TU Delft

Delft University of Technology

The use of apps to promote energy saving

A study of smart meter–related feedback in the Netherlands

Geelen, Daphne; Mugge, Ruth; Silvester, Sacha; Bulters, Annemieke

DOI
10.1007/s12053-019-09777-z

Publication date
2019

Document Version
Other version

Published in
Energy Efficiency (online)

Citation (APA)
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 (online)*, *12*(6), 1635-1660. https://doi.org/10.1007/s12053-019-09777-z

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Energy Efficiency
https://doi.org/10.1007/s12053-019-09777-z

ORIGINAL ARTICLE

The use of apps to promote energy saving: a study of smart meter–related feedback in the Netherlands

Daphne Geelen · Ruth Mugge · Sacha Silvester · Annemieke Bulters

Received: 23 May 2017 / Accepted: 13 November 2018
© The Author(s) 2019

Abstract Feedback systems with direct feedback have shown to be effective in stimulating households to change their energy consumption levels. This research is one of the first to explore the use of apps to influence household energy use. Compared to dedicated in-home displays, smartphone/tablet apps provide a low-cost and simple design solution for making energy feedback available. This research consisted of three studies conducted with different samples within a selection of households where a smart meter was installed as part of the smart meter implementation program in the Netherlands. First, for a period of 16 months, electricity and gas consumption levels were measured for a large sample of households ($n = 519$) divided into an application user group and a reference group. Second, questionnaires ($n = 270$) provided insight in how people used the applications and to what extent the applications increased households' insight in their energy consumption and stimulated behavior changes. Third, interviews ($n = 12$) were held to obtain more in-depth insight. In the sample with measured energy consumption, we did not find a significant reduction in electricity and gas consumption during this research. Yet in the questionnaires, the application users reported more energy awareness and indicated to have made more investments and changes in their behavior than the reference group. Most app users started using the first app they found and did not explore the other options. The interview results indicate that, after an initial learning period, the app was used to monitor the electricity and gas consumption levels, rather than to lower them. In line with other research into feedback, the interview results suggest that the apps could be more effective with information that is more actionable and meaningful with respect to one's own specific situation and goals for the household. Further exploration is recommended with respect to how the design of such apps can encourage a wide audience not only to monitor their consumption, but also guide them in taking action to change their consumption levels.

Keywords Apps · Smart energy meters · Direct feedback · Energy saving · Behavior change · Households

D. Geelen · A. Bulters
Enexis B.V., 's-Hertogenbosch, Netherlands

D. Geelen
e-mail: daphne.geelen@enexis.nl

A. Bulters
e-mail: annemieke.bulters@enexis.nl

R. Mugge (✉) · S. Silvester
Faculty of Industrial Design Engineering, Delft University of Technology, Delft, Netherlands
e-mail: r.mugge@tudelft.nl

S. Silvester
e-mail: s.silvester@tudelft.nl

Introduction

A quarter of the final energy consumption in the European Union is accountable to households (European Union

Published online: 28 February 2019

2016). Households are therefore an important target group for policy measures aimed at energy efficiency. In the Energy Efficiency Directive communications, the European Commission states its goal to "ensure major energy savings for consumers" and its aims to "empower energy consumers to better manage consumption. This includes easy and free access to data on consumption through individual metering." (European Commission 2017). To make consumption information available to households, EU member states have to equip households with smart meters, whenever it is cost-effective (European Union 2012).

Why is it so important to make consumption information available via smart meters? A "smart meter" is an electronic energy meter that is remotely readable and can thus communicate consumption data continuously and automatically. Smart meters are introduced to provide operational gains for suppliers and distribution network operators, such as automated and accurate billing, fraud detection, and differentiated tariffs, as well as to provide consumers with more information about their energy consumption. The underlying premise for unlocking consumption data to consumers is that when consumers receive energy feedback information, they will gain insight in their energy consumption pattern and will be able to take measures to lower their consumption. These measures include changing behavior routines and investment in energy efficient technology, such as insulation and efficient appliances. In light of the ongoing energy transition, smart meters are also expected to play an important role in facilitating products and services that enable households to (automatically) adjust their consumption patterns to the availability of renewable electrical energy and to contribute to the balancing of supply and demand in the grid (Geelen et al. 2013; Giordano et al. 2011; Kobus et al. 2015a).

