NBER WORKING PAPER SERIES

DIFFERENCES IN HOW AND WHY SOCIAL COMPARISON AND REAL-TIME FEEDBACK IMPACT RESOURCE USE: EVIDENCE FROM A FIELD EXPERIMENT

Mark A. Andor
Lorenz Goette
Michael K. Price
Anna Schulze Tilling
Lukas Tomberg

Working Paper 31845
http://www.nber.org/papers/w31845

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050Massachusetts Avenue
Cambridge, MA 02138
November 2023

We are grateful for comments and suggestions by Max Auffhammer, Fabian Dehos, Eugen Dimant, Valeria Fanghella, Stefano Della Vigna, Christian Honow, Joachim Schleich, Christoph M. Schmidt, and Colin Vance, as well as by participants of the Advances with Field Experiments Conference 2022 (AFE), the Annual Conference of the Association of Environmental and Resource Economists 2022 (AERE), the Annual Conference of the European Economic Association 2023 (EEA), the Annual Conference of the German Committee for Environmental and Resource Economics 2023 (AURO), the Mannheim Energy Conference 2023, the 2022 World Economic Science Association Conference (ESA), the CAIS Research Incubator, the 9th Retreat of the CRC TR 224, the Behavioral group at IDOS, and seminars at University of Duisburg-Essen, Grenoble Ecole de Management and the Mercator Research Institute on Global Commons and Climate Change (MCC). Furthermore, we thank Sven Hansteen, Florin Martius and Valerie Peetz for excellent research assistance. In addition, we are indebted to Michael Schild for his tremendous support in the technical implementation of the study. This project was conducted in research partnership with the Center for Advanced Internet Studies (CAIS). We gratefully acknowledge financial support by the Ministry of Culture and Science of the State of North Rhine-Westphalia and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under EPoS Collaborative Research Center Transregio 224 and Germany's Excellence Strategy – EXC 2126/1-390838866. The experimental design was pre-registered with the AEA RCT Registry as trial AEARCTR-0006747 and IRB approval was obtained from the GfeW (Certificate No. nQ9jc48B). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Differences in How and Why Social Comparison and Real-Time Feedback Impact Resource Use: Evidence from a Field Experiment
Mark A. Andor, Lorenz Goette, Michael K. Price, Anna Schulze Tilling, and Lukas Tomberg
NBER Working Paper No. 31845
November 2023
JEL No. C93,D12,Q25,Q55

ABSTRACT

We compare the behavior and welfare effects of two popular interventions for resource conservation. The first intervention is social comparison reports (SC), which primarily provide consumers with information motivating behavioral change. The second intervention is real-time feedback (RTF), which primarily provides consumers with information facilitating behavioral change. In a field experiment with around 1,000 participants, we directly observe the interventions' effects on participants' behavior. Further, we elicit participants' willingness to pay for receiving the interventions, both before and after having experienced them for one month. We find that SC leads to a reduction in water use per shower by 9.4%, RTF by 28.8%, and the combination (BOTH) by 35.0%. Our willingness to pay results show that all interventions are highly valued by participants and that willingness to pay for RTF and BOTH is significantly higher than for SC. Furthermore, we find that the valuation of the interventions do not change following one-month experience. Our results suggest that while both interventions improve welfare, providing consumers with information facilitating behavioral change achieves a higher impact and a slightly higher welfare increase than providing consumers with information motivating behavioral change.

Mark A. Andor
RWI - Leibniz Institute for Economic Research
Hohenzollernstraße 1-3
Essen NRW 45128
Germany
andor@rwi-essen.de

Lorenz Goette
National University Singapore
lorenz@lorenzgoette.org

Michael K. Price
Department of Economics, Finance,
and Legal Studies
The University of Alabama
250 Alston Hall
Box 870224
Tuscaloosa, AL 35487
and Australian National University
and also NBER
mkprice2@cba.ua.edu

Anna Schulze Tilling
Bonn Graduate School of Economics
schulzetilling@uni-bonn.de

Lukas Tomberg
RWI - Leibniz Institute for Economic Research
Hohenzollernstraße 1-3
Essen NRW 45128
Germany
lukas.tomberg@rwi-essen.de

A randomized controlled trials registry entry is available at
https://www.socialscienceregistry.org/trials/6747

1 Introduction

In the last decade, there has been increasing interest in behavioral economic interventions from research, society, and politics. A whole series of empirical studies have investigated the effectiveness of behavioral interventions in a wide range of areas,[1] behavioral insights groups have been installed around the world (Obama, 2015; OECD, 2017), several books on behavioral interventions became bestsellers (Ariely and Jones, 2008; Thaler and Sunstein, 2009; Kahneman, 2011; Halpern, 2015), and Daniel Kahneman, Robert Shiller and Richard Thaler each won a Nobel prize.

However, the welfare effects of such interventions have rarely been studied. From a cost-effectiveness standpoint, many behavioral interventions triumph (Allcott and Mullainathan, 2010; Benartzi et al., 2017). By design, they usually do not impose monetary costs on the consumer, and implementation costs are low. These characteristics make behavioral interventions particularly attractive to policymakers and even interventions with small effects are often assessed as "cost-effective". However, such cost-effectiveness analyses usually ignore non-monetary costs and benefits, including shame, pride, and social pressure (DellaVigna et al., 2012; 2016; Butera et al., 2022), (dis-)utility resulting from changes in consumption behavior (Allcott and Kessler, 2019), or the non-instrumental (dis-)utility of information provision (Sharot and Sunstein, 2020).

This paper contributes to this discussion by offering a comprehensive welfare analysis of two popular behavioral interventions affecting consumers along two different dimensions. While one intervention primarily affects consumers by providing information that *motivates* behavioral change, the other provides information that *facilitates* a change in behavior. While many interventions can be classified along similar lines (see Congiu and Moscati, 2020), it is still unclear whether targeting one or the other dimension is more effective in creating behavioral change and a positive welfare impact. To provide a clean evaluation for informational interventions,[2] we compare both

[1] For recent meta-analyses see, for example, Jachimowicz et al. (2019), Nisa et al. (2019), Cadario and Chandon (2020), and Mertens et al. (2022).

[2] Our focus in this project is on behavioral interventions providing information, which constitute a

interventions in a common context: fostering resource conservation in showering behavior.[3]

We focus on social comparison reports (SC) as a representative and well-studied[4] example of an intervention providing consumers with information that *motivates* a change in behavior. Our SC reports provide treated participants with weekly information on their average water use per shower and compare it to that of other participants. This communicates the social desirability of conserving resources, motivating participants to conserve water and energy.

We focus on real-time feedback (RTF) as a representative and well-studied[5] example of an intervention providing consumers with information that *facilitates* a change in behavior. Our RTF intervention activates participants' shower heads to glow in different colors when showering, depending on how much water has already been consumed during the current shower. This solves two problems impeding behavioral change in the absence of RTF. First, the quantity of energy and water consumed while showering is usually only imperfectly observable to consumers, leading to consumers either not being able to act upon this aspect at all or making consumers reliant on imperfect heuristics. Consumption can be thought of as a shrouded attribute in this setting.[6] Second, consumers' attention might not be focused on energy and water conservation at the moment of showering – even if a consumer had developed a perfect

subset of the behavioral interventions described in, e.g., Congiu and Moscati (2020). Our findings might not apply to other types of interventions, e.g., interventions changing product attributes or product placement.

[3]Showering consumes a considerable amount of resources, both due to the energy necessary to heat the water used and the water use itself. A typical individual in our sample uses 33 liters of hot water per shower, which requires on average 1.6 kWh to heat it up. In comparison, the average household in the European Union uses 1.0 kWh for lighting per day (Faberi et al., 2015), and a modern refrigerator uses 0.63 kWh per day (Michel et al., 2016). Furthermore, showering behavior is prone to behavioral biases, such as salience bias (Tiefenbeck et al., 2018), and thus to inefficiently high levels of resource consumption. At the same time, it is a daily activity that can be directly controlled by consumers, making it an ideal application for behavioral interventions.

[4]See, for example, Allcott (2011); Ferraro and Price (2013); Allcott and Rogers (2014); Brent et al. (2015); Goette et al. (2021b).

[5]See, for example, Tiefenbeck et al. (2018; 2019); Byrne et al. (2021); Fang et al. (2021); Goette et al. (2021a).

[6]The most prominent examples of shrouded attributes are of a financial nature, such as taxes (Chetty et al., 2009), shipping costs (Gabaix and Laibson, 2006; Brown et al., 2010), or marginal prices (Ito, 2014). In the context of resource consumption, which is what we are concerned with, it is consumption itself that is often not observable to consumers (Jessoe and Rapson, 2014).

heuristic, she would still require an unusually high level of self-control to be able to direct her attention toward this heuristic during every shower. The RTF intervention addresses both of these problems by making resource consumption salient and actionable, facilitating behavioral change.

Our experiment design allows us to draw conclusions on both the behavioral and welfare effects of both interventions SC and RTF, as well as their combination (BOTH). We equip all of our 1,003 field experiment participants in 574 households with smart shower heads to track showering behavior, capturing water use, temperature, and time of every shower. After observing baseline showering behavior for four weeks, participants in the treatment groups experience an information intervention for four weeks based on the experimental group to which they are exogenously assigned. For participants receiving RTF, we remotely activate the real-time feedback function in their shower heads. For participants receiving SC, we configure weekly e-mail reports. The third treatment, BOTH, combines both interventions: Participants receive real-time feedback through their shower head and social comparison reports. Further, we elicit participants' willingness to pay to experience the interventions for four weeks. We elicit willingness to pay twice. We first ask participants for their willingness to pay before having experienced any of the interventions, and then again after they have experienced one of the interventions due to their exogenous assignment. The willingness to pay bids are binding and thus constitute revealed preferences, as the field experiment includes a phase in which the treatments are allocated endogenously based on the elicited willingness to pay.

The willingness to pay elicitation constitutes the basis of our welfare analysis. The willingness to pay subsumes the private benefits from resource savings, the change in shower comfort, the psychological costs and benefits of the interventions, and any further factors relevant to the consumers. Adding the societal benefits from the reduction in carbon emissions due to the reduction in hot water use caused by the interventions and subtracting the program costs, taking into account the marginal cost of public funds, yields estimates of the welfare effects of the interventions.

In terms of behavioral impact, we find that the RTF intervention decreases water use per shower by 28.8%. SC reports decrease water use per shower by 9.4%, and the combination BOTH by 35.0%. RTF is thus around three times as effective as SC reports in our context, and the combination of both interventions is most effective. Our results suggest that providing consumers with information facilitating behavioral change is more effective than providing information motivating consumers to change behavior.

While all interventions are highly valued by participants, RTF and the combination of both interventions are around 10% more highly valued than SC reports alone. We find that the experience of the interventions does not substantially alter the willingness to pay for them. Furthermore, the endogenous allocation of the interventions based on willingness to pay instead of exogenous allocation based on a random draw leads to similar conservation effects. Willingness to pay is, on average, much above and beyond actual and perceived cost savings realized through the interventions, implying the existence of psychological benefits. For around 25% to 28% of participants, however, willingness to pay is lower than perceived savings, implying that these participants incur psychological costs with the intervention.

In terms of welfare impact, we estimate the highest welfare increase for the combined intervention and the lowest increase for the SC intervention. Recognizing that our sample of volunteers may be particularly motivated to conserve resources and to receive the interventions, we conduct an additional welfare analysis under particularly conservative assumptions with regard to conservation effects and consumer surplus. This conservative analysis yields positive welfare effects for both RTF and the combined intervention, while the welfare effect of SC is close to zero. Our results indicate that providing consumers with information facilitating behavioral change has a slightly more positive effect on consumer welfare than providing information motivating behavioral change.

We contribute to the literature in three ways. First, our study contributes to the relatively young literature on the welfare effects of behavioral interventions. As recent literature has pointed out (Allcott and Kessler, 2019; Butera et al., 2022), policymakers must be cautious of behavioral interventions producing the desired behavioral effect at the cost of imposing disproportionate psychological costs on consumers. In response, a dynamic strand of research has emerged. Essentially, two main approaches to welfare analysis have developed, which differ in their way of quantifying changes in consumer surplus: one that derives consumer surplus from structural models or sufficient statistics (Chetty et al., 2009; DellaVigna et al., 2012; 2016; Rodemeier, 2020; Löschel et al., 2022; Rodemeier and Löschel, 2022), and one that directly measures consumer surplus by eliciting the willingness to pay to receive a behavioral intervention, which may be combined with a structural approach (Allcott and Kessler, 2019; Allcott et al., 2022; Butera et al., 2022). Our paper aligns with this second strand of research, providing the first welfare comparison of two interventions affecting consumers along different dimensions. Importantly, the two interventions and their combination are evaluated within the same experimental context, providing an apples-to-apples comparison so far lacking in the literature.

Another difference to previous literature lies in our double elicitation of consumer willingness to pay, both before and after receiving the intervention. Previous literature on welfare effects differs not only in the consumption context studied but also in the timing of the willingness to pay inquiry. Allcott and Kessler (2019) elicit consumer willingness to pay to continue receiving an intervention, while Butera et al. (2022) elicit willingness to pay for receiving an intervention for the first time. By eliciting both measures within the same consumption context, we provide evidence on whether economically large differences exist between the two measures and provide guidance for future research on how an experiment can be designed to capture both.

Second, our paper relates to conceptual work classifying behavioral interventions. Our focus on the two dimensions of motivating vs. facilitating behavioral change most closely relates to Congiu and Moscati (2020) who classify interventions into a *message* and an *environment* dimension. Our findings can also inform earlier classifications by Johnson et al. (2012), Hansen and Jespersen (2013), and Mongin and Cozic (2018) with empirical evidence. The added value of our paper lies in our ability to provide a clean and comprehensive comparison of two information interventions affecting behavior among different dimensions. Since consumption context, participant sample, and timeline are common, differences in treatment and welfare effects can be cleanly and causally attributed to differences in the nature of the interventions.

