Global Environmental Psychology

PsychOpen GOLD

Research Articles

Check for updates

Optimal Delivery of Social Norms Feedback to Reduce Household Water Consumption

Ukasha Ramli [1], Cameron Brick [2], Benjamin Abera [3]

[1] *Department of Psychology and Behavioural Sciences, London School of Economics and Political Science, London, United Kingdom.* [2] *Department of Psychology, University of Amsterdam, Amsterdam, the Netherlands.* [3] *Department of Psychology, University of Inland Norway, Lillehammer, Norway.*

Global Environmental Psychology, 2025, Vol. 3, Article e11705, https://doi.org/10.5964/gep.11705

Received: 2023-04-05 • **Accepted:** 2024-08-16 • **Published (VoR):** 2025-12-12

Handling Editor: Karlijn van den Broek, University of Utrecht, Utrecht, the Netherlands.

Corresponding Author: Ukasha Ramli, Department of Psychology and Behavioural Sciences, London School of Economics and Political Science, London, United Kingdom. E-mail: kashramli@gmail.com

Badges for Good Research Practices: Code. Diversity Statement.

Abstract

Providing households with consumption feedback can effectively reduce water and energy use. However, the efficacy may depend on the medium and frequency of the messages. In this longitudinal study spanning over two years, we report two experiments on household water use read via meter in the United Kingdom: one manipulating the frequency of feedback (N = 13,047), and one comparing the medium (email vs. paper) of the feedback messages (N = 18,896). Overall, the feedback interventions reduced water consumption by around 2%. There were no differences in delivering the treatment monthly vs. 3-monthly, and paper messages were effective, but emails were not. Almost all households reduced their consumption, although the effect was variable. These results inform psychological interventions and cost-effective messaging for water districts and perhaps other types of utilities.

Keywords

social norms, pro-environmental behaviour, eco-feedback, water use, energy use

Non-Technical Summary

Background

Water scarcity is becoming an increasingly serious global challenge due to climate change and rising demand. One promising approach to reducing household water use is sending personalized feedback about consumption, often including comparisons to neighbours. This approach, known as "social norms" messaging, has been shown to effectively reduce both energy and water use in various contexts. However, there are many unanswered questions about the best way to deliver these messages, including how often they should be sent and whether digital or physical formats are more effective.

Why was this study done?

The goal of this study was to understand how the frequency and medium of feedback messages affect household water consumption. Although feedback interventions have been tested many times, utilities face practical decisions about implementation that research has not fully addressed. Sending paper mail is expensive, so knowing whether less frequent mailings work just as well could save significant costs. Similarly, email is far cheaper than paper mail, but it remains unclear whether households pay attention to digital messages. This study aimed to provide evidence to help water utilities design cost-effective programs that can still change behaviour.

What did the researchers do and find?

The researchers conducted two large-scale field experiments with over 30,000 households in the United Kingdom over more than two years between 2016-2019. In the first experiment, households received personalized water use reports comparing their consumption to their neighbours either monthly, every three months, or every six months. In the second experiment, households received similar reports either by paper mail or by email.

The results showed that overall, the feedback interventions reduced household water consumption by approximately 2%. Monthly and three-monthly paper reports were equally effective, meaning there was no benefit to more frequent messaging. Paper mail successfully reduced water use, but email had no effect whatsoever. The researchers also found that households with medium-to-high baseline consumption showed the largest reductions, while those with very low or very high consumption changed less. The intervention proved cost-effective, with utilities saving approximately 798 litres of water for every £1 spent on paper messaging.

What do these findings mean?

These findings suggest that for water utilities considering feedback programs, sending paper reports every three months appears to be the most cost-effective approach. Increasing the frequency to monthly mailings does not improve outcomes but does increase costs. Email, despite being much cheaper, should not be expected to change behaviour, likely because people ignore or delete these messages without reading them. The findings also suggest

PsychOpen GOLD

that targeting households with medium-to-high water consumption could maximize water savings per pound spent. However, it should be noted that these findings come from specific regions in England, and results might differ in other contexts with different climates, water use patterns, or cultural norms around mail and digital communication. Future research might test whether these results can be replicated in other countries and whether alternative digital formats could be made more engaging.

Highlights

- Social norms feedback interventions reduced household water consumption by approximately 2% in two large-scale field experiments in the United Kingdom, consistent with prior research in this context.
- There were no significant differences in effectiveness between monthly and three-monthly message delivery, suggesting that utilities can reduce messaging frequency without compromising water savings.
- Paper mail reports successfully reduced water consumption, but email-based interventions had no detectable effect, likely because digital messages are easier to ignore or require additional steps to access the feedback content.
- Heterogeneity analyses revealed that households with medium-to-high baseline consumption showed the greatest reductions, while those with very low consumption faced floor effects and the highest consumers showed attenuated treatment effects.
- The intervention proved cost-effective, saving approximately 798 litres of water for every £1 spent on paper messaging, with potential for greater efficiency when targeting specific household segments based on baseline consumption.

Water scarcity is a looming threat that could affect everyone on the planet (Caretta et al., 2023). Increasing demand combined with the effects of climate change means that conserving this limited resource is extremely important (Global Commission on the Economics of Water, 2023; Kummu et al., 2016; UN Water, 2021). Clean water is the foundation of our livelihoods, ecosystems, and economies, and preserving it requires demand-side changes in consumers. One of the most popular and consistently effective interventions on behaviour is providing feedback to consumers about their resource consumption (Andor & Fels, 2018; Karlin et al., 2015; Sanguinetti et al., 2018).

What distinguishes eco-feedback from general feedback is the goal of reducing environmental impact paired with personalisation. Eco-feedback is the delivery of personalised information during user-product interaction to encourage the adoption of energy or water-saving strategies (Sanguinetti et al., 2018). In contrast, general messaging conveys broad information to a broad audience without personalising it to specific consumption patterns, e.g., “You should reduce your consumption to help the planet”. These feedback interventions are often combined with social information about the consumption patterns of similar peers (e.g., nearby households of similar size), known as descriptive social norms. An early study found that door hangers with personalised feedback about a household's energy consumption combined with descriptive and injunctive social norms reduced household energy use by about 8% or $d = .6$ (Schultz et al., 2007).

Expanding upon these findings, the software company OPOWER was a major company in translating findings in eco-feedback research into a software product for utilities (Allcott, 2011; Bloomberg, 2023). They sent out Home Energy Report Letters containing energy-saving tips and included a comparison between a household's electricity usage and that of their neighbours. Utility companies across the world then adopted similar norm-based feedback interventions. Interventions in the United States have resulted in energy consumption reductions ranging from 2% to 2.4% (Allcott, 2011; Ayres et al., 2013; Costa & Kahn, 2013; Schultz et al., 2007). In Europe, decreases were seen at 0.7% (Andor et al., 2020), 0.4% (Bonan et al., 2020), and 6% (Dolan & Metcalfe, 2015). Interventions have reduced household water use by 1.8% to 6% (Bhanot, 2017; Brent et al., 2015; Carlsson et al., 2021; Ferraro & Miranda, 2013; Ferraro & Price, 2013; Ferraro et al., 2011; Ramli, 2021; Ramli & Laffan, 2022). Notably, Schultz et al. (2016) showed a substantial 26% reduction in water use in California with a descriptive norm message. This growing evidence base demonstrates the effectiveness and low cost of these interventions as a tool to manage consumer demand.

However, the proliferation of these interventions has highlighted the large heterogeneity in treatment effects, both within and between studies (Allcott & Mullainathan, 2010; Andor et al., 2020). Across multiple meta-analyses of feedback interventions, these differences in implementation appear to moderate overall treatment effects (Karlin et al., 2015; Kluger & DeNisi, 1996; Sønderlund et al., 2014). Possible moderators include the type of feedback medium (smart meters or in-home displays), the timing of feedback delivery (immediate or delayed), and tailoring feedback to household characteristics. These implementation differences show the importance of thoughtful design and deployment of feedback interventions. Such nuances can either amplify or undermine the desired treatment outcomes. The current study addresses this possibility by testing key design features in how interventions were delivered in two experiments on household water consumption. The first experiment tested different frequencies of message delivery, and the second experiment tested the delivery medium (email vs. paper).

Expanding on this theoretical groundwork, in the next section, we examine the determinants of household water usage, which are shaped by individual decision-making processes and demographics. We investigate the key features of feedback interventions using insights from energy conservation research. We also assess the medium and frequency of feedback, compare email and paper as intervention mediums, and consider the implications for utility messaging practices.