Energy feedback information can be provided in various ways and it is crucial that the data from the smart meter is communicated to consumers in a way that is useful and meaningful to them. The EU directives do not prescribe how the households should get insight into their energy consumption, and member states decide upon their own policies in this respect. For example, in the UK, each household in which a smart meter is installed is provided by default with a display that visualizes the energy consumption levels in real-time (Ofgem and DECC 2011). In the Netherlands, the policy is that households receive a bimonthly overview of their electricity consumption from their energy supplier (Van Elburg 2014). For more direct and detailed feedback, households can acquire one of the feedback systems that are available on the market. A market analysis looking into the development and uptake of feedback systems in the Netherlands concluded that there is a need for accessible and cheap systems to appeal to households with little interest or limited means to purchase a feedback system (Van Elburg 2014). The study also emphasized that it is important that feedback is simple to use for those consumers who are not Internet savvy or naturally interested in monitoring their energy consumption. A review of the products and services in the Dutch market (www.energieverbruiksmanagers.nl) confirms the need for cheap and accessible feedback systems. The majority of products and services on offer are smartphone/tablet software applications (apps) that are available for free. Whether these are effective for bringing about energy savings for a wide audience is still in question. There has been limited research on how energy feedback via apps can help households to save energy over an extended period. This study contributes to the knowledge on energy feedback systems by exploring how households engage with a simple and low-cost feedback system in the form of an app in a natural setting and the effects of such a system on their household energy consumption.

Literature review

Previous research has shown that providing feedback information about a household's energy consumption can lead to energy savings, see, e.g., the reviews by Abrahamse et al. (2005); Darby (2006); Fischer (2008); and Ehrhardt-Martinez et al. (2010). The reviews indicate that different ways of providing the feedback information as well as the duration of feedback intervention produce different achieved savings. Energy savings tend to be lower with longer intervention periods and in larger scale trials (Ehrhardt-Martinez et al. 2010). For example, energy savings from in-home displays are expected in the range of 3–5% when implemented in large-scale trials, rather than 6–10% found in small-scale studies (McKerracher and Torriti 2013). The reviews also show that differences in the achieved savings are also due in part to differences in the design of feedback systems and in how people engage with the feedback. The question now is therefore not so much whether feedback works, but how it works and which design factors influence the effectiveness of feedback (Karlin et al. 2015; Van Dam et al. 2010).

In the design of feedback systems, we have to consider the ways in which users engage with the feedback, how the feedback is presented, and which information is provided. One of the key factors for design and evaluation is the immediacy with which the feedback is available. Often, the distinction is made between what Darby (2006) identified as "direct feedback" versus "indirect feedback." Direct feedback is directly available to the user—on demand—and presents real-time information. Indirect feedback becomes available only with a delay that may range from a day to a year. Overall, direct feedback has shown to be more effective and the more immediate the feedback information is made available, the more effective it is (Abrahamse et al. 2005; Darby 2006; Fischer 2008; Ehrhardt-Martinez et al. 2010; Karlin et al. 2015).

The direct availability of feedback information does not imply that users frequently consult the information, which is a separate and important factor. Systems with higher use frequencies have shown to be more effective (Kobus et al. 2015b). To achieve a high-use frequency, the feedback information should be easily accessible and provide information that is relevant for the user. Furthermore, additional functionality of the feedback system is suggested to motivate frequent use, such as the combination of feedback with thermostat control, weather information, or dynamic tariffs (Kobus et al. 2015b).

From the literature, it is clear that the design of the system should facilitate users' engagement with the feedback system a normal part of daily life. This implies that the feedback system should support positive dialogs between household members about energy management, and that the feedback system is accessible and attractive to all household members (Van Dam et al. 2012; Hargreaves et al. 2010).