Finally, our study contributes to previous literature on the efficacy of behavioral interventions in reducing resource consumption. Social comparison reports have been prominently discussed in previous literature (Allcott, 2011; Ferraro and Price, 2013; Allcott and Rogers, 2014; Brent et al., 2015; Goette et al., 2019; 2021b), but are not generally cost-effective across different countries (Andor et al., 2020) and may impose substantial (intangible) costs that reduce the welfare effects of such interventions (Allcott and Kessler, 2019). Real-time feedback has been identified as a promising intervention to influence showering behavior, with previous studies reporting higher effect sizes than for social comparison reports (e.g., Tiefenbeck et al., 2018).[7] However, the existing evidence is derived from very different consumption contexts, e.g., studying the electricity consumption of a specific appliance versus aggregated electricity consumption. Our paper adds to recent research comparing different behavioral interventions within the same consumption context (Brandon et al., 2019; Fang et al., 2021). Our findings are relevant for all consumption contexts in which fine-grained data is available with smart sensors and the question is how to best utilize this data to increase the effectiveness of interventions, adding to a broader strand of literature (e.g., Jessoe and Rapson, 2014; Asensio and Delmas, 2015; Grubb and Osborne, 2015; Ito et al., 2018; Brülisauer et al., 2020; Gerster et al., 2021). They also connect to the literature focusing on shrouded attributes and the importance of making shrouded attributes salient at the moment of purchase (e.g., Gabaix and Laibson, 2006; DellaVigna and Pollet, 2009; Chetty et al., 2009; Grubb, 2009; Brown et al., 2010; Taubinsky and Rees-Jones, 2018). The shrouded attribute in our setting is the quantity consumed – making it both difficult for consumers to allocate their attention to this attribute and costly to calculate adequate estimates. Since consumers in this case need to develop heuristics to guide

[7] For more detailed reviews of behavioral interventions applied in the context of resource consumption see, for example, Andor and Fels (2018) and Khanna et al. (2021).

their behavior, and our interventions help develop effective and simple heuristics – for example, to stop showering when real-time feedback gives a specific cue – we also connect to the literature on heuristic thinking (e.g., Lacetera et al., 2012; List et al., 2023).

The paper proceeds as follows: In section 2, we describe the experimental design and our study sample. In section 3, we present the behavioral impact of the interventions and explore the mechanisms underlying treatment differences, specifically focusing on the role of social comparison reports in motivating behavioral change and the role of real-time feedback in facilitating behavioral change. We then present participants' willingness to pay for receiving the interventions, and our analysis of the welfare effects of each of the interventions. Section 4 concludes.

2 The field experiment

2.1 Technical equipment

We provided all study participants with a smart shower head, a WiFi gateway, and installation instructions. Details on the functioning and installation of the shower head are included in Appendix A5. The shower head and WiFi gateway, connected to the participants' WiFi network, formed a data infrastructure that allowed the shower head to continuously send data on the water use in liters, date, time, and temperature of every shower to the researchers. The participants did not have access to this data during the study.

2.2 Sample

Participants were recruited in collaboration with the German survey institute *forsa*, which maintains a sample of panelists that is representative of German-speaking internet users aged 14 and above. In late August 2020, *forsa* sent out an invitation e-mail to 9,376 panelists, living in the Rhine-Ruhr-Area, Germany's largest metropolitan area.[8]

[8] This area was chosen because two of the research institutions involved in the research project, RWI – Leibniz Institute for Economic Research in Essen and the University of Bonn, are located in this region, thus enabling potential technical support at the premises of the participating households.

The invitation email conveyed some general information about the environmental impact of showering, information on the compensation for participation,[9] and information about the tasks as a study participant, which included installing the shower head and data infrastructure and completing several surveys throughout the study. The study was presented to the participants under the name "Project Sustainable Showering". Those who were interested in participating were referred in the email to a project website where more detailed information was provided and where they had the opportunity to register for study participation. The invitation email and the project homepage neither explained that a randomized experiment was to be conducted, nor which interventions were planned. Registration took place via an online form on the project homepage in which we queried contact details, consent to data processing within the study, and initial information on household size and technical circumstances to ensure that participants met the requirements to take part in the study. The content of the invitation email is depicted in Appendix A2, and the content of the project website is presented in Appendix A3.

1,100 of the invited persons registered for the study, representing 11.7% of those originally invited. We selected the final study participants among the eligible registrations[10] targeting smaller households with fewer showers in the home because we consider it more likely for these households that the information provided to the participants in the course of the experiment would reach all shower users. Since we targeted a total sample size of around 600 households and expected some attrition during the study, we confirmed the participation of a total of 685 participants and sent them an invitation for the baseline survey in October 2020. 647 participants filled out the baseline survey within the given time frame. Those who did not fill out the survey after a reminder email were not considered for study participation anymore. From the

[9]Participants were compensated for study participation in two ways. First, they were allowed to keep the smart shower head after the study period and could then control it independently using an app. Secondly, the participants received a monetary amount of 35 EUR, which was paid out to them at the end of the study in the form of a voucher of their choice.

[10]We had to delete 25 registrations because they were either obvious duplicates or the participants provided an address outside the Rhine-Ruhr-Area. Further, 203 registered persons did not meet the technical requirements needed for successful installation of the shower head and the data infrastructure.

outset, we divided the participants into two study waves. The first study wave began immediately, while the second wave began in January 2021. The division into the two study waves was made to logistically simplify the roll-out of the smart shower heads. This division allowed for regional clustering of the respective waves, which in turn minimized travel time for on-site technical support. Apart from the difference in timing, the research design was identical for both study waves. The final study sample consists of 574 households.[11]

Table 1: Socio-economic summary statistics

Variable	Explanation	Mean	Std. Dev.
Age	Age of respondent	53.85	13.31
Female	Dummy: 1 if respondent is a woman	0.39	–
Household size	Number of persons living in the household	1.75	0.59
Children	Dummy: 1 if children live in the household	0.07	–
College degree	Dummy: 1 if respondent has a university degree	0.49	–
Number of showers	Number of showers installed in the household	1.25	0.52
Income	Monthly net household income in EUR	3489.93	1445.03
N	574		

Note: This table shows summary statistics for the final sample of study participants, i.e., participants who filled out the baseline survey and successfully installed the shower head infrastructure.

Table 1 provides summary statistics on the study sample. As a result of our selection process, the average household size is smaller than that of the German population (1.75 compared to 2.0; BiB, 2023) as is the share of households with children (7% compared to 29%; Destatis, 2022a). Correspondingly, the monthly net household income is smaller than in the population (3,490 EUR compared to 3,813 EUR; Destatis, 2022b). The study sample furthermore consists of older participants (54 years compared to 45 years; BiB, 2022), a higher share of participants with a college degree (49% compared to 19%; Destatis, 2022a) and a lower share of female respondents (39% females compared to 51%; Destatis, 2022a).

[11]The remaining 73 households either dropped out after the baseline survey or were unable to install the smart shower heads or the data infrastructure.

2.3 Experimental design

Our experiment consists of three phases illustrated in Figure 1: A baseline phase, an exogenous treatment phase, and an endogenous treatment phase.[12] In the baseline phase, i.e., the first four weeks of the experiment, we tracked showering behavior without introducing any behavioral intervention. The baseline phase was followed by the exogenous treatment phase, i.e., four weeks in which the behavioral interventions were exogenously assigned to the participants. In the following four weeks, the endogenous treatment phase, we reassigned the experimental groups based on participants' willingness to pay (WTP) to receive the interventions. The primary purpose of this endogenous treatment phase was to make the WTP inquiry consequential and thus incentive-compatible (Carson and Groves, 2007).[13] Furthermore, it allows us to observe the conservation effects in a setting in which participants receive the interventions based on their WTP. The experimental design was pre-registered with the AEA RCT Registry as trial AEARCTR-0006747 and IRB approval was obtained from the GfeW (Certificate No. nQ9jc48B).

2.3.1 Experimental groups

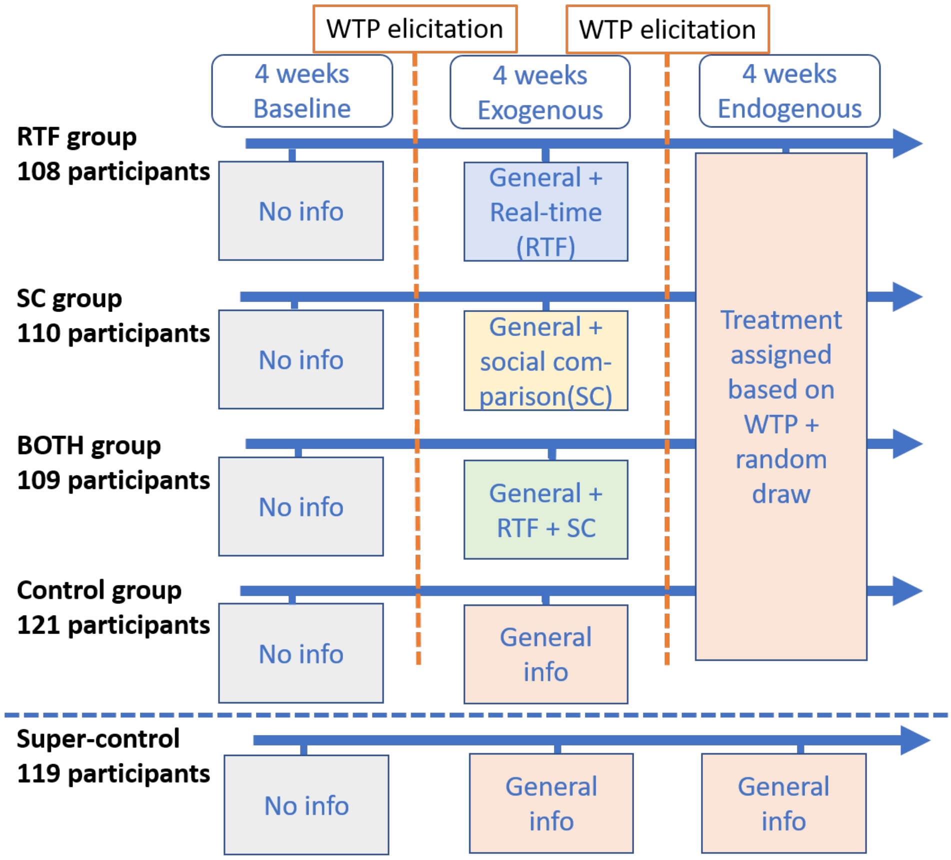
For the exogenous treatment phase, we split the sample into five equally-sized experimental groups. We allocated the experimental groups randomly, with stratification ensuring that they did not significantly differ from each other concerning baseline water use per shower, age, household size, number of showers installed in the household, environmental attitudes, perceived social norms regarding energy use, and education. The five experimental groups are characterized as follows:

Super-control group: This group received no behavioral intervention throughout the study and served two purposes: First, it provided data on showering behavior over time in the absence of any intervention and second, this group was used to form

[12] Baseline phase of wave 1 (wave 2): From October 10, 2020 to November 19, 2020 (from February 11, 2021 to March 10, 2021). Exogenous treatment phase of wave 1 (wave 2): From November 11, 2020 to December 12, 2020 (from March 11, 2021 to April 13, 2021). Endogenous treatment phase of wave 1 (wave 2): From December 21, 2020 to January 20, 2021 (from April 14, 2022 to May 13, 2021).

[13] Details on the treatment assignment mechanism in the endogenous treatment phase are provided in Appendix A1.1.3.

Figure 1: Overview of the experimental design



Note: This figure shows the experimental design. Participant treatment during the exogenous phase depended on treatment group assignment. Participant treatment during the endogenous phase depended on WTP indicated during the two elicitations (before and after the exogenous phase) and a random draw.

reference groups needed for the behavioral interventions: Each participant who was not in the super-control group was randomly assigned to a reference group consisting of nine members of the super-control group. Participants who received the social comparison reports were compared to their reference group in the comparison reports. Similarly, the thresholds in the real-time feedback intervention were based on the behavior of the reference group. Because the super-control group was not eligible for an intervention in either phase of the study, the WTP inquiries filled out by these households were purely hypothetical. Yet, to ensure that these participants stayed at a similar attention level regarding their study participation as the other participants, they received weekly emails ("newsletters") with general information about showering and its energy intensity throughout the study; as did all other study participants. The upper panel of the exemplary newsletter in Appendix A4 contains one example of the general information texts.

Control group: Like the super control-group, this group received weekly newsletters but no intervention in the exogenous treatment phase. Yet, this group was eligible for receiving an intervention in the endogenous treatment phase, making this group's WTP inquiry binding and thus incentive compatible.

Social comparison (SC) group: Participants assigned to this group received a social comparison section in their weekly newsletters in addition to the general information section. In this social comparison section, the average water use per shower (and corresponding CO_2 emissions, given the average water temperature) was reported and compared to the average of the participant's reference group constituting of untreated households from the super-control group. Furthermore, weekly averages of the water use per shower of both the participant household and its reference group were plotted in a time diagram. In addition, participants were shown how their water consumption per shower ranks compared to the nine reference households, i.e., they were assigned a rank between one (lowest use) and ten (highest use). For each rank, an additional emoticon was displayed: A smiling face for comparatively low consumption, a neutral face for medium consumption, and a frowning face for high consumption. To ensure that enough data was available to calculate the social comparison, the information was updated every two weeks, so that the second newsletter repeated the information of the first newsletter, the fourth newsletter repeated the information of the third newsletter, and so on. The middle panel of the exemplary newsletter in Appendix A4 is an example of the social comparison information.

Real-time feedback (RTF) group: For participants assigned to this intervention, the real-time feedback function was activated in participants' shower heads. The LEDs built into the shower head signaled the water consumption of the shower in real time. Shower heads were configured to light up green for the first 15 liters of a shower. Thereafter they turned blue, purple, red and eventually started to flash in the red color. The thresholds at which the colors changed were determined by the behavior of the participant's untreated reference group. In detail, the shower head would start to flash red once the reference group's average water use per shower over the past two weeks was exceeded. The thresholds of the remaining colors were allocated in regular intervals between 15 liters and the average water use per shower of the reference group. For these participants, a RTF section was added to the weekly newsletter. This section contained information on the water consumption corresponding to each of the color thresholds, but participants were not informed about the origin of the color thresholds. For each threshold in liters, the corresponding average CO_2 emissions caused by heating the water to the usual water temperature of the household were also detailed. To ensure that enough data was available to reliably construct the thresholds, the thresholds were updated every second week, as they were for the SC group so that the newsletter in between was a repetition of the newsletter of the week before. The lower panel of the exemplary newsletter in Appendix A4 is an example of how the real-time feedback functionality was communicated.