Intervention Delivery Mechanisms

Households are made up of individuals that choose when and where to use water (Lancaster, 1975). These behaviours are related to psychological factors like knowledge and motivation. However, they are also determined by demographics like income and location, which predict household size, types, and amount of appliances or features such as fountains (Bich-Ngoc & Teller, 2018; Domene & Saurí, 2006). Several meta-analyses and reviews found that average messaging effects vary widely. First, sample population (e.g., households that volunteer to participate in a study or general culture) and physical location (e.g., climate) explain some variance (Allcott, 2015; Andor et al., 2020). For example, the sample population can have an effect through site-selection bias whereby communities that are already environmentally conscious may invite pro-environmental interventions to be tested in their region (Allcott, 2015). Alternatively, the variables associated with a physical location can determine how much change can occur simply by how much consumption already occurs (e.g., temperate climates that do not require energy for cooling or heating generally consume less energy and, therefore, have less scope for change) (Andor et al., 2020). Others attribute treatment variance to intervention characteristics. For example, Kluger and DeNisi (1996) developed the Feedback Intervention Theory by conducting a meta-analysis of 607 treatment effects of feedback interventions on a wide range of behaviours, with a primary focus on performance-related behaviours across various domains. This meta-analysis was the first attempt at understanding how differences in feedback implementation impact the treatment effects, and two-thirds of the interventions appeared effective at changing behaviour.

Which reference point to use for feedback has been extensively studied, and using social norms has become standard practice. For example, the California Public Utilities Commission defined behaviour-based energy efficiency programmes as requiring social norms-based comparisons (Mahone & Haley, 2011). This choice is likely due to social norms' reliable efficacy (Karlin et al., 2015; Nolan et al., 2008). Even household water use studies that were not focused on testing social norms often contain a norms condition. For example, one study testing Information Motivation Behavioural Skills theory included an experimental condition focused on what other households were doing (Hodges et al., 2020).

In energy consumption, reviews by Fischer (2008), Karlin et al. (2015), and Sanguinetti et al. (2018) highlighted several components of feedback interventions that moderate its effects:

1. Intervention delivery frequency.
2. Medium or format of the information (e.g., letters, display units, emails).
3. Metric of use (e.g., litres, price, CO2).
4. What the feedback is compared against (e.g., historical behaviour, social norms).
5. Feedback granularity (e.g., household, room, appliance).

PsychOpen GOLD

6. Length of treatment.

From a policy perspective, it is important to focus on frequency and medium because these components have significant operational costs that inform intervention scaling. Sending paper requires printing, sorting, and mailing, and frequency multiplies this cost. Therefore, the medium and frequency of messaging interventions have large practical relevance for high impact through behaviour change.

Medium: Email vs. Paper

The first published example of using social norms for reducing energy consumption was delivered through door hangers (Schultz et al., 2007). Since then, other mediums have been explored, including 'home reports' by paper mail (e.g., Allcott, 2011), email (e.g., Dolan & Metcalfe, 2015), real-time consumption feedback on an in-home display (e.g., Schultz et al., 2015), or by a website or mobile application (e.g., Geelen et al., 2019). Email and paper mail are 'pushed' onto people, while in-home displays and websites require people to 'pull' the information (Schultz et al., 2016). 'Pull' mediums rely on users to dictate the frequency at which they are treated by the intervention and typically require smart meters, which are not as widely available for water use. Because 'pull' methods are less scalable and provide less control of treatment dosage, we focused on paper mail vs. email. Utilities are more likely to engage their customers with paper mail because they have customer home addresses but may not have up-to-date email addresses or have installed smart meters and in-home displays.

Regardless of format, messages must engage the audience. If feedback is not received or read, there can be no treatment effect. Mediums that are challenging to access or are too easily ignored would probably work less well. Dolan and Metcalfe (2015) directly tested paper vs. email as mediums of intervention delivery. Only paper mail reduced energy consumption. Other reviews found that computer-based mediums were more effective than paper and suggested it was because such messages are more interactive and engaging (Fischer, 2008; Karlin et al., 2015; Sanguinetti et al., 2018). However, this assumes people are willing to log in to access the feedback. One poll found that the average adult American has 500 unread emails (Moore, 2015). This highlights a challenge of the inundation of digital communications: messages might remain unseen or unopened in a cluttered digital space. Communications can only be effective if they are noticed.

Feedback is effective at changing behaviour, especially in areas such as energy or water consumption where people may be unaware of their usage. However, a person's motivation to monitor their goal progress (i.e., seek feedback), even those they rate as something important, may wane due to specific psychological mechanisms (Webb et al., 2013). For example, the self-verification motive proposes that individuals seek consistency with their established self-concept. In this context, if individuals believe they are not performing well, the feedback could challenge their beliefs or self-concept, and as a result, they may be motivated to dismiss it. With emails, the subject line and sender may signal the content without opening a message, and so unwelcome news may be easy to ignore. This is especially the case if the feedback is not included in the body of the email and the customer needs to log into a website to view usage. In contrast, the content of physical letters is difficult to discern without opening. Therefore, physical letters may be more likely to be read, even if briefly. This difference in engagement may be crucial for behaviour change.

Frequency

Frequency refers to how often feedback is delivered to the individual. More frequent feedback improves learning and task performance (Ilgen et al., 1979; Lam et al., 2011; Salmoni et al., 1984). Increased frequency provides more information for the individual to learn from and the opportunity to develop strategies to improve performance. Greater learning comes because individuals can better link an action with specific feedback in time. More frequent feedback could also create greater saliency of the target behaviour. Furthermore, when the effect of a feedback intervention decays over time, more frequent feedback can prevent that decay (Allcott, 2011).

It is less clear whether more frequent feedback leads to more behaviour change. While reviews by Fischer (2008) and Darby (2006) suggested so, a meta-analysis by Karlin et al. (2015) found no effect. Perhaps this null finding was because frequency was not studied in isolation and so was masked by the moderating effect of other variables. Or there may be no true effect. In one experiment on household energy consumption, households that received paper home reports each month vs. each quarter both used a similar amount of energy (Allcott, 2011). Looking at the wider pool of 17 experiments analysed by Allcott (2011) that delivered interventions at various frequencies, the unweighted mean of the average treatment effects was -2.2% for monthly and bimonthly interventions compared to -1.7% for quarterly interventions. If higher frequency were only slightly more effective, the higher frequency would not justify the increased costs of printing and mailing. It is also unknown how infrequently interventions can be delivered without compromising the treatment effect. This question is especially relevant for water utilities in the United Kingdom because the majority of U.K. households have meters that are only read every six months (UK Government, 2023), which makes frequent, personalised feedback more challenging: for example, the cost of delivering monthly feedback interventions may include installing and maintaining smart meters. In sum, the influence of feedback frequency on behaviour change remains inconclusive. It is also important to consider how feedback frequencies may affect the duration of treatment effects after treatment ends (Allcott & Rogers, 2014). Further research is needed to determine whether higher frequencies lead to longer-lasting effects, especially where household meters are read only every six months.

Aims

The current paper presents two field experiments with different social norms feedback interventions. The primary aim of Study 1 is to examine the effects of messaging frequency on household energy use by conducting a field experiment that tests monthly, quarterly, or 6-monthly messaging. Because Allcott (2011) did not find differences between monthly and quarterly messages about energy use, we expected no difference between the monthly and quarterly conditions and, by extension, between all conditions and the 6-monthly conditions. This replication would provide evidence for potentially large savings for utilities. This study is the first to test feedback frequencies in the United Kingdom and the first frequency experiment on water consumption globally.

Study 2 aims to compare the impact of email and paper mail as delivery mediums for feedback interventions on household water consumption. Because Dolan and Metcalfe (2015) found that letters were more effective than emails, we expect the same result here. We also explore the heterogeneity of the treatment effect. Lastly, we aim to assess the cost-effectiveness of the feedback interventions by evaluating the potential savings for utilities, the resulting impacts on household bills, and the efficiency of targeting specific household segments based on their water consumption.

Method

Participants

Two randomised field experiments were conducted during 2016–2017 in two regions of the United Kingdom through different regional water utility companies, Anglian Water and South East Water. Water utilities in the United Kingdom are private companies that are run as regional monopolies and, as such, are heavily regulated by the U.K. Water Services Regulation Authority. The regulators have incentives to encourage utility companies to reduce overall water demand, especially in drought-prone areas like where the two experiments were conducted.

The inclusion criteria required households to have a working water meter, sufficient meter reads (1/3/6 months) prior to the study, meter reads within an upper bound of 10,000 litres a day, not be on a special tariff for households with low income or specific medical conditions, not have opted out of all forms of communications from the utility, be a residential household account, and have an account that had not been closed. Houses were billed on a 6-monthly basis, with an average annual expenditure of £415. Mean residential water consumption in the United Kingdom was 349 litres per household per day in 2013 (Energy Saving Trust, 2013), compared to 523 litres in the United States in 2016 (DeOreo et al., 2016). We did not have additional demographic data on the households as these are not commonly collected by the utility company and purchasing that data for a sample of this size would have been too expensive. All households were automatically opted into the experiment, so those in the treatment group received the intervention. They could choose to opt-out at any point. The automatic inclusion to receive these communications was viewed to be an added service by the utility company as a means to help households be more environmentally friendly and potentially achieve financial savings. Before each treatment interval, a home report was only generated for households that were not excluded by these criteria. Even if a household that was previously included in the programme was later excluded from receiving a home report during the intervention period, their data was still collected as long as they had not closed their account nor had a meter read error. See the section, ‘Study 1: Frequency’ and ‘Study 2: Medium’ for the exclusion numbers per study.

