs. The results of this study highlight the potential of digital interventions to transform behaviour in energy-intensive activities, which can be implemented and monitored at the population level.

Methods

Experimental set-up. We conducted a natural field experiment in which we targeted showering as an example of a resource-intensive, low-involvement activity. Participants were not recruited as individuals; instead, we collaborated with six hotels that allowed us to conduct the experiments in their rooms without informing the guests upfront about the experiment. Shower data were collected in batches after several weeks and without a time stamp, which guaranteed complete anonymity of the guests' identity. The study was approved by the Internal Review Board of the University of Bamberg. Similar natural field experiments in hotel contexts have been conducted to investigate the impact of other behavioural interventions such as commitment strategies⁴⁷ or social comparisons^{48,49}.

Overall, guests staying in 265 different hotel rooms took part in our experiment. Hotel guests encountered smart shower meters as part of their rooms' bathroom equipment. The guests who stayed in rooms assigned to the treatment group received real-time feedback on how much energy and water they consumed over the course of their shower (details below). This activity-specific consumption feedback was displayed by a shower meter that had previously been used in framed field experiments in private households²⁵. Room assignment was randomized over floor levels and room categories to minimize confounding factors from differences in infrastructure (for example, water pressure) and did not change throughout the study. Approximately 40% of the rooms were assigned to the control group. They serve as the reference group to calculate the treatment effect. In those rooms, the same device was installed, but it displayed only water temperature. While the temperature reading does not convey information about the resource use and remains relatively static over the course of a shower, it indicates that the device measures data, thus reducing potential differences between the treatment and control group for Hawthorne effects to occur^{50,51}.

As the smart shower meters are powered by the water flow via a small internal generator, the screen displaying the feedback switches on as soon as water flows through the device and remains active for up to three minutes after the end of a shower. Thus, short interruptions to the water flow (for instance while soaping) still result in a single shower being recorded. Before the device switches off, it stores the final data in its internal memory, which was read out at the end of the study.

The experiment took place in six different hotels in Switzerland, recruited on the basis of existing contacts. Data were collected between February and April 2016. Four of the hotels focus on business customers and the other two on private tourists. The categorization into business and tourism was defined on the basis of information provided by the hotels' management. Of course, we cannot rule out that business hotels also had guests on private holidays, or that the tourism hotels hosted some business guests during the course of the study. Depending on their size, the participating hotels allowed us to install shower meters in 10 to 96 of their guest rooms, respectively.

Display content in the treatment group. Most feedback devices display a bundle of elements rather than a single numeric metric⁵² to put the measurement data into context; frequently used elements include historic comparisons, peer comparisons, analogies and energy savings tips. Likewise, the smart shower meters in rooms assigned to the treatment condition displayed water consumption in litres (one decimal), energy use in (kilo)watt hours, current water temperature, a dynamic rating of the current energy-efficiency class (A–G) and a four-stage animation of a polar bear standing on a melting ice floe with stage transitions at predefined

energy use thresholds. The energy consumption displayed on the screen represents the lower bound of the energy used (without losses), and is calculated using the standard engineering formula for heat energy ($Q = m \times c_p \times \Delta T$, with heat energy Q , mass of water m , heat capacity c_p , and ΔT being the difference between the measured water temperature of the ongoing shower and the average cold-water temperature). In the analysis of energy savings, we take into account the same average heating efficiency and losses as in Tiefenbeck et al.²⁵. The energy-efficiency class displayed was inspired by the (static) energy-efficiency class scale indicated on household appliances in Europe. The smart shower meter dynamically indicates the energy efficiency of the current shower based on the energy use in the ongoing shower, starting in energy efficiency class A and progressing to B, C and so on at predefined kilowatt hour thresholds; the thresholds were defined on the basis of the distribution of energy use per shower in a pilot study. The four stages of the polar bear animation are tied to the energy-efficiency class, and change with the transitions from B to C, D to E and E to F, respectively. While the polar bear may be an eye-catching and memorable display element, it does not seem to drive the savings effects. A related study specifically examined the effect of variations of the design choices of the feedback elements; the results indicate that if the polar bear animation makes any difference, it reduces rather than increases the effectiveness of the display⁵².