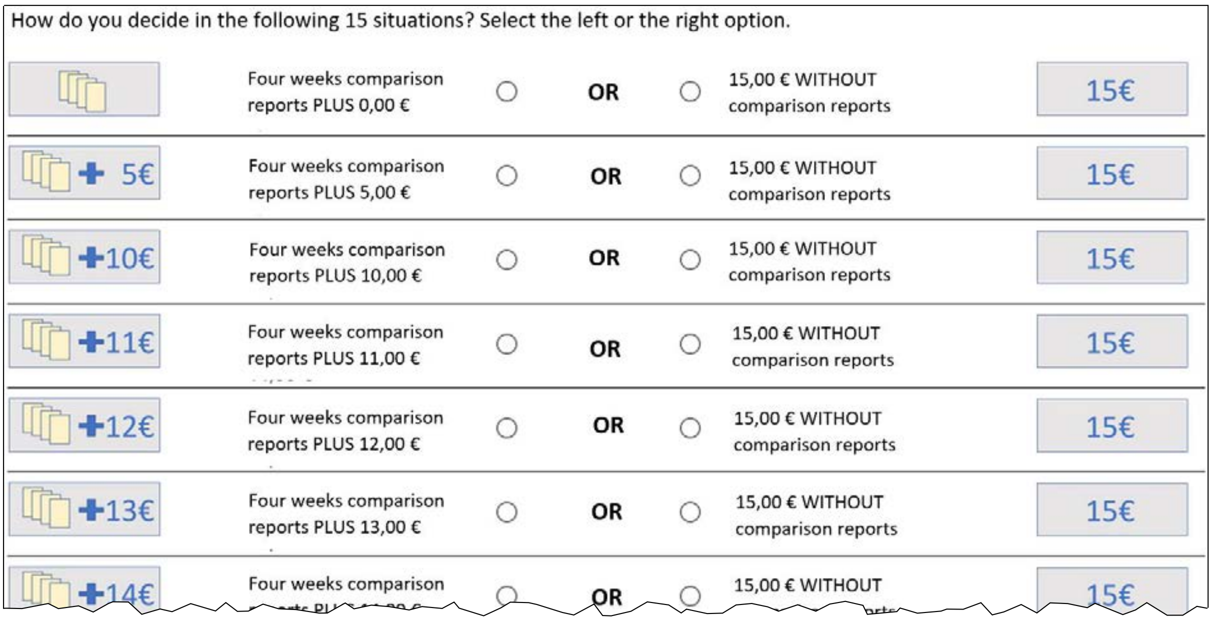
BOTH group: For participants assigned to this intervention, the SC section was added to the weekly newsletters, and the RTF function in the smart shower heads was activated. Participants thus received both the interventions of the SC group and the RTF group. The weekly newsletters of this group thus contained a SC section as well as a RTF section, and the full newsletter in Appendix A4 is a representative newsletter of a participant that received the BOTH intervention. Having the information from both interventions, the participants in this group were able to infer that their shower thresholds depended on their reference group's average water use per shower.

2.3.2 Elicitation of willingness to pay for the interventions

Towards the end of the baseline phase and the end of the exogenous treatment phase, participants filled out a willingness to pay (WTP) inquiry, in which they indicated how much they are willing to pay to receive each of the interventions (SC, RTF, and BOTH) in the endogenous treatment phase.

The WTP inquiry was conducted in the form of a survey which consisted of several multiple price lists (MPL), with one MPL for each of the three interventions. In each MPL, the participants were asked to decide if they preferred to receive four weeks of

Figure 2: Illustration of the multiple price list for SC



How do you decide in the following 15 situations? Select the left or the right option.

	Four weeks comparison reports PLUS 0,00 €	○	OR	○ 15,00 € WITHOUT comparison reports	15€
+ 5€	Four weeks comparison reports PLUS 5,00 €	○	OR	○ 15,00 € WITHOUT comparison reports	15€
+10€	Four weeks comparison reports PLUS 10,00 €	○	OR	○ 15,00 € WITHOUT comparison reports	15€
+11€	Four weeks comparison reports PLUS 11,00 €	○	OR	○ 15,00 € WITHOUT comparison reports	15€
+12€	Four weeks comparison reports PLUS 12,00 €	○	OR	○ 15,00 € WITHOUT comparison reports	15€
+13€	Four weeks comparison reports PLUS 13,00 €	○	OR	○ 15,00 € WITHOUT comparison reports	15€
+14€	Four weeks comparison	○	OR	○ 15,00 € WITHOUT	15€

Note: This figure is abridged - left side options continue with 15 €, 16 €, 17 €, 18 €, 19 €, 20 €, 25 €, 30 €. Multiple price lists for RTF and BOTH were identical apart from the left-hand-side illustrations.

the intervention in question plus a varying monetary amount or whether they preferred to forego the intervention and receive an amount of 15 EUR. An example of an MPL eliciting WTP for the social comparison reports is provided in Figure 2. The participants had to fill out an MPL for each treatment twice, once before and once towards the end of the exogenous treatment phase, totaling six MPLs. Except for the super-control group, these inquiries were binding in that one of the decisions made, i.e., the decision of a randomly drawn row (of 15) in one randomly selected MPL (out of six), was actually implemented in the endogenous treatment phase. So the participant would, depending on her decision, receive the treatment and the corresponding monetary amount or the fixed amount of 15 EUR.

3 Results

3.1 Data

As a first step, we present the shower data from the baseline phase in Table 2.[14] For the baseline phase as well as the exogenous treatment phase, we pool the super-control group and the control group to maximize power and refer to this pooled group as the "(super-)control group". We see a slight difference in baseline water use and baseline shower temperature in the (super-)control group compared to the three treatment groups. This difference is only statistically significant in one case (water temperature of the SC group) and will not affect our regression results, since we include household fixed effects in all regression specifications.

Table 2: Shower statistics in the baseline phase by experimental group

Variable	**Unit**	**(Super-) control**	**SC**	**RTF**	**BOTH**
Baseline water use	Liters per shower	34.07	32.16	31.89	32.31
			(0.89)	(1.10)	(0.80)
Baseline shower water temperature	°C	38.12	37.17	36.88	37.81
			(2.53)	(1.43)	(0.77)
Baseline water flow	Liters per Minute	9.18	9.21	9.05	8.87
			(0.12)	(0.53)	(1.23)
Baseline Shower frequency	Number of showers per day	0.96	1.12	1.08	0.99
			(1.13)	(1.31)	(0.38)
No. of households		236	110	109	109
No. of observations		8,416	4,207	4,360	4,056

Note: This table shows baseline averages. The test statistics in parentheses are t-values for comparison to the (super-)control group obtained from OLS regressions with either baseline water use, water temperature, water flow, or shower frequency as the outcome variable and group membership as the explanatory variable with standard errors clustered at the household level. 10 households in the (super-)control group were only able to properly connect their shower heads after the baseline phase, so they are excluded from these statistics and all analyses in this section.

3.2 Impact of the interventions on behavior

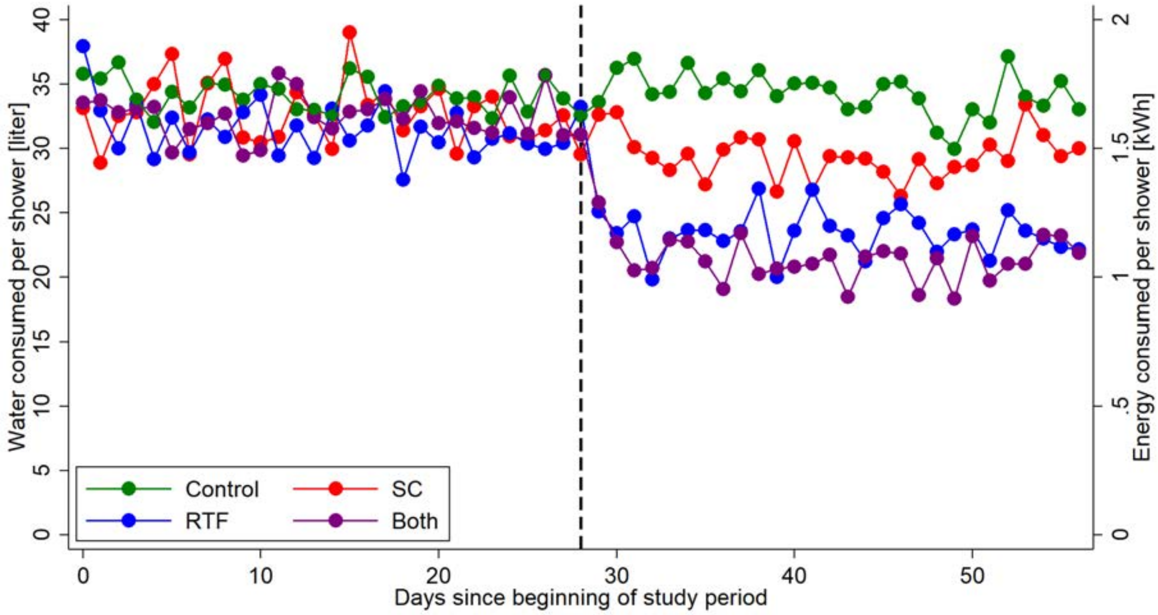
Figure 3 depicts the daily average water use per shower in the different treatment groups in the baseline phase and in the exogenous treatment phase.[15] During the first four weeks, average water use per shower is not distinguishable between the (super-)control group and the treatment groups. The start of the treatment phase, as indicated by the blue vertical line in Figure 3 leads to a sharp and immediate drop in

[14]We cleaned the raw shower data, in that we remove showers with water consumption above 200 liters or below 5 liters. This is necessary so that water withdrawals that are very likely not showers do not add additional noise to the results. A similar approach is taken by Fang et al. (2021). It should also be noted that showers taken while the WiFi gateway was not connected to the Internet were, for technical reasons, recorded with the time stamp of the next successful data transmission from the WiFi gateway and not the actual time stamp of the shower. This is the case for 17.2% of showers taken throughout the entire study period. To correct this, we redefine the time stamps of such offline showers and allocate them evenly over the period between the last and the next successful data transmission.

[15]This graph pools data from the two study waves, such that the onset of the treatment phase was on November 20, 2020 for the first study wave and on March 11, 2021 for the second study wave.

water use per shower for all groups that received treatment, while the behavior of the control group does not change systematically. Moreover, visual inspection suggests that RTF and BOTH lead to stronger reductions in water use per shower than SC.

Figure 3: Exogenous treatment phase - average water use per shower by day and treatment group



Note: This figure compares average water use per shower between treatment groups during the exogenous treatment phase, shown on the left vertical axis. The right vertical axis translates water consumption into energy consumption required to heat the water, given a water temperature of 38 °C (the average water temperature in our sample).

In a next step, we estimate the treatment effects using a difference in differences regression model of the following form:

$$y_{jit} = \alpha_i + \beta_1 SC_{it} + \beta_2 RTF_{it} + \beta_3 BOTH_{it} + \tau_t + \epsilon_{jit}, \tag{1}$$

where y_{jit} represents the outcome variable, i.e., water used during shower j by household i on day t. α_i represents household-specific fixed-effects. SC_{it}, RTF_{it} and $BOTH_{it}$ are indicators of household i receiving the respective treatment at time t. τ_t represents day-specific fixed effects and ϵ_{jit} is the error term that is clustered at the household level.

The estimates of the average treatment effects in the exogenous treatment phase are depicted in Table 3. The SC treatment reduced water use per shower by 3.19 liters, the RTF treatment by 9.79 liters, and the BOTH treatment by 11.91 liters. Relative to the average water use of the (super-)control group in the exogenous treatment phase, which is 34.1 liters, these effects correspond to a 9.4%, 28.8%, and 35.0% reduction in water use per shower. All effects are statistically significant at the 1% level. The effect of SC is significantly smaller than the effects of RTF and BOTH (p-values of the differences: <0.00) and the effect of BOTH is somewhat larger than that of RTF (p-value of the difference: 0.13). The robustness checks presented in Appendix A1.1.1 indicate that the estimates of the effects remain largely unchanged if household or time fixed effects are excluded from the regression model. In addition, the interventions have only a negligible effect on shower temperature and shower frequency.

Table 3: Average treatment effects on water use per shower in the exogenous phase

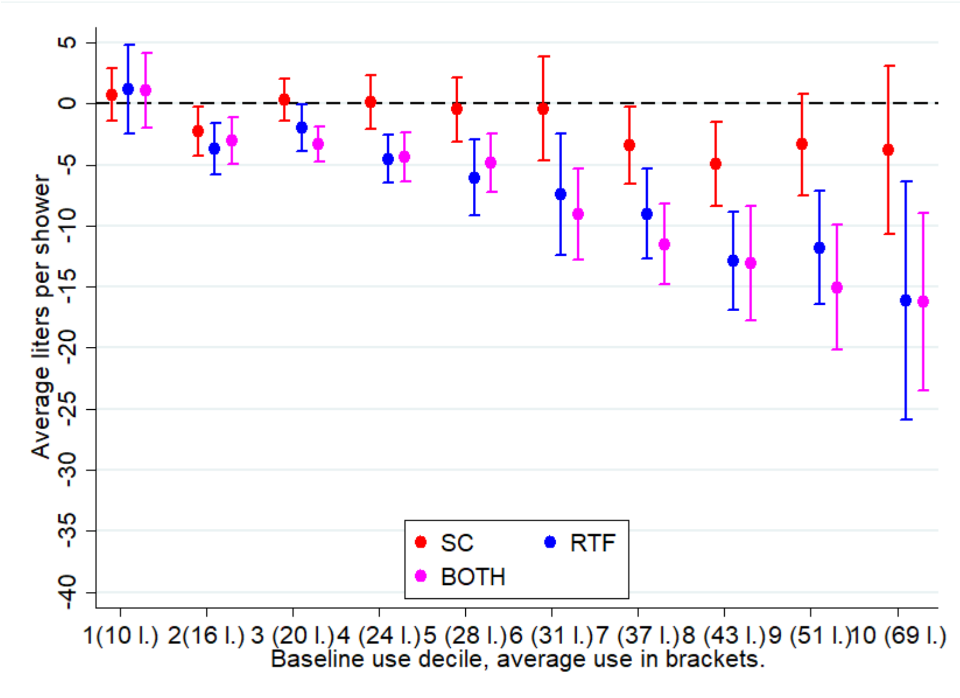
	(I) Volume (Liters)	(II) Relative to control group
SC	-3.19***	-9.4%
	(0.66)	
RTF	-9.79***	-28.8%
	(0.88)	
BOTH	-11.91***	-35.0%
	(1.22)	
(super-)control group average	34	
No. of households	564	
No. of observations	38,453	

Note: OLS estimates with household and time fixed effects included. Standard errors are clustered at the household level. ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively. The (super-)control group average corresponds to the average water use per shower of the control and super-control group in the exogenous treatment phase. Column (II) contains manual calculations obtained by dividing the treatment effects in column (I) by the (super-)control group average. These percentages are indistinguishable from those obtained when the dependent variable is divided by the average of the (super-)control group before performing the regression analysis.