To use the feedback for energy saving, the presentation of the information has to be understandable, relevant, and actionable for households (Van Dam 2013; Hargreaves et al. 2010; Geelen et al. 2013). For the measurement unit in which the feedback is expressed, users prefer costs (€, $) over energy units (kWh, m^3) (Karjalainen 2011). A note here is that Karlin et al. (2015) did not find significant moderating effects for different measurement units, which suggests that other design factors may be more influential. Showing numbers is not directly necessary for feedback to be understandable and relevant; color-based information has been found to be easily understandable and appreciated (Bonino et al. 2012) and ambient displays are also able to convey information that is actionable for users (Darby 2006). Ambient displays do not necessarily show numbers but give an impression about energy consumption based on, for example, changing colors, a flashing light, or a sound.

The granularity of the feedback, i.e., the level of detail represented, is also considered a factor that influences the effectiveness of feedback. Usually, the distinction is made between appliance-specific and aggregated consumption data. However, there is no consensus as to its effects. The results from studies by Karlin et al. (2015) and Kobus et al. (2015b) suggest that appliance-specific information is not always useful to a user, because the feedback may not tell users how to reduce consumption (Karlin et al. 2015) or because the accessibility of the information does not stimulate frequent use of the feedback (Kobus et al. 2015b).

Comparative information can help users to interpret their consumption data to create meaningful information. Historic consumption enables users to understand what their energy consumption pattern looks like (Anderson and White 2009; Fischer 2008) and to compare their current patterns (per day, week, month, year) with their past consumption. This can provide users with a personal norm as to what a normal or desirable consumption pattern is (Karlin et al. 2015). Another form of comparison is normative or peer-usage comparison, in which a household's consumption is benchmarked with that of other (similar) households. Although this form of comparison is often appreciated by users, making good comparisons is also complicated by differences among individual households (Stankovic et al. 2016). Historic consumption is considered to be more effective (Fischer 2008), though when compared with goal-based comparisons, i.e., goal-setting, Karlin et al. (2015) did not find significant effects for normative and historical comparisons. They suggest that goal-based comparison is more effective because it motivates users to focus their attention on the actions required to meet a goal that is relevant to themselves. In a similar vein, it is recommended to make interfaces goal-driven, i.e., producing actionable feedback that helps users to understand the extent to which a given goal is being approached and what actions may be required to meet the goal (Geelen et al. 2013).

The medium through which the feedback is provided largely defines the accessibility of the feedback. In this respect, in-home displays are considered most

promising, because they make feedback directly visible in the household. While in-home displays have shown promising results (e.g., Kobus et al. 2015b; Van Dam et al. 2010), we should be wary of overconfidence in these devices (Nilsson et al. 2014; Buchanan et al. 2015) because it is not the fact that people have a display, but the overall design of the display that influences the way people understand and use the feedback. Furthermore, these devices are quite expensive[1] and thus it is important to investigate whether low-cost feedback systems can also provide households with the needed insights to reduce their energy consumption.

This study contributes to the literature with insight in the effectiveness of a feedback system that uses a smartphone or tablet application as a medium to provide direct feedback. Little is known about the use and effectiveness of feedback systems in the form of an app, but an app provides a low-cost alternative to an in-home display and is therefore more attainable for a large user base than an in-home display.

Method

This research examined households in which a distribution network operator replaced their old energy meter with a smart meter. In a trial, the distribution network operator offered its customers a feedback system on the premise that this feedback system could help the households gain better insight in their energy consumption and save energy. This context provided a unique opportunity to observe the effects of a feedback system in a natural setting. To be able to see the effects of the feedback system, we made sure to collect data from a group of households with feedback systems as well as from a reference group that did not have them. In this chapter, we will discuss the characteristics of the feedback system and apps, the research design, and participant recruitment.

The feedback system and design of the apps

The apps that were tested in this research make use of electricity and gas consumption data that was measured by the newly installed smart meter in the participating households. The smart meter has a so-called P1-port, to which a P1-reader is connected. The P1-reader makes the measurement data from the smart meter available for the apps via the household's Wi-Fi-network.