Data. For each water extraction, the smart shower meter recorded energy and water consumption, average water temperature, interruptions and the duration. In addition, in 168 of the rooms, the average flow rate per shower was also measured. Using the data stored on the device, energy consumption can be converted to water consumption and vice versa. Given the high correlation between water and energy consumption per shower (0.989)²⁵, the choice of the unit of analysis does not change the results in any meaningful way. This article focuses on resource consumption in units of energy in kilowatt hours.

The raw data set included observations of 25,647 measured showers from 269 hotel rooms at 6 different hotels (see Data availability). In a first pre-processing step, the data were cleaned by removing outliers from malfunctioning devices; to this end, observations that deviated by over 3 standard deviations from the mean of the energy consumed or water volume per shower were removed from the sample—that is, only observations in the interval $[\bar{x} - 3 \times \text{s.d.}, \bar{x} + 3 \times \text{s.d.}]$ were retained. Furthermore, we removed data points that most likely did not represent showers—for example water extractions of volumes below 6.5 litres and observations deviating over 2 standard deviations from average temperature, which probably represent cleaning or other procedures. A member of the research team accompanied cleaning personnel at one hotel for several hours to gather information on cleaning practices to identify water extractions for cleaning. The specific choice of 6.5 litres was based on this assessment; we conducted robustness checks in which we changed this threshold to other values (5 litres or 10 litres), which generated very similar results.

After this pre-processing step, the final data set included 19,602 showers from 265 hotel rooms (11,384 observations in the treatment group and 8,218 in the control group). The average flow rate per shower could be measured only for 168 rooms and 8,826 observations, so only these data points are included in the estimation of models (2) and (3) in Table 2.

Since the study is a natural field experiment with uninformed participants, we were not able to collect socio-demographic data about the guests who stayed in the rooms with the smart shower meters during the study.

Data analysis. We used the data points observed in the control group to quantify the energy use per shower in the participating hotels without feedback. To estimate the treatment effect of the feedback intervention, we used a simple linear regression model and also computed its log-linear transformation:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i \quad (1)$$

$$\ln y_i = \beta_0 + \beta_1 x_i + \epsilon_i \quad (2)$$

where the dependent variable y_i is the energy consumption in shower i . The variable x_i is binary, indicating treatment (=1) or no treatment (=0), and thus coefficient β_1 estimates the treatment effect. The intercept β_0 represents the control-group mean in this model, as $x_i = 0$ for observations in the control group. The results of this analysis are reported in Table 2, column 1, and in Table 3, column 1, with the natural logarithm of the dependent variable.

For rooms in which the smart shower meter also measured the flow rate, we estimated an additional model that controls for the centred flow rate in litres f_i :

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 f_i + \epsilon_i \quad (3)$$

$$\ln y_i = \beta_0 + \beta_1 x_i + \beta_2 f_i + \epsilon_i \quad (4)$$

The results for model (3), which includes the flow rate in the regression, are reported in Table 2, column 2, and for model (4) in Table 3, column 2, with log-linear

transformation. In both model specifications, standard errors were clustered at the room level to account for infrastructural influences. Two-sided *t*-tests were conducted to test whether the coefficients were significantly different from zero.

To get an understanding of the effects of the six hotels with their different infrastructure and setting, we also computed a fixed-effects model with dummy variables for the individual hotels. In this model, the constant represents the estimates for the largest hotel (hotel 1) and dummy variables are included for the other hotels. Results are reported in Table 4.

$$y_i = \beta_0 + \beta_1 x_i + \alpha_2 h_{2i} + \alpha_3 h_{3i} + \alpha_4 h_{4i} + \alpha_5 h_{5i} + \alpha_6 h_{6i} + \epsilon_i \quad (5)$$

Code availability. The Stata code used to generate results reported in this article is available on request from the corresponding author.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data that support the findings of this study are available at <https://doi.org/10.6084/m9.figshare.6984323.v1>.