We applied three further exclusion criteria before we looked at the data. First, households that did not have any data over the baseline period were removed (see each Study below for specific dates). Second, households that did not have any data over the treatment period were removed. Finally, outliers were removed above the 99th percentile for water use because their use was unrealistically high for a single household. Mean U.K. household consumption in this period was 349 litres per day. Before this exclusion, the highest consumption was 25,000 litres/day in Study 1 and 80,000 litres per day in Study 2. Readings can vary between studies because of random meter read errors or errors in the data about types of properties (e.g., commercial). The 99th percentile was used because it balanced the removal of unrealistic use without removing many data points. After this exclusion, the highest consumption was 1,073 litres per day in Study 1 and 809 litres per day in Study 2, which are much more plausible for single households. As requested by a reviewer, we also conducted the main analyses using a threshold of +3 *SD*. The results did not vary by more than 1 litre per day, and so we used the 99th percentile as planned.

Experimental Design

Households randomly allocated to the treatment condition received a 'home report' and access to an online portal with the same content. The reports and online portal were developed by Advizzo, a software-as-a-service company that works with utilities. The authors worked with Advizzo and the utility companies to run the randomisation. The ‘home reports’ were delivered by the utility companies. The main feature of the home report was the social norms messaging in a graph that showed descriptive norms (Allcott, 2011; Ayres et al., 2013; Schultz et al., 2007; see Figure 1). This neighbour comparison graph displayed a comparison of the water consumption of the target household against the consumption of ‘average homes’ and ‘efficient homes’ (Figure 1). The message recipient would see their household placed in a condition based on use: ‘more than average’ (consuming more than the mean of similar households), ‘below the average’ (less than the mean of similar households), and ‘'most efficient’ (less than 80% of households). In this bar graph, the bars represent the amount of water consumed. Consumption was displayed as cubic metres (m^3) to align with the bill payment metric. An injunctive norm was also displayed in the form of a series of ‘smiley faces’ with three labelled levels, ‘More than average’, ‘Good’, and ‘Great’. The injunctive norm helps prevent boomerang and coasting effects (Schultz et al., 2007). The home report also included three tips to reduce water consumption (e.g., “Each minute you reduce your shower time can save you at least 9 litres”). Households received the home reports every month, three months, or six months, and through paper mail or email, depending on the condition. The control conditions in either experiment did not receive any communications from the study team. The emails did not have a read receipt; therefore, we could not determine whether they were opened. Equally, there was no way of knowing if households opened the paper mail.

Study 1: Frequency

This May 2017 study tested feedback delivery frequency (randomised into control, monthly, 3-monthly, and 6-monthly) on customers of Anglian Water, a water utility in the East of England. An initial sample of 22,000 households from the Colchester region was selected, of which 13,047 met the selection criteria. Household water meters were read at different times of the month using refuse trucks with remote meter readers as they drove around collecting refuse. The remote reading is accomplished with a radio frequency device mounted to the water meter, which is triggered by a wireless request from the nearby reader. After the metering data had been collected, a meter data management company sent us the data that showed monthly data consumption for each household. Baseline consumption data from July 2016 to June 2017 were used. There was no data available for June 2016, so for each household, we imputed their consumption data for June 2016 using data from July 2017 to ensure a full year’s data for analysis. Following the delivery of the first home report, additional reports were subsequently sent out every month, three months, or six months, depending on experimental conditions. After one year of the intervention, only households in the monthly treatment condition that had shared their email address with Anglian Water continued to receive the treatment. Households that had not shared their email address and households in the 6-monthly and 3-monthly conditions stopped receiving the treatment. Consumption data for these households continued to be collected after the interventions stopped. Figure A1 (in the Appendix) shows the number of home reports sent across the entire programme period.

Figure 1

Example of Household Home Report

Out of the 13,047 households, 425 were removed for not having data available in the baseline period, and 26 households were removed for not having any data in the treatment period. Additionally, 4,901 individual data points with water consumption higher than the 99th percentile were removed, as explained in the section above. The final data-set consisted of 12,586 households. An attrition analysis was performed to determine if households dropped out of the study vs. completed the entire treatment and whether this attrition was associated with the treatment assignment. Two regression analyses on attrition using the number of available months, as well as a binary indicator of attrition on treatment assignment, were conducted. There were no statistically significant differences in attrition between conditions (Table A1 in the Appendix), which is evidence against the presence of bias through group assignment. Furthermore, baseline consumption between the treatment and control conditions was not different, which is evidence that baseline consumption was balanced at randomisation. Descriptive statistics are in Table 1.

Table 1

Water Use by Study and Condition (Litres per Household per Day)

	Litres					
	Pre-treatment		Post-treatment		Households (*n*)	
Group	***M***	***SD***	***M***	***SD***	**Year 1**	**All years**
Study 1: Frequency						
Monthly	297	223	277	176	3,333	3,311
3-monthly	297	263	278	176	3,255	3,237
6-monthly	297	208	282	173	2,847	2,845
Control	290	216	281	176	3,141	3,122
Study 2: Medium						
Email	290	156	296	147	4,565	4,563
Paper	289	182	289	147	4,618	4,612
Control	288	150	293	147	9,120	9,108

Note. M = mean, *SD* = standard deviation, *n* = number of households in the first year (Year 1) and all years (Study 1: 35 months; Study 2: 31 months).

Study 2: Medium

Study 2 tested different mediums of feedback delivery on customers from South East Water, a water utility operating in the South East region of England. Consumption data were manually collected from households on a 6-monthly basis by the utility company as standard practice. After the metering data had been collected, a meter data management company sent us the data that showed 6-monthly data consumption for each household. Twelve months of consumption data between March 2016 and March 2017 was used as the baseline period. Home reports were sent about every eight months starting from March 2017, with a final report sent in March 2019. Consumption data through September 2019 was analysed.

An initial sample of 20,000 households that shared their email addresses with the utility were selected. Following the exclusion of households based on the criteria above, 18,896 households were randomly assigned to three conditions: receiving the home report by email, receiving the home report by paper, or a control condition that received no communication from the study team. All households at this time did not receive any other communication campaigns aside from their usual service communications (e.g., late bill payments). Half the sample was assigned to the control condition and the other half was split between the two treatment conditions. Five hundred forty-five households that had no consumption data in the baseline period were removed, and a further 37 households that had no consumption data in the treatment period were removed. The final data set, therefore, consisted of 18,314 households. There were no significant differences in attrition between the three conditions (Table A1 in the Appendix), and consumption in the three conditions was balanced over the baseline period.

Measures

Household Water Use

The dependent variable was water consumption in litres per day. The difference between two meter reads of cumulative consumption allowed us to calculate the consumption over that period (monthly for Study 1 and 6-monthly for Study 2) by dividing the meter reading by months elapsed and then dividing by days within each month.

Statistical Analysis

For a power analysis, a post-hoc sensitivity analysis was run after the study with α = .05, power = .99, and Study 1 N = 12,586 for a linear regression with three predictors using the function pwr.f2.test in the *pwr* R package. This revealed excellent power to detect changes in water use of R^2 = 0.002 or larger. Study 2 also allowed detecting effects of R^2 = .002 with excellent power.

For both studies, ANCOVA was used to estimate average treatment effects. To determine whether the effects of the different treatment conditions were equivalent, a post-hoc two-one-sided *t*-test (TOST) was used as an equivalence test.

To understand how the intervention affects people differently, heterogeneity of treatment effects was studied using quantile regressions and conditional average treatment effect (CATE). Knowing this can help improve the targeting of the intervention to maximise its impact. Quantile regressions are useful for studying heterogeneity because rather than just estimating the treatment effect on the mean of the outcome of interest, quantile regressions allow the estimation of the effect on the median of this outcome alongside the full range quantiles, revealing how the treatment effect might vary across the different aspects of the distribution of the outcome of interest. This technique is common in economics and political science, and relatively rare in psychology, despite its utility. We present a separate unconditional quantile regression, LPDit = Ti, for each treatment condition. One limitation of quantile regressions is that due to rank invariance, it is difficult to identify specific conditions within the output.

Heterogeneity can also be studied by looking at the conditional average treatment effect (CATE) of baseline consumption. This is complementary to quantile regressions, which are unable to make statements about conditions of the sample due to rank invariance. CATEs are more interpretable and enable statements about how different levels of baseline consumption are affected by the treatment. To run CATE, we used the same model specification as for the average treatment effect, with the addition of an interaction term between treatment condition and deciles of baseline consumption. The limitation of using interaction terms for subgroup analysis is lower statistical power.