The households that were offered a feedback system for this research could make use of two smartphone/tablet apps, a desktop (Windows) application, and an online application. The online application was principally used to connect the smartphone/tablet app and desktop application to the P1-reader. See Table 1 for an overview of the characteristics of the available energy feedback. The majority of the application users used a smartphone/tablet app (84%).

The smartphone/tablet apps were developed by different organizations. The researchers could not influence the design of the apps because the research project was initiated by a Dutch distribution network operator that is responsible for large parts of the electricity network in the Netherlands. The network operator's goal was to allow other organizations to develop their own apps based on their organization's preferences. As a result, the app features differed on several aspects, such as functionality and graphic design.

The apps were developed as a low-cost and low-maintenance solution with a short time for development. The developers therefore used design features that are commonly used for energy feedback. With respect to the guidelines described in the literature review, these apps included direct feedback aggregated to the total household consumption and historic comparisons. Both designs were built up in layers. The "home" screen focused on the actual consumption levels (direct feedback). By clicking on items on this screen or navigating the menu items, one could access more detailed information, such as historic consumption (indirect feedback). The two apps differed in "look and feel" and in the ways they presented energy consumption information. App A (Fig. 1) reports energy consumption with numbers and per counter of the smart meter, whereas App B (Fig. 2) reports energy consumption graphically in a dial, complemented with numbers. App B processes the raw consumption data more than App A, in order to provide meaning to the data and facilitate interpretation of the data graphically. The extra features of App B include a comparison between current consumption adding up over the day with the total consumption of the previous day (Fig. 2, left), comparison of historic daily consumption totals per week or month with the

[1] In the Netherlands, feedback systems with in-home displays are sold in the market for at least €100, sometimes in combination with a service fee of €18 to €55 per year. See www.energieverbruiksmanagers.nl (Accessed: 5 February 2018).

Table 1 Characteristics of the feedback system per test group (including reference group)

Smart meter	Reference group	Application user group			
		App A	App B	Desktop application	Online application
	Yes	Yes		Yes	
Indirect feedback	Bi-monthly overview from supplier*	Bi-monthly overview from supplier*		Bi-monthly overview from supplier*	
Direct feedback functions/-information	–	Current consumption from and delivery to the grid. Shown in a bar graph and numbers (W per 10 s, m^3 per hour)	Current consumption from the grid. Shown on a dial and numbers (kW and m^3)	Current consumption from and delivery to the grid. Shown in a bar graph and numbers (W & kWh per 10 s, m^3 per hour)	This tool must be used to enable the applications to receive data from the P1- reader.
		Historic consumption per hour, day and month per counter. Presented in numbers (kWh and m^3).	Total consumption of today and yesterday. Presented in bar graph and numbers (€)	Historic consumption per hour, day and month per counter. Presented in bar graphs and numbers (kWh and m^3).	Current consumption from and delivery to the grid. Shown in a bar graph and numbers (W & kWh per 10 s, m^3 per hour)
			Historic consumption per hour, day and month per counter. Presented in a bar graph (€), combined with consumption of the previous day/week/month.		Current meter readings (kWh and m^3)
			Tips for energy saving, accessed via a link to a website.		Download of historic consumption per counter.
			Setting a maximum consumption goal (in €) with an alert when the maximum is reached.		(No visualization of historic data)
To be obtained via	–	App Store Google Play (indicated in manual)	App Store Google Play	Feedback system website (indicated in manual)	Webpage of router (indicated in manual, required for installation of the P1-reader)
To install on	–	Smartphone Tablet	Smartphone Tablet	Computer (Windows O.S. only)	Not applicable. This is a website.