Received: 27 March 2018; Accepted: 9 October 2018;

Published online: 19 November 2018

References

- Sovacool, B. K. Diversity: energy studies need social science. *Nature* **511**, 529–530 (2014).
- Allcott, H. & Greenstone, M. Is there an energy efficiency gap? *J. Econ. Perspect.* **26**, 3–28 (2012).
- Allcott, H. & Mullainathan, S. Behavior and energy policy. *Science* **327**, 1204–1205 (2010).
- Tiefenbeck, V. Bringing behaviour into the digital transformation. *Nat. Energy* **2**, 17085 (2017).
- Darby, S. *The Effectiveness of Feedback on Energy Consumption - A Review for DEFRA of the Literature on Metering Billing and Direct Displays* (Environmental Change Institute, University of Oxford, Oxford, 2006).
- Ehrhardt-Martinez, K. et al. *Advanced Metering Initiatives and Residential Feedback Programs: A Meta-Review for Household Electricity-Saving Opportunities* (Washington, DC, American Council for an Energy-Efficient Economy, 2010).
- Davis, A. L., Krishnamurti, T., Fischhoff, B. & Bruine de Bruin, W. Setting a standard for electricity pilot studies. *Energy Policy* **62**, 401–409 (2013).
- McKerracher, C. & Torriti, J. Energy consumption feedback in perspective: integrating Australian data to meta-analyses on in-home displays. *Energy Effic.* **6**, 387–405 (2013).
- Kelly, J. & Knottenbelt, W. Does disaggregated electricity feedback reduce domestic electricity consumption? A systematic review of the literature. 1605, 00962v2 (2016).
- Allcott, H. & Mullainathan, S. *External Validity and Partner Selection Bias* Working Paper No. 18373 (NBER, 2012).
- Schleich, J., Klobasa, M., Götz, S. & Brunner, M. Effects of feedback on residential electricity demand: findings from a field trial in Austria. *Energy Policy* **61**, 1097–1106 (2013).
- Buchanan, K., Russo, R. & Anderson, B. The question of energy reduction: The problem(s) with feedback. *Energy Policy* **77**, 89–96 (2015).
- Delmas, M. A., Fischlein, M. & Asensio, O. I. Information strategies and energy conservation behavior: A meta-analysis of experimental studies from 1975 to 2012. *Energy Policy* **61**, 729–739 (2013).
- Abrahamse, W., Steg, L., Vlek, C. & Rothengatter, T. A review of intervention studies aimed at household energy conservation. *J. Environ. Psychol.* **25**, 273–291 (2005).
- Haynes, L., Service, O., Goldacre, B. & Torgerson, D. *Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials* (London Cabinet Office Behavioural Insights Team, 2012).
- Vine, E., Sullivan, M., Lutzenhiser, L., Blumstein, C. & Miller, B. Experimentation and the evaluation of energy efficiency programs. *Energy Effic.* **7**, 627–640 (2014).
- Campbell, D. T. Reform as experiments. *Am. Psychol.* **24**, 409–429 (1969).
- Clark, C. F., Kotchen, M. J. & Moore, M. R. Internal and external influences on pro-environmental behavior: Participation in a green electricity program. *J. Environ. Psychol.* **23**, 237–246 (2003).
- Sulyma, I., Tiedemann, K., Pedersen, M., Rebman, M. & Yu, M. Experimental evidence: a residential time of use pilot. In *Proc. ACEEE Summer Study Energy Effic. Build.* 292–304 (ACEEE, 2008).
- Baladi, S. M., Herriges, J. A. & Sweeney, T. J. Residential response to voluntary time-of-use electricity rates. *Resour. Energy Econ.* **20**, 225–244 (1998).
- Herter, K., Wood, V. & Blozis, S. The effects of combining dynamic pricing, AC load control, and real-time energy feedback: SMUD'S 2011 Residential Summer Solutions Study. *Energy Effic.* **6**, 641–653 (2013).
- Schleich, J., Klobasa, M., Götz, S. & Brunner, M. Effects of feedback on residential electricity demand-findings from a field trial in Austria. *Energy Policy* **61**, 1097–1106 (2013).
- Lossin, F., Kozlovskiy, I., Sodenkamp, M., & Staake, T. Incentives to go green: an empirical investigation of monetary and symbolic rewards to motivate energy savings. In *Proc. Eur. Conf. Inf. Syst.* 1–16 (2016).
- A Framework for Pro-Environmental Behaviours* (Department for Environment Food and Rural Affairs, 2008).
- Tiefenbeck, V. et al. Overcoming salience bias: how real-time feedback fosters resource conservation. *Manag. Sci.* **64**, 983–1476 (2016).