The effect of 28.8% that we find for RTF is slightly larger than that found in Tiefenbeck et al. (2018), who estimate that RTF reduces water use per shower by 22%. The 9.4% reduction in water use per shower caused by the SC intervention is substantially larger than that reported by Allcott (2011), who finds that SC reports decrease aggregate household electricity use by 2%. One reason for this difference could be that our SC reports are appliance-specific and thus could make behavioral adjustments easier.

Previous studies (e.g., Allcott, 2011; Tiefenbeck et al., 2018; Andor et al., 2020) have shown that the effectiveness of behavioral interventions on resource conservation in absolute terms increases in baseline resource use. For example, this may be because there is a higher potential for savings when initial resource use is high. We find a similar pattern with our sample. As shown in Figure 4, this pattern is particularly pronounced for RTF and BOTH, where the treatment effects increase from below 5 liters per shower in the second decile of baseline water use to around 15 liters in the 10th decile. By contrast, the effect of SC is close to zero in the first six deciles of baseline water use and increases to about 5 liters thereafter.

Figure 4: Treatment effect by baseline deciles



Note: The markers visualize effect sizes by baseline use decile, with tails indicating 95% confidence intervals.

3.2.1 Including the endogenous treatment phase

Next, we include data from the endogenous treatment phase. This analysis is insightful from several angles: First, we can investigate the treatment effects when the interventions are allocated based on individual preferences for receiving them. This is an important use case as such interventions, as long as they are not mandated by the government, will be rolled out on an opt-in basis. To avoid incorrectly attributing potential long-run effects of the interventions from the exogenous treatment phase to the endogenous treatment phase, we control for the interventions received in the exogenous treatment phase in the regression. The results of this analysis, presented in column (I) of Table 4, reveal that treatment effects are slightly lower in the endogenous treatment phase compared to the results from the exogenous treatment phase, which represent estimates of the average treatment effects (ATE) in absence of self-selection into treatment. Specifically, we find that in the endogenous treatment phase SC reduces water consumption per shower by 2.83 liters, RTF by 6.73 liters, and BOTH by 9.07 liters. These effects can still be considered quite large, showing that the interventions can lead to considerable conservation effects even with endogenous allocation.

Table 4: Average treatment effects on water use per shower in both treatment phases

	(I) Only endo. phase	(II) Pooled exo. & endo. phase	(III) Pooled exo. & endo. phase (IV approach)
SC	-2.83***	-3.10***	-3.97***
	(1.04)	(0.54)	(0.56)
RTF	-6.73***	-8.34***	-9.05***
	(1.23)	(0.73)	(0.78)
BOTH	-9.07***	-10.50***	-11.60***
	(1.09)	(0.74)	(0.77)
Controls for treatment in the exo. phase	Yes	Yes	Yes
Instrumental variables	No	No	Yes
No. of households	564	564	564
Observations	37,113	54,527	54,527

Note: OLS/IV estimates with household and time fixed effects included. Standard errors are clustered at the household level. ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively. Standard errors in column III are calculated by bootstrap. In column III, we instrument treatment receipt in the endogenous treatment phase by the factors governing which of the 90 MPL questions was chosen for implementation. These are 1) whether the first or second elicitation was chosen for implementation, 2) which of the three MPL lists was chosen, and 3) which of the 15 questions on the MPL was chosen.

As an additional step, we pool the data from the exogenous and endogenous treatment phases. This maximizes the size of our estimation sample. These results are depicted in column (II) of Table 4. The estimated coefficients are, as expected in between the coefficients from the endogenous phase presented in column (I) and those from the exogenous phase presented in Table 3. The standard errors of the estimated coefficients decrease substantially due to the larger sample. As a result, the effects of all three interventions are now significantly different from each other on the 5%-level, substantiating that SC has a complementary effect to RTF, as the joint provision of SC and RTF in the BOTH group leads to a significant increase in the effect compared to the RTF group.

In column (III) of Table 4, we correct for the endogeneity of treatment assignment in the endogenous phase by using an instrumental variables (IV) approach.[16] We find that the treatment effects increase slightly compared to the ones from the pooled model in column (II) and are similar to the ones from the exogenous phase in Table 3.

3.2.2 Exploring the mechanisms underlying treatment differences

Having established that RTF induces greater behavior change than SC and that the effect of RTF is enhanced when combined with SC, it is natural to explore how we can explain these differences in effectiveness in terms of the mechanisms through which the different interventions may operate.

Regarding the mechanisms that SC could operate through, we adopt the intuition by Allcott (2011) who argues that SC could work because "if households are uncertain about some part of their production function, the social comparisons may facilitate social learning about their privately-optimal level of energy use" (p. 1084). In our case this means that households may not have thought about how much water they would need to use to have a satisfactory shower experience, but begin to optimize when they learn that others use much less. Second, he argues that "the treatment may directly affect the 'moral cost' of energy use" (p. 1084), i.e., in our case, excessive water use while showering becomes associated with (increased) moral costs once a household receives SC reports. Both mechanisms boil down to the idea that the SC reports increase the motivation and thus the intention to conserve water and energy while showering.

However, it is well known that people tend to exhibit an "intention-behavior gap", meaning that the right intentions are not always followed by changes in behavior that

[16] For details on this approach see Appendix A1.1.3.

are in line with those intentions. Two of the key problems that drive the intention-behavior gap, as summarized by Sheeran and Webb (2016, p. 507), are "fail to get started", e.g., not remembering or ignoring the intention to conserve when getting into the shower, and "fail to keep goal pursuit on track", e.g., failing to put the intention to take shorter showers into practice while enjoying a pleasant shower, or simply not knowing when to stop showering in order to achieve the personal conservation goals.[17] Combining the SC reports with the RTF intervention has the potential to overcome both of these challenges, thus bridging the intention-behavior gap: Because the light signals are directly emitted by the shower heads and could not be turned off during the study, they reminded participants of their conservation intentions each time they began to shower, and the color changes reminded them continuously while they were showering, making it unlikely to be distracted from those intentions.

Even if participants were able to fully bridge the intention-behavior gap without RTF, they would still have difficulties achieving their conservation goals because the exact amount of water and energy they use while showering would remain a "shrouded attribute" (Gabaix and Laibson, 2006). This is because the standard deviation of water flow (liters per minute) within a household is quite high in our sample, being 0.71 at the median.[18] Thus, even participants with perfect self-control and the use of heuristics such as using a stopwatch to monitor shower time would not be able to accurately control their water use per shower. For example, a person who uses an average of 35 liters of water per shower, which takes about 3.5 minutes at 10 liters per minute, would be on average about 2.5 liters off if she sets a stopwatch to 3.5 minutes while showering. RTF addresses this problem by providing information on the amount of water used while showering, which should allow the person showering to directly relate shower duration to the amount of water and energy used, thus facilitating the precise pursuit of her conservation goals. This reasoning is in line with Jessoe

[17]Note that we use the term "conservation goal" here to describe whatever abstract goal the participants have in their mind when they start changing their behavior. This is not to be confused with experimentally induced goal setting as for example investigated in Goette et al. (2021a).

[18]This variation cannot be explained by different people using a shower, as the median standard deviation is still 0.64 even for one-person households.

and Rapson's (2014) finding that high-frequency electricity consumption information allows households to better understand electricity use quantities and thereby to more accurately respond to price signals.

Table 5: Average treatment effect on the standard deviation of water use per shower within households

	(I) Volume (Liters)	(II) Relative to control group
SC	-2.23***	-13.1%
	(0.68)	
RTF	-4.29***	-25.2%
	(0.71)	
BOTH	-5.83***	-34.3%
	(0.74)	
(super-)control group average	17	
No. of households	564	
No. of observations	1,127	

Note: OLS estimates with household fixed effects included. Standard errors are clustered at the household level. ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively. The (super-)control group average corresponds to the standard deviation of water use per shower of the control and super-control group in the exogenous treatment phase. Column (II) contains manual calculations obtained by dividing the treatment effects in column (I) by the (super-)control group average.

Table 5 shows the effect of the interventions on variation in water use per shower, comparing within-household variation in water use in the pre-intervention period with within-household variation in the intervention period. Treatment effects are significantly stronger for RTF and BOTH compared to SC. While all three interventions reduce this variation, RTF and BOTH reduce it more than twice as much as SC. This suggests that RTF makes it significantly easier for participants to achieve a certain target behavior in terms of water use per shower than SC alone.

Similarly, we show the treatment effects on the proportion of showers ended within a given color threshold in Table 6. These results show that, on average, participants who received SC decreased the proportion of showers that ended while the shower head was flashing red by 6 percentage points and, in turn, increased the proportion of showers that ended while the shower head was still green by 4 percentage points. In contrast, participants that received RTF (BOTH) decreased their share of showers ended while the shower head was flashing red to a much larger extent, i.e., 19 (25) per-

Table 6: Average treatment effect on the proportion of showers ended within a certain color threshold

	(I)	(II)	(III)	(IV)	(V)
	Effect on the proportion of showers ended while the shower head was...				
	...green	...blue	...purple	...red	...flashing red
SC	0.04*	0.03	-0.02	0.02	-0.06***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
RTF	0.12***	0.06**	0.01	0.01	-0.19***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
BOTH	0.22***	0.06*	-0.02	-0.00	-0.25***
	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)
(super-)control group average	0.26	0.11	0.15	0.12	0.35
No. of households			446		
No. of observations			882		

Note: OLS estimates with household fixed effects included. Standard errors are clustered at the household level. ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively. The dependent variable is the share of showers per shower head ended within a given color threshold. We calculated the hypothetical color thresholds in the baseline phase as we did in the treatment phases, that is, we based them on the baseline average water consumption per shower of a participant's reference group, which consisted of participants from the super-control group. Therefore, participants in the super-control group are not included in these analyses because there are no valid hypothetical color thresholds available for them. See Table A3 in the Appendix for an overview of the individual transitions between color thresholds.

centage points. In turn, they increased the share of showers ended while the shower head was still green by 12 (22) percentage points and the share of showers ended while the shower head was blue by 6 (6) percentage points. Again, these results suggest that RTF allowed participants to condition their behavior much more closely on actual water use during showering than participants in the SC group were able to do.

3.3 Impact of the interventions on willingness to pay

As a further step, we analyze the WTP data, which we expect to subsume the utility loss and the perceived monetary savings due to shorter showers as well as any psychological costs and benefits caused by the interventions. Our study is unique in that we elicit participants' WTP both before and after utilizing the interventions. By doing so, we gather information on the anticipated utility of the interventions for the participants, or in other words, how much they are willing to pay in the hopes of gaining from the interventions. By comparing the WTP data collected before and after the participants have used the interventions for a month, we can ascertain if the participants' perception of the interventions has shifted.

Since WTP was collected with an MPL that provides interval-censored data, we take a conservative approach and set the WTP values to the respective lower limit of the respective interval, providing lower-bound values of WTP. Thus, those who always opt for the intervention in an MPL would receive a WTP value of 15, those who indicated WTP between 10 and 15 EUR received a WTP value of 10, and so forth. Since the lower end of the MLP is an open interval limit for those who always chose the voucher without an intervention, we set the WTP to -20 in this case.

Table 7 depicts the average WTP values and the shares of participants with negative WTP for the full sample (panels I and II) as well as separately for the respective treatment groups (panels III and IV) and for the control group (panels V and IV). For each panel, the first column depicts WTP elicited in the baseline phase and the second column depicts WTP elicited in the exogenous treatment phase. From panel I, we find that WTP for SC is significantly lower than that for RTF (10%-level) and BOTH (1%-level), but all average WTP figures are at around 10 EUR. In addition, we find that WTP is slightly lower in the exogenous treatment phase than in the baseline phase in all panels. However, as indicated by the small and non-significant regression results in panel VII, this decrease does not appear to be due to treatment experience.

Focusing on the proportion of participants who indicated a negative WTP for the interventions (panels II, IV, VI, and VIII), i.e., participants who were willing to forgo money to avoid receiving the interventions, we find that these shares vary between 3.4% and 9.2%, with hardly any systematic differences between the interventions.