Lastly, the cost-effectiveness of the interventions was assessed by calculating the amount of water saved per £1 spent per household and comparing this to the cost of water production.

Results

Anonymised data can be shared upon request. All analysis code (Python) is at Ramli et al. (2025).

Average Treatment Effects

The average treatment effects for both experiments were estimated with an ANCOVA with specifications similar to those of Allcott and Rogers (2014),

$$LPD_{it} = T_i + LPDpre_{it} + \pi_m + \epsilon$$

where LPD_{it} is household i's water consumption in litres per day at month t, T_i is the treatment indicator, $LPDpre_{it}$ is household water consumption in the matching calendar month in the baseline period, and π_m are month and year fixed effects. All analyses included the clustering of standard errors on the household level.

Study 1: Frequency

To address the first aim, we investigated the effects of messaging frequency on household water consumption. Specifically, we examined whether there were differences in water usage when households were receiving feedback at monthly, quarterly, and 6-monthly intervals. The treatment was delivered to the majority of the sample for the first 12 months only. Therefore, the analysis was split to focus first on this initial period, then the subsequent 23 months, and then all the months together.

Only households in the monthly and 3-monthly treatment conditions reduced their water consumption over the first 12 months when the reports were being sent (Table 2;

Figure 2). Households in both these treatment conditions reduced their consumption by 6 litres per day, or -2%, compared to the control condition. The subsequent 23 months showed more reductions, with households in the monthly condition reducing their consumption by 8 litres per day and the 3-monthly condition by 7 litres per day. A pairwise *t*-test for equality of coefficients (with Holm-Sidak correction for multiple tests) showed no differences between these three treatment conditions.

The post-hoc TOST equivalence test was conducted to determine if the differences in water consumption reductions between the 1-monthly and 3-monthly intervention groups were within the predefined equivalence margins of ±0.5 litres per day. The results suggest that equivalence for both the upper bound $t(75442) = 0.514$, $p = .303$ and lower bound $t(75442) = -0.514$, $p = .303$ could not be established. Consequently, the data did not provide sufficient evidence to assert that the differences in effectiveness between the two frequencies fall within the narrow range considered practically equivalent. The equivalence margin of ±0.5 litres per day was selected based on the observed effect sizes and the cost implications of the interventions, with the rationale that differences smaller than this threshold would not meaningfully alter the cost-effectiveness of the interventions. Additionally, the residuals from the ANCOVA of the main effect were used as the basis for the equivalence test to control for the influence of baseline water consumption and other covariates on the outcome variable of average daily water usage. This approach allowed for the isolation of the effect of the intervention frequency on water consumption independent of these covariates.

Table 2

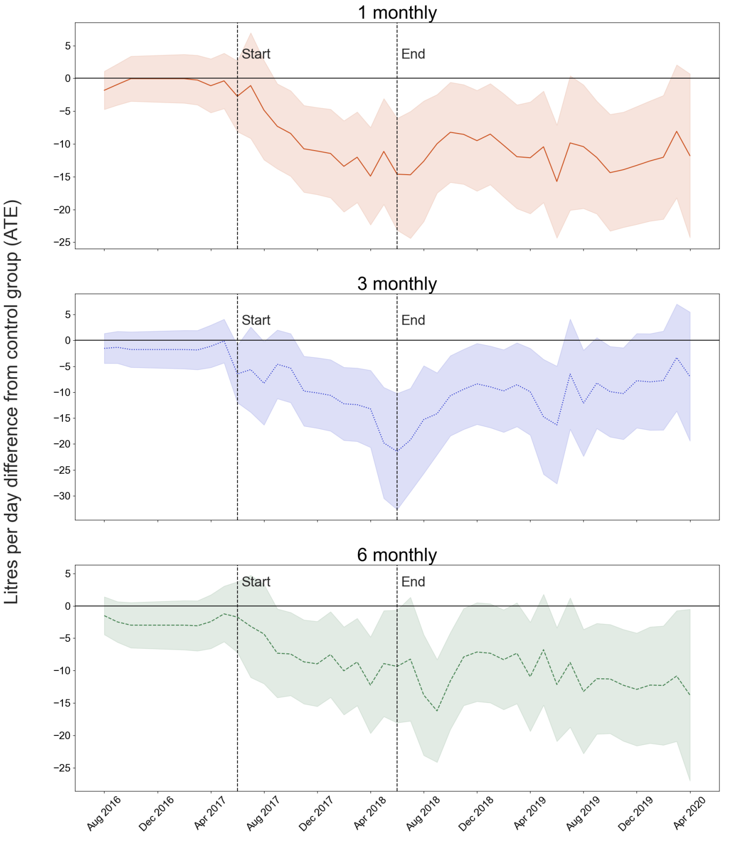
Linear Regression of Message Frequency on Water Use (Study 1)

	Household water use (Litres/day)	
Group	**Treatment period (12 months)**	**All months (35)**
1 monthly	-6.33** [-10.41, -2.25]	-7.42** [-11.78, -3.05]
3-monthly	-6.61** [-10.80, -2.43]	-7.02** [-11.45, -2.60]
6-monthly	-1.71 [-5.83, 2.40]	-3.24** [-7.60, 1.13]
3-monthly vs 1 monthly	-2.85 [-4.51, 3.94]	0.39 [-4.09, 4.87]
6-monthly vs 1 monthly	4.62 [0.59, 8.64]	4.18 [-0.11, 8.47]
6-monthly vs 3-monthly	-4.90 [0.65, 9.16]	-3.79 [-0.68, 8.26]
Number of households	12,515	12,576
Observations	145,299	363,016
R^2	.47	.41
F statistic	69.0	56.0

Note. Average treatment effect for the frequency experiment, controlling for baseline consumption and time, with confidence intervals in brackets. The dependent variable was household water consumption in litres per day. ‘Treatment period’ refers to the first 12 months of the experiment where the treatment was delivered to all households, and ‘All months’ refers to the inclusion of all the months of the experiment combined for analysis. Rows 4 – 6 are pairwise *t*-tests for equality of coefficients to identify differences between conditions.
$^{**}p < .001$.

Figure 2

Water Use by Frequency Condition (Study 1)



Note. Event study graph of the average treatment effect in litres per day across all months before and after the intervention for all three treatment conditions in the frequency experiment. The coefficients and confidence intervals were generated from interacting treatment assignments with monthly dummy variables while controlling for baseline consumption.

Study 2: Message Medium

To address the second aim, we compared the impact of email and paper mail as delivery mediums for feedback interventions on household water consumption. Table 3 shows that only households that received paper mailings reduced water use compared to the control. Paper mailings were also more effective than email. Households in the paper condition reduced their consumption by 5 litres a day (-1.8%). Households that received paper reports saw an initial reduction in consumption that persisted over time, while those that received email reports never meaningfully reduced consumption (Figure 3).

PsychOpen GOLD

Table 3

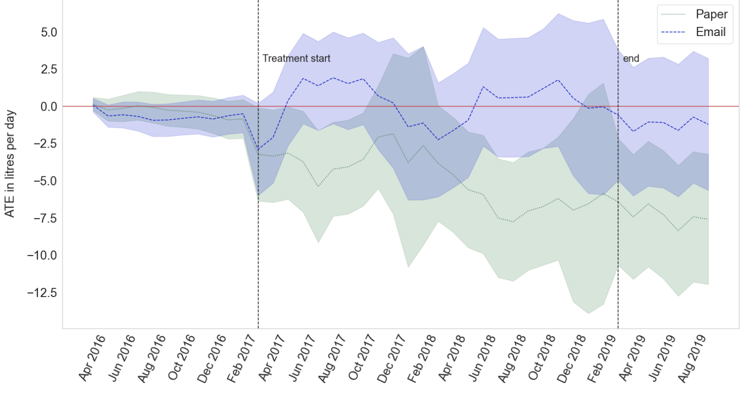
Linear Regression of Message Medium on Water Use (Study 2)

	Household water use change in litres per day	
Group	**First 12 months**	**All 31 Months**
Email	1.41 [-0.86, 3.68]	1.20 [-1.15, 3.56]
Paper	-3.45** [-5.83, -1.08]	-5.14** [-7.51, -2.77]
Email = Paper	-4.86** [-7.71, -2.01]	-6.35** [-9.15, -3.55]
Number of households	18,283	18,303
Observations	225,737	489,977
R^2	.52	.49
F statistic	27.4	28.2

Note. Average treatment effect for the medium experiment with confidence intervals in brackets. The dependent variable was household water consumption in litres per day. 'Email = Paper' refers to pairwise *t*-tests for equality of coefficients to test differences between conditions.
**p < .001.

Figure 3

Water Use by Message Medium (Study 2)



Note. This event study graph shows the pre- and post-treatment period for paper and email messages. Coefficients and confidence intervals were generated by interacting treatment assignment with monthly dummy variables while controlling for baseline consumption. Dashed vertical lines indicate when home water reports were delivered (about 6-monthly).