- Frederiks, E. R., Stenner, K., Hobman, E. V. & Fischle, M. Evaluating energy behavior change programs using randomized controlled trials: Best practice guidelines for policymakers. *Energy Res. Soc. Sci.* **22**, 147–164 (2016).
- Faruqui, A., Sergici, S. & Sharif, A. The impact of informational feedback on energy consumption-A survey of the experimental evidence. *Energy* **35**, 1598–1608 (2010).
- Schwartz, D., Bruine de Bruin, W., Fischhoff, B. & Lave, L. Advertising energy saving programs: The potential environmental cost of emphasizing monetary savings. *J. Exp. Psychol. Appl.* **21**, 158–166 (2015).
- Allcott, H. & Sweeney, R. L. The role of sales agents in information disclosure: evidence from a field experiment. *Manag. Sci.* **63**, 21–39 (2017).
- Energy Efficiency Campaign: 'Deutschland Macht's Effizient'* (German Federal Ministry for Economic Affairs and Energy, 2016); <http://www.deutschland-machts-effizient.de/>
- Customer Incentives for Energy Efficiency Through Program Offerings* (US Environmental Protection Agency, 2010).
- Landry, C. E., Lange, A., List, J. A., Price, M. K. & Rupp, N. G. Toward an understanding of the economics of charity: evidence from a field experiment. *Q. J. Econ.* **121**, 747–782 (2006).
- Karlan, D. & List, J. A. Does price matter in charitable giving? Evidence from a large-scale natural field experiment. *Am. Econ. Rev.* **97**, 1774–1793 (2007).
- Goette, L., Stutzer, A. & Frey, B. M. Prosocial motivation and blood donations: a survey of the empirical literature. *Transfus. Med. Hemotherapy* **37**, 149–154 (2010).
- Olmstead, S. M. & Stavins, R. N. Comparing price and nonprice approaches to urban water conservation. *Water Resour. Res.* **45**, 1–10 (2009).
- Borenstein, S. The long-run efficiency of real time electricity pricing. *Energy J.* **26**, 93–116 (2005).
- Maki, A., Burns, R. J., Ha, L. & Rothman, A. J. Paying people to protect the environment: a meta-analysis of financial incentive interventions to promote proenvironmental behaviors. *J. Environ. Psychol.* **47**, 242–255 (2016).
- Bamberg, S. Is a residential relocation a good opportunity to change people's travel behavior?: Results from a theory-driven intervention study. *Environ. Behav.* **38**, 820–840 (2006).
- Frey, B. S. & Oberholzer-Gee, F. The cost of price incentives: an empirical analysis of motivation crowding-out. *Am. Econ. Rev.* **87**, 746–755 (1997).
- Sandel, M. J. *What Money Can't Buy: The Moral Limits of Markets* (Macmillan, London, 2012).
- Gneezy, U. & Rustichini, A. A fine is a price. *J. Legal Stud.* **29**, 1–17 (2000).
- Thøgersen, J. Monetary incentives and environmental concern. Effects of a differentiated garbage fee. *J. Consum. Policy* **17**, 407–442 (1994).
- Karlin, B., Zinger, J. F. & Ford, R. The effects of feedback on energy conservation: a meta-analysis. *Psychol. Bull.* **141**, 1205–1227 (2015).
- Schopfer, S., Tiefenbeck, V. & Staake, T. Economic assessment of photovoltaic battery systems based on household load profiles. *Appl. Energy* **223**, 229–248 (2018).
- Tiefenbeck, V., Tasic, V., Schöb, S. & Staake, T. Long-lasting effects or short-term spark? On the persistence of behaviour change induced by real-time feedback on resource consumption. In *Proc. Eur. Conf. Inf. Syst.* 1–17 (2016).
- I'm Not Surprised. *Nat. Energy* **2**, 17101 (2017).
- Baca-Motes, K., Brown, A., Gneezy, A., Keenan, E. A. & Nelson, L. D. Commitment and behavior change: evidence from the field. *J. Consum. Res.* **39**, 1070–1084 (2013).
- Goldstein, N. J., Cialdini, R. B. & Griskevicius, V. A room with a viewpoint: using social norms to motivate environmental conservation in hotels. *J. Consum. Res.* **35**, 472–482 (2008).
- Schultz, P. W., Khazian, A. M. & Zaleski, A. C. Using normative social influence to promote conservation among hotel guests. *Soc. Influ.* **3**, 4–23 (2008).
- Schwartz, D. et al. The Hawthorne effect and energy awareness. *Proc. Natl Acad. Sci. USA* **110**, 15242–15246 (2013).
- Tiefenbeck, V. On the magnitude and persistence of the Hawthorne effect - evidence from four field studies. In *Proc. European Conference on Behaviour and Energy Efficiency* 1–6 (2016).