Table 7: Average willingness to pay for the interventions

		WTP elicited in the baseline phase		WTP elicited in the exogenous treatment phase		
		Mean	95%-Conf. int.	Mean	95%-Conf. int.	Obs.
Full sample						
	Average WTP amounts					
I	WTP for SC	9.58	[8.90 – 10.27]	8.92	[8.22 – 9.61]	439
I	WTP for RTF	10.12	[9.42 – 10.83]	9.53	[8.82 – 10.25]	439
I	WTP for BOTH	10.31	[9.63 – 10.98]	9.76	[9.06 – 10.45]	439
	Share of participants with negative WTP					
II	WTP for SC	6.8%	[4.5% – 9.2%]	5.9%	[3.7% – 8.1%]	439
II	WTP for RTF	5.9%	[3.7% – 8.1%]	5.7%	[3.5% – 7.9%]	439
II	WTP for BOTH	6.1%	[3.9% – 8.4%]	5.5%	[3.3% – 7.6%]	439
WTP in the respective treatment groups						
	Average WTP amounts					
III	WTP for SC	8.93	[7.48 – 10.38]	8.59	[7.21 – 9.97]	109
III	WTP for RTF	9.63	[7.97 – 11.29]	9.13	[7.54 – 10.72]	105
III	WTP for BOTH	11.13	[9.91 – 12.35]	9.67	[8.34 – 11.01]	107
	Share of participants with negative WTP					
IV	WTP for SC	6.4%	[1.7% – 11.1%]	9.2%	[3.7% – 14.7%]	109
IV	WTP for RTF	8.6%	[3.1% – 14.0%]	6.7%	[1.8% – 11.5%]	105
IV	WTP for BOTH	3.7%	[0.1% – 7.4%]	3.7%	[0.1% – 7.4%]	107

		WTP elicited in the baseline phase		WTP elicited in the exogenous treatment phase		
		Mean	95%-Conf. int.	Mean	95%-Conf. int.	Obs.
WTP in the control group						
	Average WTP amounts					
V	WTP for SC	9.64	[8.26 – 11.03]	9.44	[8.18 – 10.70]	118
V	WTP for RTF	10.80	[9.63 – 11.97]	10.25	[9.01 – 11.50]	118
V	WTP for BOTH	10.58	[9.29 – 11.88]	10.14	[8.90 – 11.39]	118
	Share of participants with negative WTP					
VI	WTP for SC	7.6%	[2.8% – 12.5%]	4.2%	[0.5% – 7.9%]	118
VI	WTP for RTF	3.4%	[0.1% – 6.7%]	4.2%	[0.5% – 7.9%]	118
VI	WTP for BOTH	5.9%	[1.6% – 10.3%]	5.9%	[1.6% – 10.3%]	118
Difference in differences analysis of treatment experience on WTP						
	Average WTP amounts					
VII	WTP for SC			-0.14	[-2.07 – 1.79]	878
VII	WTP for RTF			0.05	[-1.79 – 1.88]	878
VII	WTP for BOTH			-1.02	[-2.51 – 0.47]	878
	Share of participants with negative WTP					
VIII	WTP for SC			6.1%	[-2.9% – 15.1%]	878
VIII	WTP for RTF			-2.8%	[-10.5% – 5.0%]	878
VIII	WTP for BOTH			0.0%	[-7.0% – 7.0%]	878

Note: The sample in panels I and II consists of all participants who filled out both WTP inquiries and who were not part of the super-control group, as the WTP inquiries for the super-control group were purely hypothetical. Six participants did not fill out any of the WTP inquiries, and three of the remaining participants did not fill out the second WTP inquiry. The sample in panels III and IV consists of those who filled out both WTP inquiries and who received the treatment for which the WTP is reported in the respective row in the exogenous treatment phase. The sample in panels V and VI consists of all participants who filled out both WTP inquiries and who were part of the control group (not the super-control group) in the exogenous treatment phase. The sample in panels VII and VIII is the same as in panels I and II, but it combines the data from the first and the second WTP elicitation. The depicted coefficients are estimates using a difference in differences regression, controlling for individual fixed effects.

An important aspect of the WTP is what the participants think they can save financially through the interventions. In the next step, we therefore leverage a survey question about the perceived monetary savings that we elicited at the end of the exogenous treatment period.[19] On average, participants in the SC group reported perceived savings of 4.12 EUR in the exogenous treatment month. Average perceived savings in the RTF group were 5.48 EUR and perceived savings in the BOTH group were 6.23 EUR. Using the average number of 31 showers per household in the exogenous treatment phase, the treatment effects in liters per shower from Table 3, and an estimated cost of about 1 Eurocent per liter of warm shower water, we estimate the actual monetary savings for one month of treatment to be around 1 EUR for SC, 3 EUR for RTF, and 3.7 EUR for BOTH. This indicates that participants tended to overestimate the monetary savings from the interventions.

Table 8: Comparison of willingness to pay with perceived savings for those experiencing treatment

Treat. (N)	Post-treatment WTP		Perceived savings		Net non-monetary value		Share negative net non-monetary value	
SC (89)	8.59	[7.21 – 9.97]	4.12	[2.73 – 5.52]	4.34	[2.24 – 6.43]	28.1%	[18.6% – 37.6%]
RTF (92)	9.13	[7.54 – 10.72]	5.48	[3.90 – 7.05]	3.59	[1.35 – 5.83]	25.0%	[16.0% – 34.0%]
BOTH (93)	9.67	[8.34 – 11.01]	6.23	[4.85 – 7.61]	3.38	[1.60 – 5.17]	26.9%	[17.7% – 36.1%]

Note: In contrast to Table 7, only participants who answered the survey question eliciting the perceived savings are included in this table.

In Table 8, we contrast the perceived savings with WTP stated in the second WTP elicitation by those who experience the respective treatment. Subtracting the perceived savings from WTP provides us with an estimate of the net non-monetary value of the interventions (third column of Table 8). While this net non-monetary value is significantly positive on average, the share of participants with negative net non-monetary value is between 25% and 28% (fourth column), with no significant difference between the interventions.

[19]The question read: "Do you think INTERVENTION has helped you reduce your electric and water bills?" and if yes "Please estimate: "By how many euros did your energy and water costs decrease last month thanks to the INTERVENTION?"; where INTERVENTION was the specific intervention experienced by the respondent during the exogenous treatment phase.

3.4 Impact of the interventions on welfare

To estimate the effect of the interventions on welfare, we need to consider not only the change in consumer surplus but also the changes in externalities and the implementation costs of the interventions. In doing so, we lean on the approach by Allcott and Kessler (2019) and calculate the change in welfare ΔW as follows:

$$\Delta W_j = \Delta V_j - \phi_R \times \Delta R_j - (MCF - MU_j) \times c_j, \tag{2}$$

where ΔV_j is the experienced change in consumer surplus, which we measure as WTP elicited after treatment experience of intervention j.[20] ΔR_j is the change in resource use while showering, i.e., the treatment effect of intervention j. ϕ_R is the external effect associated with one unit of resource use and c_j is the cost associated with the intervention. Assuming that these costs are borne by the government, we multiply the cost of the intervention with the marginal costs of public funds (MCF). Since the revenues of the firms providing the interventions do not represent a welfare loss, the unit cost of the interventions must be reduced by the per-unit markup (MU_j) on the retail price of the interventions.

We use three scenarios to parameterize the welfare equation. The first one represents the "base scenario", which relates directly to the average results of our experiment, the second one represents a "conservative" scenario, and the third is an "opt-in" scenario, in which we assume that only those receive the interventions who have a positive WTP for the interventions and who are therefore likely to opt-in for an intervention even in a completely voluntary setting.

For the base scenario, the values for ΔV_j are provided in Table 7 (WTP elicited in the exogenous treatment phase). With regard to the external effects, we focus on the effects of carbon emissions associated with the energy use required to heat water for showering. Based on the average shower temperature of 38 °C and averages for the

[20]Consumer surplus is defined as the difference between maximum WTP and the price paid. We assume in our welfare analysis that the interventions are tax-funded and therefore consumers do not bear a direct cost, which is why we consider WTP to be an appropriate measure of consumer surplus.

type of energy used for water heating in Germany, we assume that one liter of water use is associated with 15 grams of CO_2 (see Appendix A6 for details). To arrive at a monetary value, we multiply this with the social cost of carbon estimated in Rennert et al. (2022) amounting to 185 Dollars / 160 EUR per ton of CO_2.[21] ΔR_j is the effect of the respective intervention in liters per shower taken from Table 3 and multiplied by 31, which is the average number of showers taken in the exogenous treatment month. The costs associated with the intervention are not easily quantified. To approximate these costs, we assume the retail costs of the smart shower heads of 80 EUR and a lifetime per shower head of 5 years. Thus, we assume costs per month of $80/(5 \times 12) = 1.33$ EUR. Operational costs are abstracted from, as they are expected to be low on a large scale since the interventions can be fully automated. For the marginal costs of public funds (MCF), we assume the factor of 1.85 found by Kleven and Kreiner (2006).[22] We assume a 35% per-unit markup on the retail price, which is the value that De Loecker and Eeckhout (2018) estimate as the average markup for Germany in 2016.

The results of the welfare change due to the interventions (ΔW) in the base scenario are depicted in the upper panel of Table 9 and indicate that the average welfare effect of SC is 6.83 EUR per month, the effect of RTF is 7.86 EUR per month and the effect of BOTH is largest at 8.56 EUR per month. This means that all three interventions lead to substantial welfare gains.

For the conservative scenario, we assume that markups (MU_j) are either zero or do not contribute to welfare. In addition, we reduce our emission factor for one liter of hot water consumption by 16%, which is the share of energy used to heat water in private households that comes from renewable energy or district heating (BMWi, 2020) and that thus is associated with low or zero carbon emissions.[23] Recognizing that our sample of volunteers may be particularly motivated to conserve resources and

[21] Assuming the long-term average exchange rate of Dollars to Euros over the past 10 years, which is 0.8627 (ECB, 2023). This calculation results in external costs of 0.0024 EUR / liter.

[22] Kleven and Kreiner (2006) estimate the MCF based on several labor supply elasticity scenarios. We use the result of Scenario 6, which they denote as the "natural baseline scenario".

[23] While district heating does in principle entail carbon emissions, there may be cases where it is fed entirely by industrial waste heat, so that the use of this waste heat does not result in further emissions, which is what we assume in the conservative scenario. This calculation results in external costs of 0.002 EUR / liter.

Table 9: Welfare effects per month of intervention

	ΔV	$\phi_R \times \Delta R$	$(MCF - MU) \times c$	ΔW	$MVPF$
Base scenario					
SC	8.59	−0.24	2.00	6.83 EUR	6.62
RTF	9.13	−0.73	2.00	7.86 EUR	7.39
BOTH	9.67	−0.89	2.00	8.56 EUR	7.92
Conservative scenario					
SC	2.42	−0.08	2.47	0.04 EUR	1.87
RTF	2.57	−0.24	2.47	0.34 EUR	2.11
BOTH	2.72	−0.29	2.47	0.54 EUR	2.27
Opt-in scenario					
SC	11.98	−0.26	2.00	10.24 EUR	9.18
RTF	12.21	−0.73	2.00	10.94 EUR	9.70
BOTH	11.85	−0.88	2.00	10.73 EUR	9.55

Note: Welfare changes and MVPF are calculated as follows: $\Delta W = \Delta V - \phi_R \times \Delta R - (MCF - MU) \times c$, $MVPF = \frac{\Delta V - \phi_R \times \Delta R}{c}$.

to receive the interventions, we scale our WTP values and treatment effects down. To this end, we compare our average WTP estimate for the SC intervention, which is 8.59 EUR, to the average WTP in Allcott and Kessler (2019), which to our knowledge is the only other study that has elicited WTP for a social comparison intervention. They find that the average WTP for receiving four home energy reports on residential natural gas consumption is 2.81 Dollars, or 2.42 EUR,[24] which is 72% lower than the WTP for the SC intervention in our study. Since there are no comparable WTP estimates available for the RTF and BOTH interventions, we reduce all of our WTP values in the conservative scenario by 72%. To scale down the treatment effects, we draw on the results by Tiefenbeck et al. (2019), who conducted a study on the effectiveness of RTF while showering among hotel guests, i.e., in a population without volunteer selection bias. They find a conservation effect of RTF of 11.4%. Given the control group's water use of 34.1 liters during the exogenous treatment period, an 11.4% reduction would correspond to water savings of 3.89 liters per shower, which is 60% less than the effect we found in the exogenous treatment phase (Table 3). In our conservative scenario, we thus assume 3.89 liters as the treatment effect of RTF and also reduce the effects of the

[24] Assuming the long-term average exchange rate of Dollars to Euros over the past 10 years, which is 0.8627 (ECB, 2023).

other interventions by 60.1%, resulting in a SC effect of 1.27 liters and a BOTH effect of 4.73 liters per shower.

The results of the conservative scenario, which we interpret as a lower bound on the average welfare effects of the interventions, are shown in the middle panel of Table 9 and indicate that all three interventions still yield modest welfare gains ranging from 4 cents (SC) to 54 cents (BOTH) per month even under conservative assumptions.

For the opt-in scenario, we use the same parameters as in the base scenario. However, we prune the sample so that only individuals with a positive WTP after treatment experience remain in the sample.[25] The results of the opt-in scenario are shown in the bottom panel of Table 9 and can be interpreted as an upper bound for the welfare effects. As expected, these are much larger than in the base scenario, ranging from 10.24 (SC) to 10.94 EUR (RTF) per month. However, these welfare effects are not average effects for the entire sample, but only for those with positive WTP and thus for a smaller overall group.

In addition to the conventionally calculated welfare effects of the interventions (ΔW), we also report the marginal value of public funds (MVPF) in Table 9. This statistic, the use of which is encouraged by Hendren and Sprung-Keyser (2020), uses the same input values as ΔW, but standardizes the welfare gains, or in Hendren and Sprung-Keyser's (2020) terms, the societal willingness to pay, by the cost of the policy. Thus, this statistic provides us with an estimate of the social benefits per unit of public funds, which has the advantage of allowing a unified analysis of the welfare effects of different policies and making them directly comparable across domains.

We find that, with the exception of the values in the conservative scenario, all MVPFs exceed the value of 5. This is for example in the order of magnitude of several measures targeting children's health and education outcomes, which are among the most welfare-efficient social policy measures analyzed by Hendren and Sprung-Keyser (2020). Concluding, the MVPF analysis underlines that the interventions in our experiment promise large welfare gains.

[25] Table A4 in the Appendix reports the treatment effects for this sub sample, which we use to calculate ΔR in this scenario.

4 Conclusion

This paper provides evidence comparing popular behavioral economic interventions in terms of resource conservation and welfare impacts. In a field experiment, we compare two information interventions aimed at reducing energy and water consumption while showering and their combination. One intervention provides social comparison information motivating consumers to conserve, while the other provides real-time information while showering, facilitating resource conservation for participants. Both interventions are highly effective in reducing resource use, but the facilitating intervention and the combination of both interventions lead to significantly greater reductions than the motivating intervention alone. Participants' revealed willingness to pay for the interventions and our welfare analysis show that both intervention types improve welfare, but that the facilitating and the combined intervention do so to a greater extent. Thus, our findings from a resource conservation setting suggest that information that facilitates behavior change is more effective and welfare-enhancing than information that motivates behavior change.

Our results are in line with previous findings highlighting the importance of making shrouded attributes salient to consumers (e.g., Gabaix and Laibson, 2006; Chetty et al., 2009; DellaVigna and Pollet, 2009; Grubb, 2009; Brown et al., 2010; Taubinsky and Rees-Jones, 2018). Our findings corroborate that in environments such as showering and energy use in the home, quantity consumed is a shrouded attribute. Making quantities salient to consumers removes barriers to behavioral change, facilitating reactions to prices (Jessoe and Rapson, 2014) and knowledge about environmental impact (Fang et al., 2021).