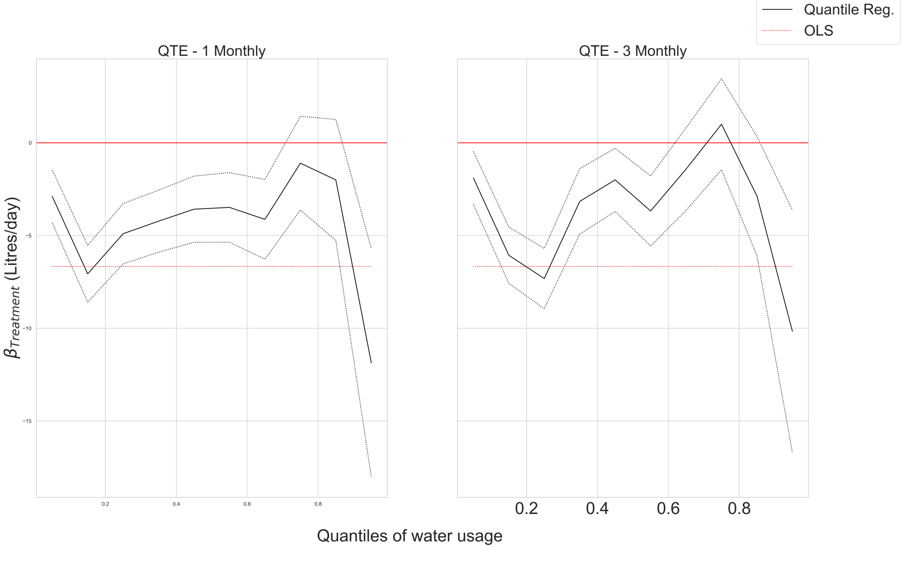
Heterogeneity

To understand how the intervention affects people differently, heterogeneity of treatment effects was studied using quantile regressions and conditional average treatment effect (CATE). Studying the heterogeneity of treatment effects is important for understanding how the intervention affects people differently.

Study 1: Frequency — Figure 4 of the quantile treatment effects for both the monthly and 3-monthly conditions shows that the entire distribution of consumption is not shifting downwards, and so there was no aggregate reduction in consumption across all households. Rather, the distribution between the 0.5 and 0.8 quantiles seems to be attenuating the treatment effect, with a small amount of mass in the 3-monthly conditions moving above zero, suggesting an increase in consumption. The bottom tail is moving downwards but not by much, which is expected given that there is a lower bound to consumption. The largest reduction is seen in the upper tail of the distribution. This upper tail could be households with high consumption, but due to rank and invariance, it is difficult to draw such conclusions. Instead, we can look at the results of interacting baseline consumption with the different experimental groups to get the conditional average treatment effect (CATE). Figure 5 of the CATE shows that for the monthly experimental group, the greatest reduction in consumption was in the fourth to fifth and the seventh deciles. The 3-monthly group shows the most reduction in the eighth and ninth deciles, but these are not statistically significant, probably due to being underpowered. There was no statistically significant main effect for the 6-monthly group, and it was therefore omitted from the analysis of heterogeneity.

Figure 4

Quantile Treatment Effects of the Monthly and 3-monthly conditions (Study 1)

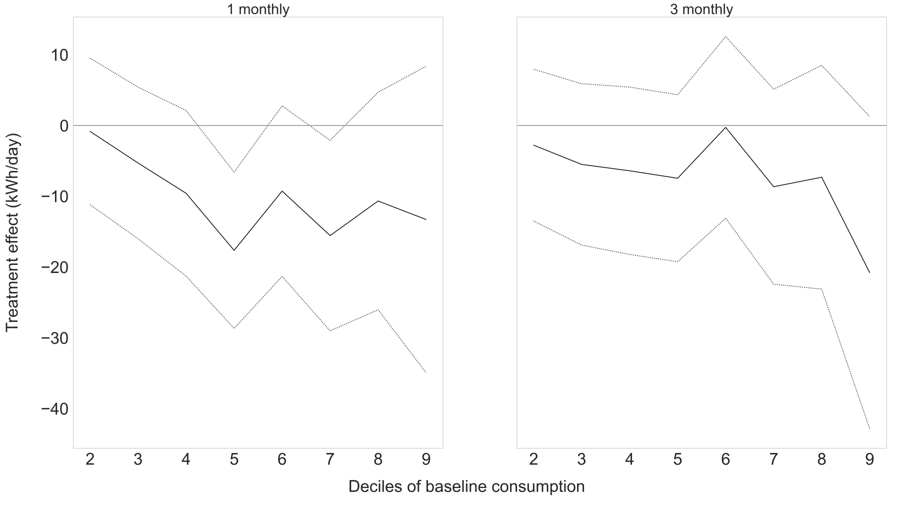


Note. The dashed red line refers to the coefficient of the average treatment effect of the standard OLS. The dashed black line refers to the confidence intervals of the quantile treatment effect. The lower quantiles represent lower levels of consumption.

https://doi.org/10.5964/gep.11705

Figure 5

Conditional Average Treatment Effect of the Monthly and 3-Monthly Conditions (Study 1)



Note. The dashed black line refers to the confidence intervals of the conditional average treatment effect (CATE). The lower the decile, the lower the consumption at baseline.

Study 2: Medium — As expected, the quantile regression in Study 2 showed that the lowest-use households reduced their use the least due to a floor effect (Figure A2). The middle distribution shows an S-shape, but it should not be over-interpreted as this minor pattern could be noise. Assuming the middle of the distribution was flat as expected, households used 6 fewer litres per day on average. The top of the distribution shows an attenuation of the treatment effect. This is further demonstrated by looking at the CATE in Figure A3, which shows that while the top and bottom deciles of baseline consumption did not reduce consumption, the greatest reductions came from households in the middle deciles, with those in the sixth decile reducing by 16 litres per day.

Cost Effectiveness

In addition to evaluating intervention efficacy, we gauged the cost-benefit paradigm of the feedback mechanisms to provide insights into potential savings and efficiencies for utility companies. Based on the results above, receiving a paper message every three months for one year reduced water use by 7 litres per household per day (-2,555 litres per year). The cost of generating, printing and posting a report was £0.80 per household per report (£3.20 per year). Therefore, utility companies could save 798 litres for every £1 spent per household on paper messages. Based on Anglian Water's pricing for 2021–2022 of £1.60 per cubic metre, an intervention like this would have reduced a household water bill by £4 per month. This is small per household but large in aggregate. According to Nauges and Whittington (2019), the cost savings for a water utility from a similar feedback intervention come from requiring less electricity and chemicals to purify and deliver water. At most, these electricity and chemical costs are 25% of the operation and maintenance, or 10% of the mean cost of producing one cubic meter of clean, delivered water (Nauges & Whittington, 2019).

Cost-effectiveness can also be increased by targeting households based on previous use. For example, Study 1 households with medium-high use (7th decile of baseline consumption) reduced consumption by 11 litres per day, equivalent to saving 1,255 litres for every £1 spent on messaging. If the utility were to only target this segment of 943 households, they could save 3,786,145 litres of water annually for £3,018[1] in messaging costs, resulting in -1,255 litres saved per £1 (comparable to Hodges et al., 2020).

Discussion

The current paper tested how frequency and message medium affect social norms messaging on household water use. Below, we summarise and clarify how message and medium affected household water consumption in the United Kingdom. To consider the generalisability of the results, we compare them to international studies and anchor the findings within previous findings. We also discuss limitations, suggestions for future research, and implications for practitioners.

Main Intervention Effects

Both interventions successfully reduced household water consumption in the United Kingdom. Households in Study 1 reduced water consumption by 2%, and in Study 2 by 1.8%, consistent with the 1.8% reduction found by a previous study in the United Kingdom that similarly tested out the effect of social norms feedback on household water consumption (Ramli, 2021). These are smaller effects than the 5% observed in the United States (Brent et al., 2015; Ferraro & Price, 2013). Several factors might account for this disparity. Firstly, U.S. households use more water than U.K. households, which means more opportunities for large reductions (Andor et al., 2020). With a greater baseline consumption, even identical behavioural interventions may result in larger decreases. Furthermore, cultural, weather, infrastructure, and regulation differences may also determine water use. Future research can investigate this complexity, but it is already clear that context is necessary to interpret intervention effectiveness (Andor et al., 2020).

1) £3.20 (cost of delivering intervention per household per year) x 943 households (in the 7th decile of baseline consumption) = £3017.60

PsychOpen GOLD

Heterogeneity

Heterogeneity was observed in all five treatment conditions for both studies using both quantile regressions and CATE. Unfortunately, due to rank invariance and insufficient sample size for the multiple interaction effects, no statistical conclusions can be made about the nature of the heterogeneity.