52. Ableitner, L., Schöb, S., Tiefenbeck, V. & Fridgen, G. Real-world impact of information systems: the effect of seemingly small design choices. In *Proc. Work. Inf. Technol. Syst.* 1–16 (2017).

Acknowledgements

We thank S. Häcki and T. Bachmann from Swiss Mobiliar insurance for their great efforts in reaching out to hotels and managing the on-site study implementation and data collection. We would also like to express our gratitude to the management of the six Swiss hotels for the opportunity to run the study. Funding for this work was provided by the MTEC foundation of ETH Zurich (data analysis) as well as by Swiss Mobiliar insurance (hardware, deployment).

Author contributions

V.T. and T.S. designed the study. S.S. wrote the software of the study devices and enabled the technical side of the data collection. V.T. oversaw the study implementation. A.W. and V.T. analysed the data. V.T. and A.W. drafted the manuscript; T.S. and E.F. edited the manuscript. V.T., T.S. and E.F. secured funding for the study.

Competing interests

V.T., A.W. and E.F. declare no competing financial interests. T.S. and S.S. are co-founders of and hold shares in Amphiro AG, the SME that manufactures the smart shower meters. T.S. and S.S. were not involved in the data analysis, hotel selection or room assignment.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41560-018-0282-1>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to V.T.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2018

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated
- Clearly defined error bars
State explicitly what error bars represent (e.g. SD, SE, CI)

Our web collection on [statistics for biologists](#) may be useful.

Software and code

Policy information about [availability of computer code](#)

Data collection

Custom software for the shower meter Amphiro was used to measure resource consumption in the showers in 265 rooms in 6 different hotels and for the data read-out

Data analysis

Stata SE 14.1 was used to preprocess the resource consumption data, to conduct regression analyses, and plot the presented graphs

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data that support the findings of this study is available on figshare.com: <https://doi.org/10.6084/m9.figshare.6984323.v1>.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Behavioural & social sciences

Study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Our study is a randomized controlled trial at six hotels, to examine resource consumption during showering in a natural field experiment in a setting of complete absence of monetary incentives among a sample of uninformed participants who did not self-select into the study. We collected quantitative data from 19,602 showers (energy and water consumption, temperature, flowrate).
Research sample	The study sample contains data on resource consumption in the shower from the guests of 265 rooms of 6 different tourist and business hotels in Switzerland. Due to the the hotels' privacy regulations we were not able to collect any demographical data, but this sampling strategy had the advantage that our sample did not self-select into the study. This point is discussed in the article.
Sampling strategy	We ran the study as a natural field experiment, thus guests encountered the measurement and feedback device as part of their bathroom equipment. We conducted powertests based on the data of a similar previous study in households (Tiefenbeck et al. (2016), Management Science).
Data collection	A smart shower meter was used to collect data on resource consumption from every water extraction during the study period. The device called "Amphiro" is depicted in the article. The data is stored on the device and was read out by the researchers at the end of the study.
Timing	Data collection took place from February to April 2016.
Data exclusions	The data was cleaned by removing outliers from malfunctioning devices; to this end, observations that deviated by over 3 standard deviations from the mean of the energy consumed or water volume per shower were removed. Furthermore, we excluded data points that most likely did not represent showers – e.g. water extractions of volumes below 6.5 litres and observations deviating over 2 standard deviations from average temperature, which probably represent cleaning or other procedures .
Non-participation	We are not aware of any hotel guests that asked to have the device removed, so to the best of our knowledge, data on all guests of the participating rooms during the period of the study are present in our raw dataset.
Randomization	At each hotel, we randomly allocated rooms to the two experimental conditions; the randomization was done by floor level to account for heterogeneity in the flow rates in the different floors of the buildings.