* Households with a smart meter receive a bi-monthly overview of their electricity and gas consumption by default from their energy provider

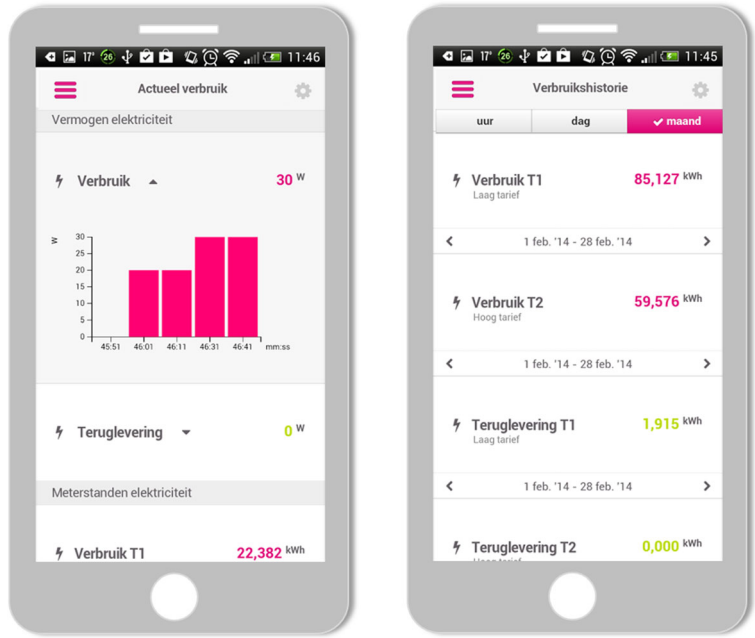

Fig. 1 Main screens of app A. Left: Current consumption of electricity and gas in kWh and m^3, depicted in a bar graph. Right: Historical consumption per hour, day, or month for each counter of the smart meter

consumption levels of the week or month before (Fig. 2, right), setting of an alarm for when a certain level of consumption costs for the day is reached (different screen), and tips for energy saving (directing to a website).

It is important to note that both the application user group and the reference group received bi-monthly consumption overviews (indirect feedback) throughout the study (see Table 1). By default, households in the Netherlands with smart meters each receive a bi-monthly

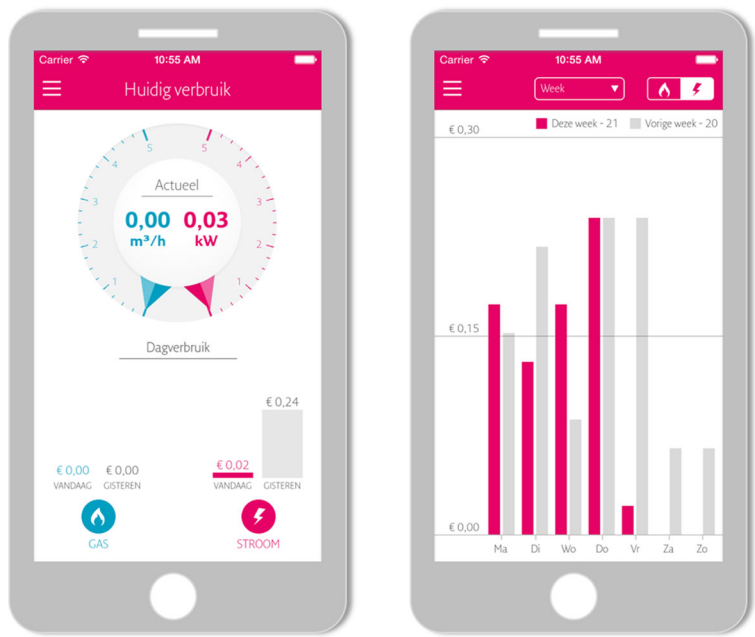

Fig. 2 Main screens of app B. Left: Current consumption of electricity and gas in kWh and m^3 depicted as a speedometer and comparison between today's and yesterday's consumption in Euros. Right: Historical consumption in Euros per day (pink) compared with days of the previous weeks (gray)

consumption overview. This overview is provided by the household's energy provider. Because this bi-monthly overview is provided to all participants in the research, we have included questions in the questionnaires and interviews about it in order to evaluate the effect of the apps in comparison to the bi-monthly overview.