Frequency

We aimed to explore whether there is a moderating effect of treatment frequency and, similar to Allcott (2011), we did not find significant differences between treatment frequencies. One argument for higher frequency is found in Allcott and Rogers (2014), where the treatment effect is strongest when households receive the reports and decay soon after until the next report. Therefore, by delivering the treatment at a higher frequency, there is less decay of the treatment effect, and so greater overall savings can be achieved. This could not be formally tested here as daily meter reads were not available, and we were unable to stop treatment delivery for a random selection of households in both the monthly and 3-monthly groups. Additionally, while the 6-monthly group did not have an effect in Study 1, delivering the treatment every 6 months did have an effect in Study 2. This difference may have been due to small context differences.

Medium

The results from Study 2 confirmed that interventions through email were not effective, consistent with Dolan and Metcalfe (2015). Emails may be easier to ignore than letters since one can often assume the content of an email just by looking at the sender's name or the subject line. Furthermore, the digital realm can be distracting, with users frequently multitasking, which reduces their focus on any specific action, such as reading the intervention email. In contrast, a physical letter in a plain envelope is more ambiguous as to the sender and content until it is opened, and letters may convey more importance or officialness than an email. Furthermore, the email treatment required people to click through to a website to receive the feedback, and this additional layer of friction may have deterred some of those, even with initial interest, thus decreasing how many households were treated. Another dimension is that, whereas a single person normally reads emails, letters might be placed in a common area, such as a table, where multiple household members could see them. This communal exposure could bring about conversations or plans that boost the impact of the message, which might also explain the effectiveness of the physical mailing in Hodges et al. (2020).

Limitations

There are notable limitations to the current study. Firstly, the lack of demographics about the households means it was not possible to identify heterogeneities across other dimensions. Second, the period after the last paper report was sent in Study 1 had some bias: households in the monthly condition that shared their email address continued to receive the treatment via email. These numbers were very low (~500, seen in Figure A1), and the monthly condition had an equivalent reduction from the other treatment conditions, so this additional email treatment was likely unimportant. Finally, the high cost of mailing is another major limitation. Future studies could explore ways to make emails more engaging. This could reduce the cost of norm-based interventions. For example, email-based interventions could be more effective if they are personalised or made easier to access the feedback (e.g., not needing a login to a separate website). Study 2 only compared the mediums of email and paper. However, given how fast digital communication is changing, future studies should explore alternative methods that are engaging and cost-effective, potentially such as apps using real-time feedback. Finding a medium that balances engagement and cost-effectiveness would be an important contribution to the evidence on norm-based interventions.

Future Directions

Longer-term studies that extend the intervention over numerous years and that could randomise the stopping of intervention delivery could determine the longevity and decay rates of feedback effects at different messaging frequencies. This could reveal whether more frequent feedback leads to more persistent, long-term behavioural change. Another possibility is to use higher granularity metering with daily or even hourly meter scans. These high frequency meters are common in energy (e.g., Allcott & Rogers, 2014) but have yet to be rolled out widely in the water sector. Such an approach could help evaluate how consumption patterns change immediately after feedback and how they change over shorter time periods. Moreover, as smart meters and/or smart devices become more common, feedback could also be delivered in real-time at the point of the behaviour, for example, during a shower (e.g., Tiefenbeck et al., 2019), and new research can compare periodic messages like that in the current studies with real-time feedback.

Conclusion

In the English towns during this period, these results suggest the most effective intervention for social norms-based feedback on water use is to deliver paper mail reports every 3 or 6 months over an extended period and to target households with medium-to-high consumption. Those with low baseline consumption might require different interventions. It is unknown whether the same frequency and medium results would hold for other types of persuasive messages or in different towns within the United Kingdom, and other locations globally. The effects of a physical letter may be less generalisable across locations. We speculate that the tactile experience and communal visibility of letters in a household environment could increase engagement. However, as digital communication continues to dominate and traditional mail becomes less common, future research should track changes in how recipients interact with physical mail. For example, people may engage differently with physical mail as it becomes less common.

In conclusion, we found that frequency and medium can be important and cost-effective elements of successful messaging interventions on household water use. A robust, generalizable set of findings about the most effective message content and delivery across diverse demographics and areas could reduce environmental damage and costs for utility companies. Such demand-side changes are critical for mitigating the changing climate and ensuring water security on a thriving planet (Creutzig et al., 2022).

Openness and Transparency Statements

The present article has been checked by its handling editor(s) for compliance with the journal's open science and transparency policies. The completed *Transparency Checklist* is publicly available at: http://doi.org/10.23668/psycharchives.21505

Author Contributions.
Ukasha Ramli: Conceptualization. Data curation. Formal analysis. Funding acquisition. Investigation. Methodology. Project administration. Resources. Software. Visualization. Writing – original draft.
Cameron Brick: Project administration. Supervision. Validation. Writing – review & editing.
Benjamin Abera: Project administration. Validation. Writing – review & editing.

Acknowledgments. The authors have no additional (i.e., non-financial) support to report.

Funding. The authors have no funding to report.

Competing Interests. Ukasha Ramli served as a consultant for Advizzo during the course of these experiments. Cameron Brick was a member of the Editorial Board for Global Environmental Psychology at the time this manuscript was submitted and accepted. The authors affirm that these roles had no influence on the peer review process or the editorial decision regarding this paper.

Ethics Statement. This study received ethical approval from the London School of Economics's Department of Psychology and Behavioural Sciences. The researchers received data of this experiment from the utility company. Data from all participating households were anonymised and all data transfer process and storage complied with the EU's Global Data Protection Regulation. Participating households could opt-out of receiving the communications at any time over the course of the study.
These studies were also reported in the doctoral dissertation of the principal author. The analyses and writing have been improved and updated here through peer review.
Additionally, we have reported all measures, conditions, data exclusions, and the calculations used to determine power based on available sample size. All of these points can be found in the Methods section.

Diversity Statement. In the list below, the check mark (☑) indicates which steps were taken to increase diversity within the context of this paper. Steps that were not taken or did not apply are unmarked (☐).

☐ Ethnically or otherwise diverse sample(s)
☐ Gender balanced sample(s)

- [ ] Inclusive gender measure
- [x] Inclusive materials
- [x] Sampling justification
- [ ] Extensive sample description
- [x] Discussion of generalizability
- [ ] Diverse reference list
- [x] Underprivileged / minority author(s)
- [x] Early career author(s)
- [ ] Degree of privilege/marginalization considered in authorship order
- [ ] Author(s) from sampled population (avoiding 'helicopter science')

Data Availability. The data is owned by the utility company and the tech company that ran the experiment. Anonymised version of the data could be shared on request.

Supplementary Materials. The following table provides an overview of the accessibility of supplementary materials (if any) for this paper.

Type of supplementary material	Availability/Access
Data The anonymised data can be shared upon request to the corresponding author.	–
Code Analysis code (Python)	Ramli et al. (2025)
Material No material was provided from the study.	–
Study/Analysis preregistration Study was not preregistered.	–

Badges for Good Research Practices.
Open data: NO.
Open code: YES.
Open materials: NO.
Preregistration: NO.
Diversity statement: YES.
Note: YES = the present article meets the criteria for awarding the badge. NO = the present article does not meet the criteria for awarding the badge *or* the criteria are not applicable.

References

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics, 95*(9–10), 1082–1095. https://doi.org/10.1016/j.jpubeco.2011.03.003

Allcott, H. (2015). Site selection bias in program evaluation. *Quarterly Journal of Economics, 130*(3), 1117–1165. https://doi.org/10.1093/qje/qjv015

Allcott, H., & Mullainathan, S. (2010). Behavior and energy policy. *Science, 327*(5970), 1204–1205. https://doi.org/10.1126/science.1180775

Allcott, H., & Rogers, T. (2014). The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review, 104*(10), 3003–3037. https://doi.org/10.1257/aer.104.10.3003

Andor, M. A., & 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. https://doi.org/10.1016/j.ecolecon.2018.01.018

Andor, M. A., Gerster, A., Peters, J., & Schmidt, C. M. (2020). Social norms and energy conservation beyond the US. *Journal of Environmental Economics and Management, 103*, Article 102351. https://doi.org/10.1016/j.jeem.2020.102351

Ayres, I., Raseman, S., & Shih, A. (2013). Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage. *Journal of Law Economics and Organization, 29*(5), 992–1022. https://doi.org/10.1093/jleo/ews020

Bhanot, S. P. (2017). Rank and response: A field experiment on peer information and water use behavior. *Journal of Economic Psychology, 62*, 155–172. https://doi.org/10.1016/j.joep.2017.06.011

Bich-Ngoc, N., & Teller, J. (2018). A review of residential water consumption determinants. In O. Gervasi, B. Murgante, S. Misra, E. Stankova, C. M. Torre, A. M. A. C. Rocha, D. Taniar, B. O. Apduhan, E. Tarantino, & Y. Ryu (Eds.), *Computational science and its applications – ICCSA 2018. ICCSA 2018. Lecture notes in computer science* (Vol. 10964). Springer. https://doi.org/10.1007/978-3-319-95174-4_52