Our findings also showcase the potential for behavioral change in consumption settings in which fine-grained data can be made available. While information interventions motivating behavioral change, such as social comparison reports (Allcott and Mullainathan, 2010; Benartzi et al., 2017), can achieve effects in some settings even in the absence of fine-grained data, appliance-specific data has the potential to generate much larger effects. Increasing technological development and technological diffusion is making low-cost solutions to assessing real-time information more available than ever before. For example, a large German retailer has recently started selling a RTF shower head at very affordable prices (Tchibo, 2023). Policymakers should be aware of these developments and the potential they hold for behavioral interventions made possible through such technology.

While the application we examine in detail – showering behavior – is in itself an everyday and energy-intensive behavior and thus important to consider in energy conservation efforts, there is potential for future research to assess whether the pattern we observe also holds for other water- and energy-intense behaviors. For example, future research could compare motivating and facilitating information interventions to address consumers' room heating and cooling behaviors. We would expect patterns to be similar since the energy used for room heating and cooling is – similar to the energy used for heating water when showering – a shrouded, non-salient attribute to the consumer. We thus expect a similar energy-saving potential for information interventions making this attribute salient.

In contrast, the distinction between motivating and facilitating information interventions might be less pronounced in an environment where salience bias is less of an issue. This would apply to environments in which consumers are better able to develop heuristics for their consumption or where the consumption itself is more salient (e.g., taking a plane). Comparing motivating and facilitating information interventions in such applications would be an interesting avenue for future research.

More generally, our findings contribute to the still very limited literature assessing the welfare effects of (behavioral) interventions by directly eliciting consumers' willingness to pay for experiencing the interventions (Allcott and Kessler, 2019; Allcott et al., 2022; Butera et al., 2022). In our setting, the behavioral interventions on average create a win-win situation: They reduce environmental externalities and simultaneously create a psychological benefit to consumers. We also show that the intervention experience did not change willingness to pay.

References

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, 95(9-10):1082–1095.

Allcott, H., Cohen, D., Morrison, W., and Taubinsky, D. (2022). When do "nudges" increase welfare? National Bureau of Economic Research (NBER) Working Paper No. w30740.

Allcott, H. and Kessler, J. B. (2019). The welfare effects of nudges: A case study of energy use social comparisons. *American Economic Journal: Applied Economics*, 11(1):236–76.

Allcott, H. and Mullainathan, S. (2010). Behavior and energy policy. *Science*, 327(5970):1204–1205.

Allcott, H. and Rogers, T. (2014). The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review*, 104(10):3003–37.

Andor, M. A. and Fels, K. M. (2018). Behavioral economics and energy conservation – a systematic review of non-price interventions and their causal effects. *Ecological Economics*, 148:178–210.

Andor, M. A., Gerster, A., Peters, J., and Schmidt, C. M. (2020). Social norms and energy conservation beyond the US. *Journal of Environmental Economics and Management*, 103:102351.

Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434):444–455.

Ariely, D. and Jones, S. (2008). *Predictably irrational*. HarperCollins New York.

Asensio, O. I. and Delmas, M. A. (2015). Nonprice incentives and energy conservation. *Proceedings of the National Academy of Sciences*, 112(6):E510–E515.

Benartzi, S., Beshears, J., Milkman, K. L., Sunstein, C. R., Thaler, R. H., Shankar, M., Tucker-Ray, W., Congdon, W. J., and Galing, S. (2017). Should governments invest more in nudging? *Psychological Science*, 28(8):1041–1055.

BiB (2022). Durchschnittsalter der Bevölkerung in Deutschland (1871-2019). Bundesinstitut für Bevölkerungsforschung. `https://www.bib.bund.de/DE/Fakten/Fakt/B19-Durchschnittsalter-Bevoelkerung-ab-1871.html`, accessed on March 3, 2022.

BiB (2023). Zahl der Privathaushalte und durchschnittliche Haushaltsgröße in Deutschland (1871-2021). Bundesinstitut für Bevölkerungsforschung. `https://www.bib.bund.de/DE/Fakten/Fakt/L49-Privathaushalte-Haushaltsgroesse-ab-1871.html`, accessed on February 2, 2023.

BMWi (2020). Energiedaten. As of: May 2020. Table 7b: Endenergieverbrauch nach Anwendungsbereichen II. Bundesministerium für Wirtschaft und Energie [Federal Ministry for Economic Affairs and Energy]. `https://www.bmwk.de/Redaktion/DE/Artikel/Energie/energiedaten-gesamtausgabe.html`, originally retrieved on September 22, 2020; the URL was last accessed on January 26, 2023 and contains a newer version of the document.

Brandon, A., List, J. A., Metcalfe, R. D., Price, M. K., and Rundhammer, F. (2019). Testing for crowd out in social nudges: Evidence from a natural field experiment in the market for electricity. *Proceedings of the National Academy of Sciences*, 116(12):5293–5298.

Brent, D. A., Cook, J. H., and Olsen, S. (2015). Social comparisons, household water use, and participation in utility conservation programs: Evidence from three randomized trials. *Journal of the Association of Environmental and Resource Economists*, 2(4):597–627.

Brown, J., Hossain, T., and Morgan, J. (2010). Shrouded attributes and information suppression: Evidence from the field. *Quarterly Journal of Economics*, 125(2):859–876.

Brülisauer, M., Goette, L., Jiang, Z., Schmitz, J., and Schubert, R. (2020). Appliance-specific feedback and social comparisons: Evidence from a field experiment on energy conservation. *Energy Policy*, 145:111742.

Butera, L., Metcalfe, R., Morrison, W., and Taubinsky, D. (2022). Measuring the welfare effects of shame and pride. *American Economic Review*, 112(1):122–68.

Byrne, D. P., Goette, L., Martin, L. A., Delahey, L., Jones, A., Miles, A., Schob, S., Staake, T., Tiefenbeck, V., et al. (2021). The habit-forming effects of feedback: Evidence from a large-scale field experiment. CRC TR 224 Discussion Paper No. 185. University of Bonn and University of Mannheim.

Cadario, R. and Chandon, P. (2020). Which healthy eating nudges work best? A meta-analysis of field experiments. *Marketing Science*, 39(3):465–486.

Carson, R. T. and Groves, T. (2007). Incentive and informational properties of preference questions. *Environmental and Resource Economics*, 37(1):181–210.

Chetty, R., Looney, A., and Kroft, K. (2009). Salience and taxation: Theory and evidence. *American Economic Review*, 99(4):1145–77.

Congiu, L. and Moscati, I. (2020). Message and environment: A framework for nudges and choice architecture. *Behavioural Public Policy*, 4(1):71–87.

De Loecker, J. and Eeckhout, J. (2018). Global market power. National Bureau of Economic Research (NBER) Working Paper No. w24768.

DellaVigna, S., List, J. A., and Malmendier, U. (2012). Testing for altruism and social pressure in charitable giving. *Quarterly Journal of Economics*, 127(1):1–56.

DellaVigna, S., List, J. A., Malmendier, U., and Rao, G. (2016). Voting to tell others. *Review of Economic Studies*, 84(1):143–181.

DellaVigna, S. and Pollet, J. M. (2009). Investor inattention and friday earnings announcements. *Journal of Finance*, 64(2):709–749.

Destatis (2022a). Bevölkerung und Erwerbstätigkeit - 2020 (Endergebnisse) - Fachserie 1 Reihe 3. Statistisches Bundesamt (Destatis). `https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Haushalte-Familien/Publikationen/Downloads-Haushalte/haushalte-familien-endergebnisse-2010300207004.pdf;jsessionid=392F985EFEDBA446D5AD6BF47C06CAB0.live741?__blob=publicationFile`, accessed on March 3, 2022.

Destatis (2022b). Laufende Wirtschaftsrechnungen: Einkommen, Einnahmen und Ausgaben privater Haushalte - 2021 - Fachserie 15 Reihe 1. Statistisches Bundesamt (Destatis). `https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Einkommen-Konsum-Lebensbedingungen/Einkommen-Einnahmen-Ausgaben/Publikationen/Downloads-Einkommen/einnahmen-ausgaben-privater-haushalte-2150100217004pdf?__blob=publicationFile`, accessed on February 13, 2023.

ECB (2023). Euro foreign exchange reference rates: US dollar (USD). Change from 24 January 2013 to 24 January 2023. European Central Bank (ECB). `https://www.ecb.europa.eu/stats/policy_and_exchange_rates/euro_reference_exchange_rates/html/eurofxref-graph-usd.en.html`, accessed on January 24, 2023.

Faberi, S., Paolucci, L., Lapillonne, B., and Pollier, K. (2015). Trends and policies for energy savings and emissions in transport. Report, ODYSSEE-MURE project. `https://www.odyssee-mure.eu/publications/archives/energy-efficiency-trends-policies-transport.pdf`, accessed on October 12, 2023.

Fang, X., Goette, L., Rockenbach, B., Sutter, M., Tiefenbeck, V., Schoeb, S., Staake, T., et al. (2021). Complementarities in behavioral interventions: Evidence from a field experiment on energy conservation. CRC TR 224 Discussion Paper No. 149. University of Bonn and University of Mannheim.

Ferraro, P. J. and Price, M. K. (2013). Using nonpecuniary strategies to influence behavior: evidence from a large-scale field experiment. *Review of Economics and Statistics*, 95(1):64–73.

Gabaix, X. and Laibson, D. (2006). Shrouded attributes, consumer myopia, and information suppression in competitive markets. *Quarterly Journal of Economics*, 121(2):505–540.

Gerster, A., Andor, M. A., and Goette, L. (2021). Disaggregate consumption feedback and energy conservation. *SFB 823 Discussion Paper #14*.

Goette, L., Han, H.-J., Lim, Z. H., et al. (2021a). The dynamics of goal setting: Evidence from a field experiment on resource conservation. CRC TR 224 Discussion Paper No. 283. University of Bonn and University of Mannheim.

Goette, L., Jiang, Z., Schmitz, J., and Schubert, R. (2021b). The effects of upward and downward social comparisons on energy consumption behavior. In Quah, E. and Schubert, R., editors, *Sustainability and Environmental Decision Making*, pages 409–440. Springer, Singapore.

Goette, L., Leong, C., and Qian, N. (2019). Motivating household water conservation: A field experiment in Singapore. *PLOS ONE*, 14(3):e0211891.

Grubb, M. D. (2009). Selling to overconfident consumers. *American Economic Review*, 99(5):1770–1807.

Grubb, M. D. and Osborne, M. (2015). Cellular service demand: Biased beliefs, learning, and bill shock. *American Economic Review*, 105(1):234–271.

Halpern, D. (2015). *Inside the nudge unit: How small changes can make a big difference*. Random House.

Hansen, P. G. and Jespersen, A. M. (2013). Nudge and the manipulation of choice: A framework for the responsible use of the nudge approach to behaviour change in public policy. *European Journal of Risk Regulation*, 4(1):3–28.

Hendren, N. and Sprung-Keyser, B. (2020). A unified welfare analysis of government policies. *Quarterly Journal of Economics*, 135(3):1209–1318.

Ito, K. (2014). Do consumers respond to marginal or average price? evidence from nonlinear electricity pricing. *American Economic Review*, 104(2):537–563.

Ito, K., Ida, T., and Tanaka, M. (2018). Moral suasion and economic incentives: Field experimental evidence from energy demand. *American Economic Journal: Economic Policy*, 10(1):240–267.

Jachimowicz, J. M., Duncan, S., Weber, E. U., and Johnson, E. J. (2019). When and why defaults influence decisions: A meta-analysis of default effects. *Behavioural Public Policy*, 3(2):159–186.

Jessoe, K. and Rapson, D. (2014). Knowledge is (less) power: Experimental evidence from residential energy use. *American Economic Review*, 104(4):1417–1438.

Johnson, E. J., Shu, S. B., Dellaert, B. G., Fox, C., Goldstein, D. G., Häubl, G., Larrick, R. P., Payne, J. W., Peters, E., Schkade, D., et al. (2012). Beyond nudges: Tools of a choice architecture. *Marketing Letters*, 23:487–504.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Khanna, T. M., Baiocchi, G., Callaghan, M., Creutzig, F., Guias, H., Haddaway, N. R., Hirth, L., Javaid, A., Koch, N., Laukemper, S., Löschel, A., Zamora Dominguez, M. d. M., and Minx, J. C. (2021). A multi-country meta-analysis on the role of behavioural change in reducing energy consumption and CO2 emissions in residential buildings. *Nature Energy*, 6(9):925–932.

Kleven, H. J. and Kreiner, C. T. (2006). The marginal cost of public funds: Hours of work versus labor force participation. *Journal of Public Economics*, 90(10-11):1955–1973.

Lacetera, N., Pope, D. G., and Sydnor, J. R. (2012). Heuristic thinking and limited attention in the car market. *American Economic Review*, 102(5):2206–2236.

List, J. A., Muir, I., Pope, D., and Sun, G. (2023). Left-digit bias at lyft. *Review of Economic Studies*, rdad014.

Löschel, A., Rodemeier, M., and Werthschulte, M. (2022). Can self-set goals encourage resource conservation? Field experimental evidence from a smartphone app. ZEW - Centre for European Economic Research Discussion Paper No. 20-039.

Mertens, S., Herberz, M., Hahnel, U. J., and Brosch, T. (2022). The effectiveness of nudging: A meta-analysis of choice architecture interventions across behavioral domains. *Proceedings of the National Academy of Sciences*, 119(1):e2107346118.

Michel, A., Attali, S., and Bush, E. (2016). Energy efficiency of white goods in Europe: monitoring the market with sales data. Final report. Study realised on behalf of ADEME by SOWATT and Bush Energie. `https://storage.topten.eu/source/files/Market-Monitoring-2016-EN-Topten.eu.pdf`, accessed on October 12, 2023.

Mongin, P. and Cozic, M. (2018). Rethinking nudge: Not one but three concepts. *Behavioural Public Policy*, 2(1):107–124.

Nisa, C. F., Bélanger, J. J., Schumpe, B. M., and Faller, D. G. (2019). Meta-analysis of randomised controlled trials testing behavioural interventions to promote household action on climate change. *Nature Communications*, 10(1):1–13.

Obama, B. (2015). Executive Order – Using Behavioral Science Insights to Better Serve the American People. Media Release, White House, Washington, DC, September 15, 2015.

OECD (2017). *Behavioural insights and public policy: Lessons from around the world*. OECD Publishing, Paris.