Research setup

The conceptual model for this research (Fig. 3) follows the reasoning underpinning the policy for smart meter implementation, namely that the use of the feedback system results in increased insight in one's energy consumption and thereby enables a household to change daily energy-related behavior and/or make changes to the home that result in reduced energy consumption. The continued use of the feedback system enables the user to increase the learning (insight) about the household's energy consumption and perform further changes in daily behavior or investments. This conceptual model is used as a basis for the data collection.

This research is quasi-experimental, given that we were bound to the possibilities and decisions made by the network operator for the research methods and participant recruitment participants per method. The data collection was organized as three studies with different though complementary goals, methods, and samples.

The three studies include the following:

1. Measurement of electricity and gas consumption to investigate, with a large sample, whether the availability of the feedback system has an effect on energy consumption levels. The consumption levels of households with a feedback system were compared with those of a reference group.
2. Questionnaires to gain quantitative insight in the effects of the feedback system on energy awareness, behavior change, and evaluation of the feedback, as well as demographic characteristics. To be able to determine effects, the questionnaires were executed with households with a feedback system, as well as with households in a reference group. To evaluate differences between the app user groups, the questionnaire included questions about app use and app characteristics.
3. Interviews with app users to gain insight into how people used the apps. The interviews were used to complement the quantitative insights from the questionnaires.

In the following sections, we describe the research procedure and recruitment process for each study.

Procedure

Measurement of electricity and gas consumption

For a period of 16 months, the electricity and gas consumption per day were measured for households that consented to have their meters read by the distribution network operator. The energy consumption data were quantitatively analyzed, comparing the differences between the application user group and the reference group. Before the analysis, the dataset was cleaned by removing double entries and correcting erroneous data points due to errors in registration of the data. Additionally, cases with extreme values (M ± 2SD) were removed from the sample because these may relate to incorrect meter readings, and we did not want these extreme values to influence the overall results. This

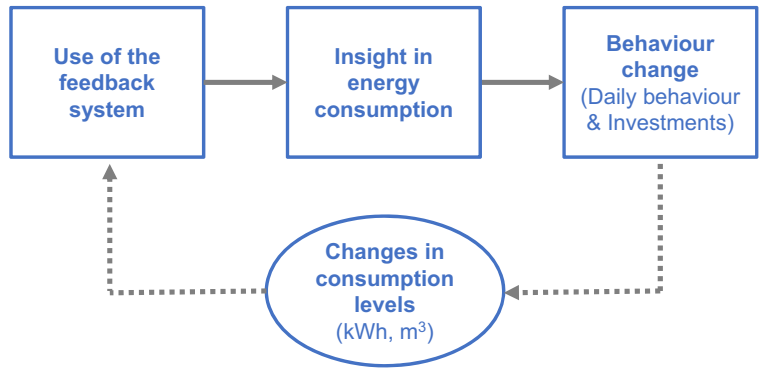

Fig. 3 Conceptual model

meant that all cases with values above 20-kWh daily electricity consumption were removed from the set of electricity consumption data (4% of the total cases was removed) and all cases with values above 8 m^3 daily gas consumption were removed from the set of gas consumption data (3% for application user group and reference group combined). No cases were removed due to extreme low values (M-2SD) because this corresponded to negative consumption (− 1 kWh and − 4 m^3 respectively) and these were not present in the sample. Additionally, the yearly consumption was collected for each household for the year before the intervention (2013) and the years the intervention ended (2014 and 2015).

Questionnaires about application use, insight in consumption, and energy-saving behavior

Questionnaires were developed for telephone questionnaires of approximately 10 min. We opted for telephone questionnaires because of the higher response rates for this method compared to online questionnaires. The questionnaires were held up to 5 months (T1) and up to 11 months (T2) after installation of the smart meter. They were conducted over a period of 3 to 4 weeks. The questionnaire for T1 focused on the short-term effects of the feedback system and the questionnaire for T2 on the effects in the longer term. The second questionnaire also addressed the default bi-monthly overview. For the second questionnaire, only the people who had responded to the questionnaire for T1 were contacted.