Bloomberg. L. P. (2023). *OPOWER Inc.* www.bloomberg.com/profile/company/OPWR:US#xj4y7vzkg

Bonan, J., Cattaneo, C., d'Adda, G., & Tavoni, M. (2020). The interaction of descriptive and injunctive social norms in promoting energy conservation. *Nature Energy, 5*(11), 900–909. https://doi.org/10.1038/s41560-020-00719-z

Brent, D. A., Cook, J. H., & 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. https://doi.org/10.1086/683427

Caretta, M. A., Mukherji, A., Arfanuzzaman, M., Betts, R. A., Gelfan, A., Hirabayashi, Y., Lissner, T. K., Liu, J., Lopez Gunn, E., Morgan, R., Mwanga, S., & Supratid, S. (2023). Water. In H.-O. Pörtner, D. C. Roberts, M. Tignor, E. S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, & B. Rama (Eds.), *Climate Change 2022: Impacts, adaptation and vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 551–712). Cambridge University Press. https://doi.org/10.1017/9781009325844.006

Carlsson, F., Jaime, M., & Villegas, C. (2021). Behavioral spillover effects from a social information campaign. *Journal of Environmental Economics and Management, 109*, Article 102325. https://doi.org/10.1016/j.jeem.2020.102325

Creutzig, F., Niamir, L., Bai, X., Callaghan, M., Cullen, J. M., Díaz-José, J., Figueroa, M. J., Grübler, A., Lamb, W. F., Leip, A., Masanet, E., Mata, É., Mattauch, L., Minx, J. C., Mirasgedis, S., Mulugetta, Y., Nugroho, S. B., Pathak, M., Perkins, P. E., . . .Ürge-Vorsatz, D. (2022). Demand-side solutions to climate change mitigation consistent with high levels of well-being. *Nature Climate Change, 12*(1), 36–46. https://doi.org/10.1038/s41558-021-01219-y

Costa, D. L., & Kahn, M. E. (2013). Energy conservation 'nudges' and environmentalist ideology: Evidence from a randomized residential electricity field experiment. *Journal of the European Economic Association, 11*(3), 680–702. https://doi.org/10.1111/jeea.12011

Darby, S. (2006). *The effectiveness of feedback on energy consumption* (p. 24). Environmental Change Institute. www.eci.ox.ac.U.K./research/energy/downloads/smart-metering-report.pdf

DeOreo, W. B., Mayer, P., Dziegielewski, B., & Kiefer, J. (2016). *Residential end uses of water, Version 2.* Water Research Foundation. www.circleofblue.org/wp-content/uploads/2016/04/WRF_REU2016.pdf

Dolan, P., & Metcalfe, R. (2015*). Neighbors, knowledge, and nuggets: Two natural field experiments on the role of incentives on energy conservation* (Centre for Economic Performance Discussion Papers, 1222). Centre for Economic Performance. https://doi.org/10.2139/ssrn.2589269

Domene, E., & Saurí, D. (2006). Urbanisation and water consumption: Influencing factors in the Metropolitan Region of Barcelona. *Urban Studies, 43*(9), 1605–1623. https://doi.org/10.1080/00420980600749969

Energy Saving Trust. (2013). *At home with water.* Defra/Procter and Gamble/Thames Water/ Consumer Council for Water/Save Water Save Money Ltd. www.energysavingtrust.org.uk/sites/default/files/reports/AtHomewithWater%287%29.pdf

Ferraro, P. J., & Miranda, J. J. (2013). Heterogeneous treatment effects and mechanisms in information-based environmental policies: Evidence from a large-scale field experiment. *Resource and Energy Economics, 35*(3), 356–379. https://doi.org/10.1016/j.reseneeco.2013.04.001

Ferraro, P. J., Miranda, J. J., & Price, M. (2011). The persistence of treatment effects with norm-based policy instruments: Evidence from a randomized environmental policy experiment. *American Economic Review, 101*(3), 318–322. https://doi.org/10.1257/aer.101.3.318

Ferraro, P., & Price, M. (2013). Using non-pecuniary strategies to influence behavior: Evidence from a large-scale field experiment. *Review of Economics and Statistics, 95*(1), 64–73. https://doi.org/10.1162/REST_a_00344

Fischer, C. (2008). Feedback on household electricity consumption: A tool for saving energy? *Energy Efficiency, 1*(1), 79–104. https://doi.org/10.1007/s12053-008-9009-7

Geelen, D., Mugge, R., Silvester, S., & Bulters, A. (2019). The use of apps to promote energy saving: A study of smart meter–related feedback in the Netherlands. *Energy Efficiency, 12*(6), 1635–1660. https://doi.org/10.1007/s12053-019-09777-z

Global Commission on the Economics of Water. (2023). *Turning the tide: A call to collective action.* https://watercommission.org/wp-content/uploads/2023/03/Turning-the-Tide-Report-Web.pdf

Hodges, H., Kuehl, C., Anderson, S., Ehret, P., & Brick, C. (2020). How managers can reduce household water use through communication: A field experiment. *Journal of Policy Analysis and Management, 39*(4), 1076–1099. https://doi.org/10.1002/pam.22246

Ilgen, D., Fisher, C., & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology, 64*(4), 349–371. https://doi.org/10.1037/0021-9010.64.4.349

Karlin, B., Zinger, J. F., & Ford, R. (2015). The effects of feedback on energy conservation: A meta-analysis. *Psychological Bulletin, 141*(6), 1205–1227. https://doi.org/10.1037/a0039650

Kummu, M., Guillaume, J., de Moel, H., Eisner, S., Flörke, M., Porkka, M., Siebert, S., Veldkamp, T. I. E., & Ward, P. J. (2016). The world's road to water scarcity: Shortage and stress in the 20th century and pathways towards sustainability. *Scientific Reports, 6*, Article 38495. https://doi.org/10.1038/srep38495

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254–284. https://doi.org/10.1037/0033-2909.119.2.254

Lancaster, K. (1975). The theory of household behavior: Some foundations. In S. V. Berg (Ed.), *Annals of economic and social measurement* (Vol. 4, No. 1; pp. 5–21). National Bureau of Economic Research. www.nber.org/chapters/c10216

Lam, C. F., DeRue, D. S., Karam, E. P., & Hollenbeck, J. R. (2011). The impact of feedback frequency on learning and task performance: Challenging the 'more is better' assumption. *Organizational Behavior and Human Decision Processes, 116*(2), 217–228. https://doi.org/10.1016/j.obhdp.2011.05.002

Mahone, A., & Haley, B. (2011). *Overview of residential energy feedback and behavior-based energy efficiency.* Energy and Environmental Economics.

Moore, P. (2015, June 4). *Inbox Zero: Most hate having unread e-mails, a few let them pile up.* YouGov. today.yougov.com/topics/lifestyle/articles-reports/2015/06/04/inbox-zero

Nauges, C., & Whittington, D. (2019). Social norms information treatments in the municipal water supply sector: Some new insights on benefits and costs. *Water Economics and Policy, 5*(3), Article 1850026. https://doi.org/10.1142/S2382624X18500261

Nolan, J. M., Schultz, P. W., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2008). Normative social influence is underdetected. *Personality & Social Psychology Bulletin, 34*(7), 913–923. https://doi.org/10.1177/0146167208316691

Ramli, U. (2021). Social norms based eco-feedback for household water consumption. *Sustainability, 13*(5), Article 2796. https://doi.org/10.3390/su13052796

Ramli, U., Brick, C., & Abera, B. (2025). *Optimal delivery of social norms feedback to reduce household water consumption* [Code: Colchester AnalysisGEP.ipynb]. OSF. https://osf.io/j6xmc/

Ramli, U., & Laffan, K. (2022). Double trouble: Concurrently targeting water and electricity using normative messages in the Middle East. *Energy Research & Social Science, 88*, Article 102496. https://doi.org/10.1016/j.erss.2022.102496

Salmoni, A. W., Schmidt, R. A., & Walter, C. B. (1984). Knowledge of results and motor learning: A review and critical reappraisal. *Psychological Bulletin, 95*(3), 355–386. https://doi.org/10.1037/0033-2909.95.3.355

Sanguinetti, A., Dombrovski, K., & Sikand, S. (2018). Information, timing, and display: A design-behavior framework for improving the effectiveness of eco-feedback. *Energy Research & Social Science, 39*, 55–68. https://doi.org/10.1016/j.erss.2017.10.001

Schultz, P. W., Estrada, M., Schmitt, J., Sokoloski, R., & Silva-Send, N. (2015). Using in-home displays to provide smart meter feedback about household electricity consumption: A randomized control trial comparing kilowatts, cost, and social norms. *Energy, 90*(1), 351–358. https://doi.org/10.1016/j.energy.2015.06.130