Rennert, K., Errickson, F., Prest, B. C., Rennels, L., Newell, R. G., Pizer, W., Kingdon, C., Wingenroth, J., Cooke, R., Parthum, B., Cromar, K., Diaz, D., Moore, F. C., Müller, U. K., Plevin, R. J., Raftery, A. E., Sevcikova, H., Sheets, H., Stock, J. H., Tan, T., Watson, M., Wong, T. E., and Anthoff, D. (2022). Comprehensive evidence implies a higher social cost of CO2. *Nature*, 610(7933):687–692.

Rodemeier, M. (2020). Buy baits and consumer sophistication: Theory and field evidence from large-scale rebate promotions. CAWM Working Paper 124.

Rodemeier, M. and Löschel, A. (2022). Information nudges, subsidies, and crowding out of attention: Field evidence from energy efficiency investments. Available at SSRN: `https://ssrn.com/abstract=4213071orhttp://dx.doi.org/10.2139/ssrn.4213071`.

Sharot, T. and Sunstein, C. R. (2020). How people decide what they want to know. *Nature Human Behaviour*, 4(1):14–19.

Sheeran, P. and Webb, T. L. (2016). The intention–behavior gap. *Social and Personality Psychology Compass*, 10(9):503–518.

Taubinsky, D. and Rees-Jones, A. (2018). Attention variation and welfare: Theory and evidence from a tax salience experiment. *Review of Economic Studies*, 85(4):2462–2496.

Tchibo (2023). LED shower head with water consumption indicator. Tchibo GmbH online store. `https://www.tchibo.de/products/402100280/led-duschkopf-mit-wasserverbrauchsanzeige`, accessed on September 19, 2023.

Thaler, R. H. and Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.

Tiefenbeck, V., Goette, L., Degen, K., Tasic, V., Fleisch, E., Lalive, R., and Staake, T. (2018). Overcoming salience bias: How real-time feedback fosters resource conservation. *Management Science*, 64(3):1458–1476.

Tiefenbeck, V., Wörner, A., Schöb, S., Fleisch, E., and Staake, T. (2019). Real-time feedback promotes energy conservation in the absence of volunteer selection bias and monetary incentives. *Nature Energy*, 4(1):35–41.

UBA (2019). Emissionsbilanz erneuerbarer Energieträger: Bestimmung der vermiedenen Emissionen im Jahr 2018. Table 62: Emission factors of heat supply from fossil fuels in private households, the tertiary sector and industry. Umweltbundesamt [German Environment Agency]. https://www.umweltbundesamt.de/sites/default/files/medien/1410/publikationen/2019-11-07_cc-37-2019_emissionsbilanz-erneuerbarer-energien_2018.pdf, accessed on September 22, 2020.

VZ-RLP (2020). Warmwasser: Komfortables Sparen – So geht's! As of: February 2020. Verbraucherzentrale Rheinland-Pfalz [Consumer Organisation of Rhineland-Palatinate]. https://www.verbraucherzentrale-rlp.de/sites/default/files/2020-03/VZ_WarmWasser_2020_12_web.pdf, accessed on September 22, 2020.

Appendix

A1 Supplementary analyses

A1.1 Further analyses on the impact of the interventions behavior

A1.1.1 Robustness checks

Table A1: Treatment effects robust to different specifications

	Volume (Liters)		
SC	-3.79**	-4.66**	-3.19***
	(1.55)	(1.98)	(0.66)
RTF	-9.91***	-10.79***	-9.79***
	(1.26)	(1.68)	(0.88)
BOTH	-12.01***	-12.88***	-11.91***
	(1.33)	(1.74)	(1.22)
Time FE	No	Yes	Yes
Household FE	No	No	Yes
No. of households	564	564	564
No. of observations	38,453	38,453	38,453

Note: Standard errors are clustered at the household level and are reported in parentheses. ***, *** and ** denote statistical significance at the 1%, 5% and 10% level, respectively.

Table A2: Negligible effect on shower temperature and frequency

	Shower temperature	Showers per day
SC	-0.17	0.05
	(0.15)	(0.04)
RTF	-0.19	-0.03
	(0.18)	(0.03)
BOTH	-0.29*	-0.02
	(0.15)	(0.03)
No. of households	564	564
No. of observations	38,560	32,148

Note: OLS estimates with household and time fixed effects included. For Spec. (2), we construct a data set including the number of showers per household per day during the baseline phase and exogenous phase of the study. Standard errors are clustered at the household level and are reported in parentheses. ***, *** and ** denote statistical significance at the 1%, 5% and 10% level, respectively.

A1.1.2 Analysis of the transition between color thresholds due to the interventions

Table A3: Transition matrix: Changes in the color threshold in which a participant's showers ended most often

			to					
			green	blue	purple	red	flashing red	N
Cont	**from**	green	92%				5%	22
		blue	30%	50%	10%	10%		10
		purple	17%	8%	58%		17%	12
		red	6%		19%	19%	56%	16
		flashing red	4%		2%	2%	92%	48
SC	**from**	green	92%	3%		3%	3%	39
		blue	83%	17%				6
		purple	17%	42%	25%	8%	8%	12
		red	20%	40%	40%			5
		flashing red	8%	8%	2%		81%	48
RTF	**from**	green	100%					31
		blue	89%	11%				9
		purple	22%	44%	22%		11%	9
		red	22%	56%	11%	11%		9
		flashing red	27%	6%	16%	8%	43%	51
BOTH	**from**	green	100%					25
		blue	87%	13%				15
		purple	67%	33%				12
		red	33%	33%	33%			3
		flashing red	28%	19%	9%	6%	39%	54

Note: We present the changes between the baseline phase and the exogenous treatment phase. The sum of the individual rows is 100% (if not, this is due to rounding error). As an example, we can interpret the last row of the table as follows: Of the 54 participants in the BOTH group who finished most of their showers in the baseline phase while the shower head would have flashed red, 28% (19%, 9%, 6%, 39%) finished most showers in the exogenous treatment phase while the shower head was green (blue, purple, red, flashing red). We calculated the hypothetical color thresholds in the baseline phase as we did in the treatment phases, that is, we based them on the average baseline water consumption per shower of a participant's reference group, which consisted of participants from the super-control group. Therefore, participants in the super-control group are not included in this table because there are no valid hypothetical color thresholds available for them.

A1.1.3 Details on treatment assignment in the endogenous phase and the IV approach

To explain the functioning of our instrumental variables, we first explain the details of treatment assignment in the endogenous treatment phase:

1. It was randomly drawn which of the two WTP inquiries was relevant for the allocation.

2. From the relevant WTP inquiry, it was randomly drawn which of the three MPLs was relevant for the allocation, i.e., whether the MPL concerning WTP for SC, RTF or BOTH was the relevant MPL.

3. One of the 15 rows in the relevant MPL, i.e., one decision, was randomly selected. If the participant chose the intervention in this row the participant received the intervention in the endogenous treatment phase and if the participant did not choose the intervention, she was allocated to the control group in the endogenous treatment phase.

Given this treatment assignment mechanism, we can divide our participants into two groups: The first group is the group of "compliers", i.e., those who received the intervention randomly drawn in step 2 of the treatment assignment mechanism. This group is composed of two subgroups, the first consisting of those participants who reported the maximum WTP for all three interventions in both WTP surveys, which is the case for 43% of our participants, and the second consisting of those who reported different WTP across interventions or WTP inquiries but whose decision in the specific MPL row drawn in step 3 of the assignment mechanism was in favor of the intervention. The second group consists of the so-called "never takers", who are those who were assigned a specific intervention in step 2 of the allocation mechanism, but decided against receiving the intervention in the MPL row drawn in step 3. Our approach of using the random draws from all three steps of the assignment mechanism as instrumental variables allows us to estimate the causal local average treatment effect (LATE) for the group of compliers (Angrist et al., 1996).

A1.1.4 Treatment effect for the opt-in scenario (sub sample of those with WTP>0)

Table A4: Average treatment effects on water use per shower in the opt-in scenario (sub sample of those with WTP>0)

	(I) Volume (Liters)
SC	-3.46***
	(0.85)
RTF	-9.79***
	(1.04)
BOTH	-11.79***
	(1.44)
Control group average	33
No. of households	355
No. of observations	25,323

Note: OLS estimates with household and time fixed effects included. The results show treatment effects in the exogenous treatment phase for households who would choose to take up an intervention in an opt-in scenario, i.e., the estimation sample consists of those who stated a positive WTP after treatment experience or who are part of the control group in the exogenous treatment phase and who stated a positive WTP for the SC intervention in the second WTP inquiry. The super-control group is not part of the estimation sample. Standard errors clustered at the household level. ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

A2 Invitation for study participation

Dear XYZ!

With this email we would like to invite you to participate in a new study by our long-time scientific partner, RWI - Leibniz Institute for Economic Research*. The study is financed by the Ministry of Culture and Science of the State of North Rhine-Westphalia and examines ways to reduce resource consumption. In a first step, only our panelists have the opportunity to participate.

WHAT EXACTLY IS THE STUDY ABOUT?

As a core element of the study, participants will be provided with a high-quality shower head that measures water consumption per shower. During the four-month study phase, this information will be automatically shared with RWI. Additionally, you will be provided with various information about your showering behavior during the course of the study. At the end of the study, you will be able to keep the shower head, which then ceases to share information. This will allow you to keep track of your water consumption yourself on a permanent basis.

HOW CAN I SIGN UP?

You can find all further information and registration form on the project website: [URL]

YOUR BENEFITS

As an **expense allowance**, you will receive vouchers or additional resource saving tools worth a total of € 35 in addition to the shower head before the end of the study. You will receive these vouchers as part of your participation in short, monthly surveys conducted by RWI to accompany the study. At the end of the study, we will offer you a selection of popular suppliers where you can redeem your vouchers. Furthermore, you will of course be informed about the results of the study in the form of a short report. So again in a nutshell:

- Modern shower head with innovative functions
- Vouchers as an expense allowance - choose your preferred provider from our list of providers
- Report with the results of the study
- Participation in a study, the findings of which will make the world a little bit more sustainable

WHY SHOWERS?

What's not commonly known is that showering involves a relatively high consumption of resources, as not only water but also significant amounts of energy are used to heat the water. For example, with the energy consumed by an ordinary shower, you could light an average household for two and a half days, or watch TV for 26 hours.

HOW IS MY DATA PROTECTED?

The study is subject to strict data protection regulations. More detailed information can be found on the project website: [LINK].

DON'T FORGET TO REGISTER!

You can register on the project homepage [LINK].

If you have any questions about the study, please contact the project team at [E-MAIL ADDRESS].

Your forsa.omninet Team

* RWI - Leibniz Institute for Economic Research in Essen is an independent, non-profit scientific research institution. As one of the leading economic research institutes in Germany, the institute is a member of the Leibniz Association.

A3 Website to register for the study

The project website aimed at participants included the following subpages:

- **Home page:** Welcome message and general introduction to the study
- **Project goals:** Short overview of the experiment's objectives
- **Your contribution:** Information on environmental potential of the study
- **FAQ:** Answers to typical questions
- **Who we are:** Short overview of the research team
- **Sign up:** Overview of the general terms & conditions and signup form

The following subsection features a translated version of selected pages.

A3.1 Website text content

Home Page

Welcome to the "sustainable showering" project homepage!

We are pleased that you are interested in our research project! On this website, we would like to introduce the project to you and hope to win you as a study participant.

This project is a scientific research project of RWI - Leibniz Institute for Economic Research (Essen) in research partnership with the Center for Advanced Internet Studies (CAIS).

The project is funded by the Ministry of Culture and Science of the State of North Rhine-Westphalia under the title "Digitization of Sustainable Behavior." [Click here to go to the project page on the RWI homepage. Please note that you can only register on the current page].

As a core element of your study participation, you will be provided with a high-quality shower head that measures water consumption per shower. During the four-month study phase, this information will be automatically shared with RWI. Additionally, you will be provided with various information about your showering behavior during the course of the study. At the end of the study, you will be able to keep the shower head, which then ceases to share information. This will allow you to keep track of your water consumption yourself on a permanent basis.

As an expense allowance, in addition to the showerhead, you will receive optional vouchers and/or additional resource saving tools worth a total of €35 until the end of the study. You will receive the vouchers as part of your participation in short, monthly surveys that will be conducted to accompany the study.

Feel free to take a further look at the project page and sign up!

Project goals

"Sustainable showering":
Of course, showering is an indispensable part of personal hygiene. Simultaneously however, showering is a very resource-intensive activity and hardly perceived as such. Many people underestimate the associated energy consumption, caused in particular by heating water. For example, an average shower consumes about 2.6 kilowatt hours. The same amount of energy could be used to light an average household for two and a half days, or to watch television for 26 hours. So the question is whether it is possible to shower more "sustainably."

The goals of the project:

- Gather information on the exact amount of energy and water used in showering.
- Development of helpful measures to reduce resource consumption.

By saving energy consumption while showering, an individual's CO_2 emissions can be significantly reduced. An average shower taken causes annual emissions of 325 kg of CO_2. It takes 26 trees to reabsorb this amount of CO_2 within one year. Or put another way: Just 14 showers emit as much CO_2 as a tree can absorb within a year.

Your contribution

Your contribution to climate protection:
By participating in this study, you are making a concrete contribution to climate protection. This happens in two ways:

- You gain exclusive access to the savings measures investigated in this study, which can be used to reduce your personal resource consumption and the associated energy costs and CO_2 emissions.
- The knowledge generated in the study can be used to formulate concrete recommendations for resource conservation in the context showering and other energy-intensive activities. In this way, policy measures can be developed to effectively save CO_2 emissions.

Every reduction in CO_2 emissions makes a valuable contribution to combating climate change. Further information on the causes and consequences of climate change can be found, for example, on the website of the Federal Agency for Civic Education: https://www.bpb.de/gesellschaft/umwelt/klimawandel/.

Our project is primarily focused on saving energy, i.e., measures are being developed to save hot water. However, water is also a valuable resource in itself and is scarce in many parts of the world, as well as regionally in Germany, at least at times. See, for example, here: https://www.umweltbundesamt.de/presse/pressemitteilungen/wassersparen-sinnvoll-ausgereizt-uebertrieben. Thus, with the study, we will identify scientific findings that are also of great interest to other parts of the world.

FAQ

Do I receive an expense allowance for my participation?