The main topics for respondents in both the application user group and the reference group were as follows:

- Perceived insight in one's household energy consumption with the feedback system
- Extent to which the used feedback system helps to save energy
- Effects on energy-related behavior, i.e., changes in daily activities and implementation of energy efficiency measures

In case of the reference group, the feedback system refers to the smart meter and the bi-monthly overview.

The application users were additionally asked about:

- Which application was (most) used
- How often they used the application

In the Appendix Table 5, a list of the questions and answer options is provided. Five-point scales were used as much as possible for the scale variables. Other questions were categorical or included an open answer option. With the survey results, we aimed to compare the groups' perceived insight in energy consumption and behavior changes.

Interviews

To gain deeper insight into how people used the apps, semi-structured face-to-face interviews with a duration of about 1 h were held. These interviews were used to complement and explain the findings from the energy data measurements and questionnaires. The interviews topics included the following: how people used the app in their daily lives, how the app helped households gain insight into their energy consumption, how people dealt with energy-saving measures for their home, and the extent to which the other household members were involved in the use of the app and energy saving.

The interviews were held in two rounds, within 6 weeks after the first and second questionnaire respectively. The interviews were fully transcribed, coded, and analyzed to explain results from the questionnaires.

Because of the qualitative nature of this study, a small sample size could suffice, as long as we reached data saturation. When there are repeatedly only one or two new insights (or codes) for new interviews, one has reached "data saturation" (Guest et al. 2006). Guest and colleagues (Guest et al. 2006) concluded that, with a group of relatively heterogeneous individuals, 12 interviews should suffice to reach data saturation. Hagaman and Wutich (2017) suggest that 12 to 16 interviews are sufficient for a focused topic and heterogeneous group.

Participant recruitment

For the recruitment of research participants, we adhered to the standard process applied by the distribution network operator so that the research conditions would be as close as possible to the default process of smart meter installation and evaluation. We therefore first made a selection of addresses for installation of a smart meter according to the standard process, namely about 5000 addresses. We only included single-family houses in small cities and villages in the Netherlands, because the majority of Dutch dwellings are single-family

houses and these are thus an important target group for energy-saving measures. A portion of these addresses was designated for recruitment to the application user group (approx. 3500) and a portion for the reference group (approx. 1500). The selection for the application user group was larger because we had to take into account a response rate for applying to the P1-reader as well as for participation in the research. For the selection of the reference group, we only had to consider the response rate for participation in the research.

All households were sent a letter announcing the installation of the smart meter in their home according to the standard procedure of the distribution network operator. After installation of the smart meter, they were sent a smart meter information package by mail. For households assigned to the application user group, this package also included a flyer with the offer to apply for the feedback system, and they were sent follow-up letters reminding them to apply for the system. The feedback system had to be requested via an online registration form. After registering, P1-readers were sent to the homes of the applicants, who could install them themselves with guidance from an installation manual. The applicants could also download and install the apps themselves and connect the apps to the P1-reader as soon as the latter was connected to the smart meter. The households were free to choose an app.

The installation of smart meters took place over a period of several months according to the availability of installers and householders. In the end, 1428 households applied for the feedback system and could thus be approached for energy consumption measurement, questionnaires, and interviews as part of the application user group. For the reference group, we could reach out to approximately 1500 households, provided that the installation of the meter had been successful.

Recruitment for energy consumption measurement

With respect to energy consumption measurement, households had to consent explicitly to daily meter reading by the distribution network operator. To obtain this permission, the households with a feedback system were sent an e-mail. For the reference group, a random selection of households was called by phone with the request to sign up via an online form. The response was 18% for the households with feedback system and 17% for the reference group (see also Table 2). The gas and electricity consumption of 519 households could be measured for a period of 16 months in 2014 and 2015 (application user group $n = 264$; reference group $n = 255$).