Schultz, P. W., Messina, A., Tronu, G., Limas, E. F., Gupta, R., & Estrada, M. (2016). Personalized normative feedback and the moderating role of personal norms: A field experiment to reduce residential water consumption. *Environment and Behavior, 48*(5), 686–710. https://doi.org/10.1177/0013916514553835

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science, 18*(5), 429–434. https://doi.org/10.1111/j.1467-9280.2007.01917.x

Sønderlund, A. L., Smith, J. R., Hutton, C., & Kapelan, Z. (2014). Using smart meters for household water consumption feedback: Knowns and unknowns. *Procedia Engineering, 89*, 990–997. https://doi.org/10.1016/j.proeng.2014.11.216

Tiefenbeck, V., Wörner, A., Schöb, S., Fleisch, E., & 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. https://doi.org/10.1038/s41560-018-0282-1

UK Government. (2023, October 25). *Appendix A: Smart metering in draft water resources management plans.* GOV.UK. www.gov.uk/government/publications/a-review-of-englands-draft-regional-and-water-resources-management-plans/appendix-a-smart-metering-in-draft-water-resources-management-plans

UN Water. (2021). *Summary Progress Update 2021 — SDG 6: Water and sanitation for all.* United Nations. https://www.unwater.org/publications/summary-progress-update-2021-sdg-6-water-and-sanitation-all

Webb, T. L., Chang, B. P. I., & Benn, Y. (2013). The ‘Ostrich Problem’: Motivated avoidance or rejection of information about goal progress. *Social and Personality Psychology Compass, 7*(11), 794–807. https://doi.org/10.1111/spc3.12071

Appendix

Figure A1

Time and Number of Reports (Study 1)

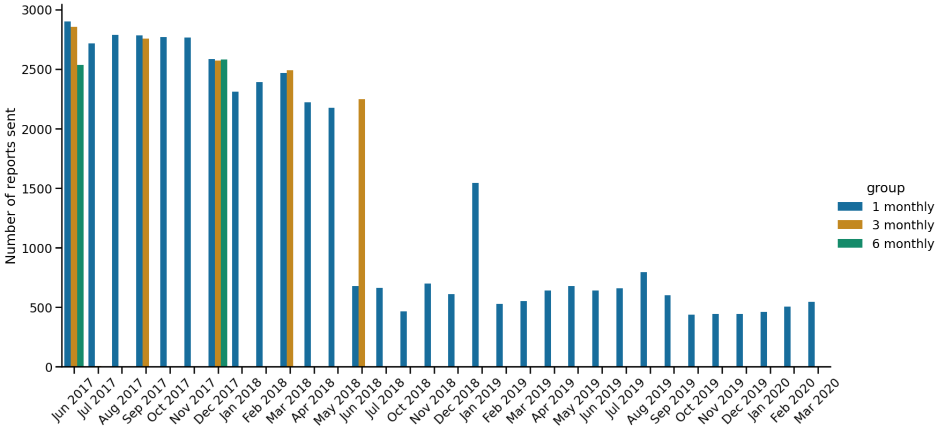
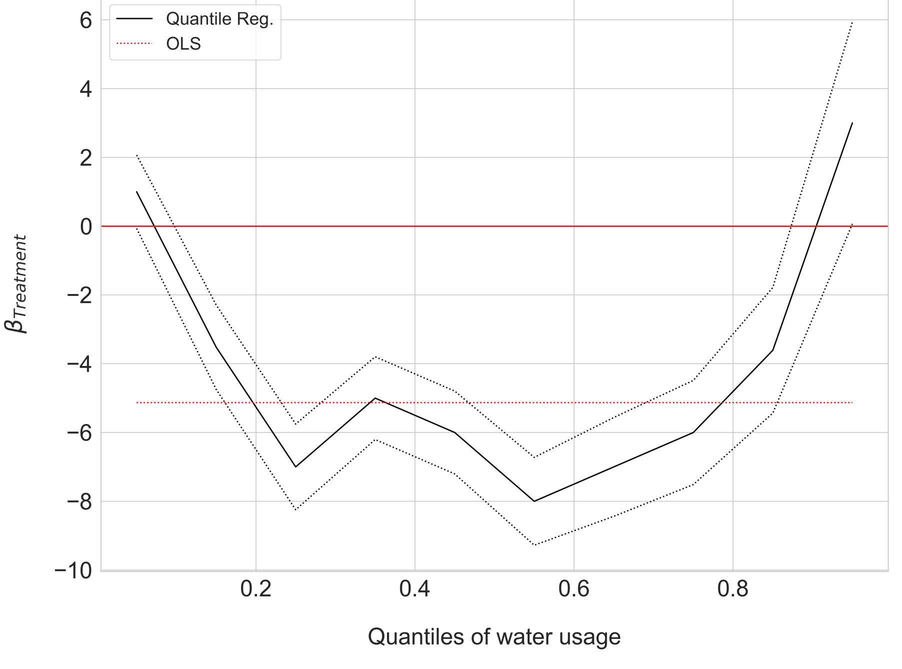


Figure A2

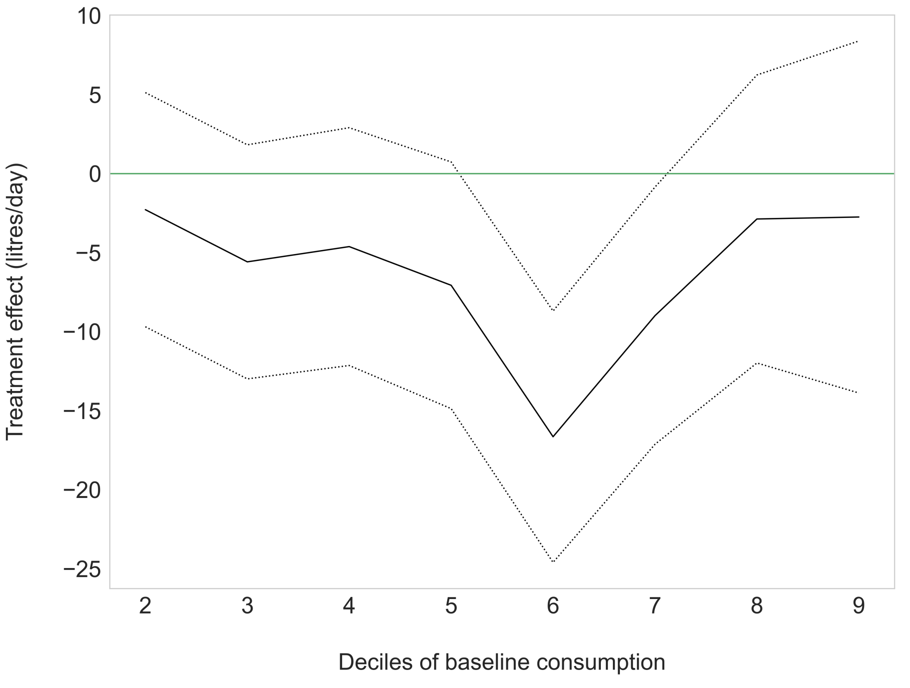
Quantile Treatment Effects of the Paper Condition



Note. The dashed red line refers to the coefficient of the average treatment effect of the standard OLS. The dashed black line refers to the confidence intervals of the quantile treatment effect. The lower quantiles represent lower levels of consumption.

Figure A3

Conditional Average Treatment Effect of the Paper Mail Condition



Note. The dashed black line refers to the confidence intervals of the conditional average treatment effect (CATE).

Table A1

Balance Checks for Attrition and Baseline Consumption (Studies 1 & 2)

	Regression Coefficient (Standard Errors)		
Experiment/condition	**Baseline consumption**	**Attrition: # treatment months**	**Attrition: binary of complete treatment months**
Frequency (Year one)			
1 monthly	-6.31 (4.6)	0.167 (0.10)	-0.00 (0.00)
3-monthly	-5.36 (4.62)	0.080 (0.10)	0.00 (0.00)
6-monthly	-2.33 (4.40)	0.177 (0.10)	0.00 (0.00)
3-monthly = 1 monthly	-0.94 (4.68)	0.087 (0.10)	0.00 (0.00)
6-monthly = 1 monthly	-3.98 (4.47)	0.34** (0.10)	0.00 (0.00)
6-monthly = 3-monthly	-3.03 (4.52)	0.26 (0.10)	0.00 (0.00)

	Regression Coefficient (Standard Errors)		
Experiment/condition	**Baseline consumption**	**Attrition: # treatment months**	**Attrition: binary of complete treatment months**
Medium			
Email	1.83 (2.01)	0.07 (0.06)	-0.00 (0.00)
Paper	-1.21 (2.02)	0.09 (0.06)	-0.00 (0.00)
Email = Paper	-3.05 (2.32)	0.02 (0.07)	-0.00 (0.00)

Note. The attrition by number of months are *t*-tests using the number of treatment period months available as the dependent variable. The binary attrition was created by categorising households based on whether they had complete data in the treatment period.

$^{**}p < .01$.