Yes. You will be compensated in two ways for the time and effort you have incurred. First, you get to keep the high-quality shower head used in the study after completion. With this shower head, you can continue to keep track of your water consumption by using a smartphone app. Second, you will receive optional vouchers and/or additional resource saving tools worth €35 until the end of the study. You will receive the vouchers as part of your participation in short, monthly surveys that accompany the study. At the end of the study, we will offer you a selection of popular suppliers where you can redeem your vouchers.

How much effort will it take for me to participate?

At the beginning of the study, we will send you a shower head and a WiFi gateway, which you can easily install yourself with the help of the enclosed instructions. If you encounter any difficulties, we will gladly assist you with the installation. During the course of the study, we will send you monthly online surveys by e-mail, each of which can be completed in about 15 minutes. In addition, you will receive newsletters from us at regular intervals, which you only need to read. The study involves no further effort on your part and is completed after four months.

How is my data collected?

The study is subject to strict data protection regulations. For details, please read the project's privacy policy.

I am going on vacation within the study phase. Is this a problem?

No. However, it would be nice if you let us know when you register. You have the option to do this on the registration form under "Notify the Project Team" or by emailing [E-MAIL ADDRESS].

There is more than one shower in my household. Can I still participate?

Of course. In this case, we ask you to install the new shower head at the shower that you personally use most often.

My shower is not compatible with conventional shower heads.

If you for example have a Rain Dance shower that requires a special shower head, participation is unfortunately not possible. If you are unsure about your shower head model, please feel free to contact us. You can find the contact information here.

A3.2 Signup form content

1. We kindly ask you to provide information about the number of persons living in your household permanently:

Household size

- Total number of persons (adults and children)
- Number of children (age below 18)

2. We kindly ask you to provide information about your shower setup:

Bild 1: Dusche und Badewanne kombiniert

Bild 2: Dusche und Badewanne separat

How many separate or combined showers and bath tubs are there in your home?

- Number of bath tubs with shower (combination, see picture 1)
- Number of separate showers (not in bath tub, as seen in picture 2)
- Number of bath tubs (without shower or not used as shower, as seen in picture 2)

3a. This question is about the shower you use most frequently:

How many household members use this shower at least once a week?

- Number of shower users

Bild 1: Die Duschbrause ist auf einen Duschschlauch aufgeschraubt

Bild 2: Die Duschbrause ragt aus der Wand

Bild 3: Es gibt zwei Duschbrausen. Eine ist auf einen Duschschlauch aufgeschraubt und die andere ragt aus der Wand

Which of these descriptions fits the shower in question?

- No selection made
- Number of separate showers (not in bath tub, as seen in picture 2)
- Number of bath tubs (without shower or not used as shower, as seen in picture 2)
- Number of bath tubs (without shower or not used as shower, as seen in picture 2)

Visually, the shower head we are going to provide you with resembles an ordinary hand shower (see picture below). It can be screwed onto ordinary 1/2 inch shower hoses, which can be found in most households.

3b. We kindly ask for your assessment about the shower you personally use the most

- No selection made
- I think my shower is compatible with this shower head
- I am unsure whether my shower is compatible with this shower head or not
- I do not think my shower is compatible with this shower head

4. Technical requirements

To set up data transmission, you need to connect the shower head to your WiFi network using a smartphone app. This step can be completed in a minute and is necessary for study participation. To perform this step, you will need your WiFi password and a smartphone. The WiFi password will not be shared with us.

- No selection made
- I own a smartphone and have access to my WiFi password
- I own a smartphone and am unsure about my WiFi password
- I do not own a smartphone and/or do not have access to my WiFi password

In addition to the smart shower head, we will provide you with a WiFi gateway, which is required to connect the shower head to your network. This device is very small (dimensions: 4.5 x 4.5 x 1cm, see picture below). It should be located within a radius of about 5 meters around your shower (ideally in the same room) and should be continuously powered.

- I have access to an electrical outlet within a radius of 5 meters around my shower

A4 Exemplary newsletter for the BOTH group

Dear Mr/Ms XYZ,

do you remember how much energy is consumed in a five-minute shower at a water temperature of 38 degrees Celsius? That's right, it's 2.2 kWh!
With these 2.2 kWh you could also

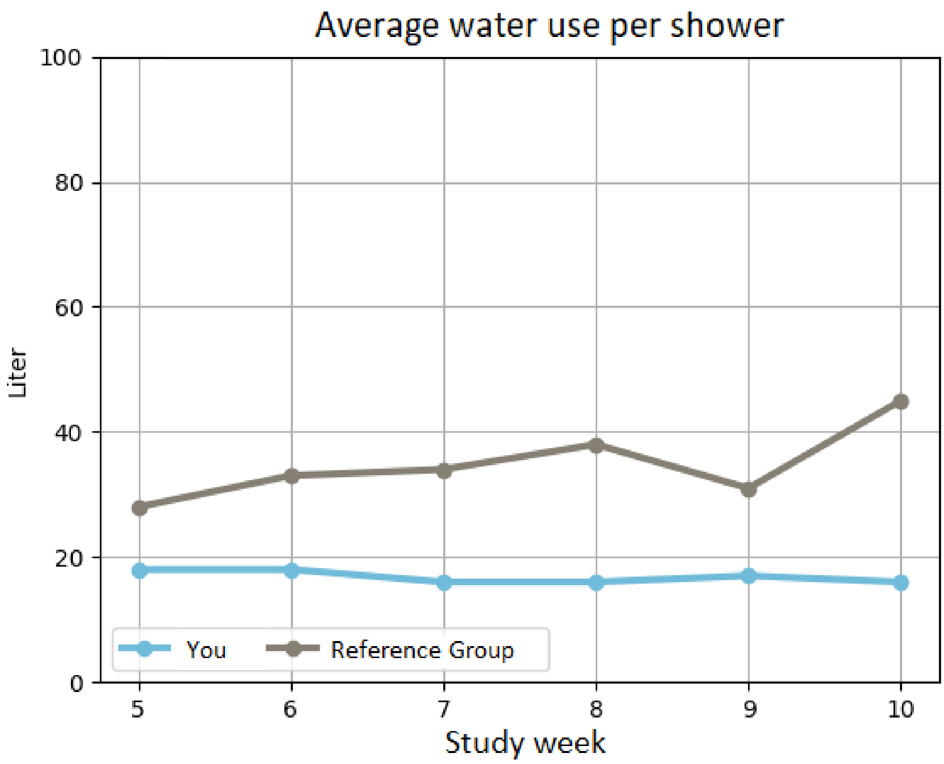
- toast 132 slices of bread
- boil 16.5 litres of water at 100 degrees Celsius with an electric kettle

The generation of 2.2 kWh of energy emits 650 g of CO2. It takes 19 beech trees to reabsorb this amount of CO2 within one day.

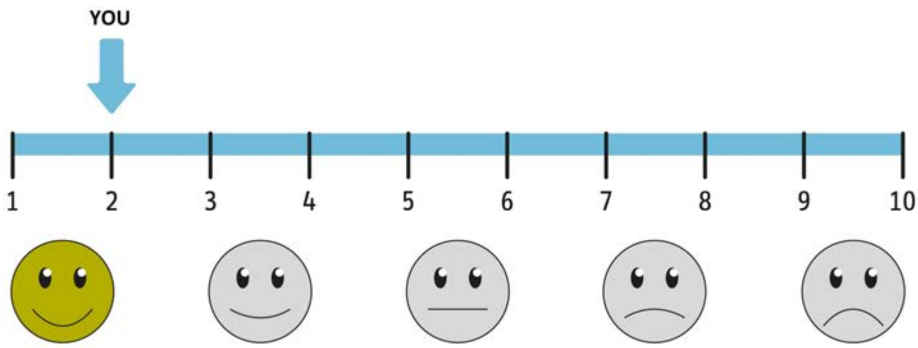
In the last two weeks, you have used an average of **17 litres of water** per shower. At your average water temperature and the type of water heating used by an average German citizen, this corresponds to emissions of **244.5 g of CO2 per shower**. It takes 7.1 beeches to reabsorb this amount of CO2 within one day.

The average consumption of 9 other randomly selected participating households was **36 litres**. The average CO2 emission per shower was **537.0 g CO2 per shower**. It takes 15.7 beeches to reabsorb this amount of CO2 within one day.

In the following figure you can see how your average consumption and the average consumption of your comparison group developed in the last weeks. The last two weeks were weeks 9 and 10 of the study.



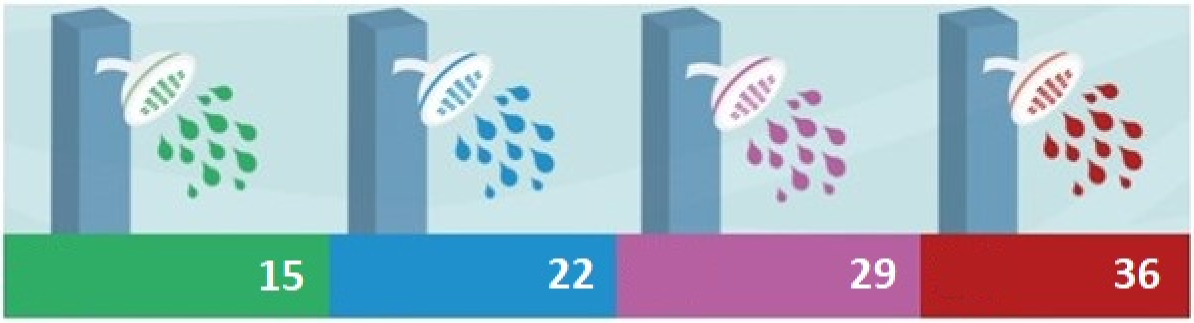
Compared to 9 other participating households, your rank is 2nd place:



One of these households had lower average water use per shower than you. Eight of these households had a higher average water consumption.

Ihr smarter Duschkopf

Your shower head will help you keep track of your water consumption while showering. It will light up in color and as your water consumption increases, the color will change. Finally, once your water consumption exceeds 36 litres, it will start flashing red. 36 litres of water consumption corresponds to 537 g of CO2 emissions based on your average water temperature and the energy consumption of an average German citizen. It takes 15.7 beech trees to reabsorb this amount of CO2 within one day. The exact sequence of colours is as follows:



During the first 15 litres of your shower, the shower head will glow green.
After **15 litres** (225 g CO_2 emissions), your shower head will begin to glow blue.
After **22 litres** (330 g CO_2 emissions), your shower head will begin to glow purple.
After **29 litres** (435 g CO_2 emissions), your shower head will start to glow red.
Finally, after **36 litres** (537 g CO_2 emissions), your shower head will begin to flash red.

Sincerely,
Your "nachhaltiges Duschen" team

Please do not reply to this mail. For questions and comments, please contact us at [E-MAIL ADDRESS].

A5 Technical details and installation of the smart shower heads

Participants received a package with the smart shower head and the WiFi gateways. In parallel, participants received the installation instructions in digital form via email. Upon request, a printed version of the instructions was enclosed in the package.
The installation of the smart shower heads and the infrastructure for data transmission took place in the following steps.

1. The existing shower head had to be unscrewed from the shower hose and the smart shower head had to be screwed on.
2. The proper functioning of the shower head could be confirmed by brief light signals, which the shower head emits during start-up. Apart from this short signal, the shower head to the participants functioned like any regular shower head during the Baseline Phase.
3. The WiFi gateway had to be plugged into a power outlet near the shower head. Participants then had to download an app to connect the WiFi gateway to their home network. This was done selecting the correct WiFi network and entering the network key. A successful connection of the gateway was indicated by light signals.
4. If problems arose, participants could reach the project team by e-mail. The problems were then solved either by e-mail, phone call or personal visit. Only for a minority of the study participants were the problems not solvable. In most cases, the issues resulted from WiFi configuration errors or too much distance between the gateway and the smart shower head, which could be easily solved.

After successful installation, the shower head transmitted information for every shower taken (time stamp, amount of water used, average water temperature, water flow, length of shower breaks) via Bluetooth to the WiFi gateway, which then transmitted the information to the research team via the Internet. Participants had no way to access this information during the study.

If the WiFi gateway was not plugged in or had no internet connection during a shower, the data from that shower was stored in the shower head (up to 200 showers can be stored) and transmitted with the next successful connection. Interruptions of showers of less than 3 minutes were interpreted by the smart shower head as shower breaks, while an interruption of more than 3 minutes signaled the start of a new shower.

A6 Details on the calculation of shower costs and carbon intensity

We draw on calculations by the Consumer Organisation of the German state of Rhineland-Palatinate (VZ-RLP, 2020) on the average cost and energy consumption per shower at 38°C, the average baseline temperature in our control group (Table 2). We convert these values to the cost and energy demand per liter of shower water, distinguishing between water heating by electricity, gas or oil. We multiply the energy demand per liter of shower water with the estimates provided by the German Federal Environment Agency for the CO_2 intensity of one kilowatt-hour of heat supply with the respective energy sources (UBA, 2019). We then use data from the German Federal Ministry for Economic Affairs and Energy (BMWi, 2020) on the distribution of final energy consumption for water heating in private households among these three energy sources considered to calculate a representative CO_2-factor per liter of shower water. In this calculation, we make the assumption that shower water in Germany is heated only by electricity, gas or oil. This is a simplification in the absence of more detailed data, but it includes a large 84% of the final energy consumption for water heating in German households in 2018. The remaining households obtained district heating or heated water with renewable energy. Our estimate therefore tends to be slightly too high. As a representative CO_2-factor, this calculation yields a value of 0.015 kg CO_2 per liter of shower water.

Table A5: Estimation of CO_2 per liter

	(I) Energy cost per shower (40 L at 38°C)	(II) Cost per kWh	(III) kWh per L at 38°C	(IV) CO_2 per kWh	(V) CO_2 per L	(VI) PJ used for water heating in 2018	(VII) Share	(VIII) CO_2 per L (weighted avg.)
Elec.	0.40 EUR	0.28 EUR	0.036	0.551	0.020	55.307	17.9%	
Gas	0.10 EUR	0.05 EUR	0.050	0.246	0.012	190.384	61.6%	0.15
Oil	0.15 EUR	0.07 EUR	0.054	0.318	0.017	63.307	20.5%	
	VZ-RLP (2020)	VZ-RLP (2020)	$\frac{(I)/40}{(II)}$	UBA (2019)	$(III) \times (IV)$	BMWi (2020)	$\frac{(IV)}{\sum(VI)}$	$\sum[(V) \times (VII)]$