Recruitment for the questionnaires

For the questionnaires at T1, addresses were selected randomly from the application user group and the reference group. To avoid a bias in our sample, we did not prioritize approaching the households who participated in the meter readings study for participation in the questionnaire study. At T2, only households that had participated at T1 were approached, and in the application user group only those who were using an application at T1. This resulted in complete questionnaires at T2 of $n = 119$ for the application user group and $n = 151$ for the reference group. There was a small number of households for which both questionnaire results and energy consumption measurements were available at T2, namely $n = 42$ for the application user group and $n = 55$ for the reference group.

Recruitment for the interviews

Participants for the interviews were selected from the questionnaire respondents who had a feedback system and indicated willingness to participate in an interview. A total of 12 interviews with application users were included in the analysis. Six of the interviews were conducted at T1, and six at T2. Because the interview setup was the same at T1 and T2, and no comparisons were made across time, we treated the set of interviews as one. After the 9th interview, we noted only incidental new insights (codes) in our analysis (see Fig. 4), which assured us that we had reached saturation. Thus, the sample of 12 interviews provided sufficient insights to understand the application users' experiences.

Results

In this section, we describe the composition of the sample to control for differences in demographic variables between the application user group and the reference group. Then we present the results from the analysis of consumption levels, followed by the results from the questionnaires concerning the use of the applications, effects on the respondent's insight in energy consumption, and behavior changes in the household.

Table 2 Number of participants per study. Response rate between parentheses

Study group	Meter read-out	Questionnaires		Interviews
		T1	T2	
Application users	264 (18%)	303 (34%)	119 (42%)	12
Reference group	255 (17%)	454 (17%)	151 (74%)	–
Total	519	757	270	12

Lastly, the results from the interviews are described, which provide insights that help explain the effects found in the analysis of the consumption data and questionnaires. The number of respondents per study is summarized in Table 2.

Demographics

The demographic differences between the application user group and the reference group were assessed with independent t tests and chi-square tests, based on the responses in the questionnaires. Significant differences were found only for gender. The reference group consisted of fewer male respondents than did the application user group (70% vs. 90%; $\chi^2(1,\ n = 224) = 12{,}564$, $p < 0.001$). It is noteworthy that there are more male than female respondents. Given that we requested the person in the household that had been most involved with the smart meter and feedback system to respond to the questionnaire, it appears that in multiple person households, the men were most often using the system (assuming that most of the households with two or more household members include male-female couples). See Table 3 for the averages of the sample. Because there are no significant differences on the other variables and given that the majority of respondents in both groups is male, we assume for this study that the application user group and the reference group are comparable with respect to the variables related to the use of the feedback, insight in energy consumption, and behavior change.[2]

Effects on energy consumption

To find out whether energy saving took place in the application user group, the consumption levels of the households were normalized based on the consumption the year before the intervention (after/before*100). This normalization allowed for a comparison of the situation before and after the intervention an independent t test with the consumption levels as dependent variable and the group as independent variable was performed. No significant differences between the two groups were found, for electricity ($t(476) = 0.1228$; $p > 0.05$) and gas consumption ($t(511) = -0.353$; $p > 0.05$). Thus, no significant difference in energy consumption was found between the application user group and the reference group.[3]

Figures 5 and 6 show the gas and electricity consumption for a period of 16 months during the study, aggregated to an average per month of the daily consumption. Note that the electricity consumption for the application user group is structurally higher than for the reference group, not only during the study but also in the years preceding the study. The difference over the total consumption during the study period is 7.7% and significant (independent t test; $t(461) = 2.053$; $p < 0.05$). The graphs in Figs. 5 and 6 illustrate that no energy saving took place in the application user group, because if energy saving had occurred in that group the consumption levels of the application users would have decreased over time in comparison to the reference group's levels.

[2] A comparison between users of App A vs. users of App B did not show significant differences.

[3] This analysis was also conducted for the electricity and gas consumption of the respondents to the questionnaire. No significant differences between the application user group and the reference group for electricity ($t(80) = -1.155$; $p > 0.05$) and gas consumption ($t(76) = -0.564$; $p > 0.05$).

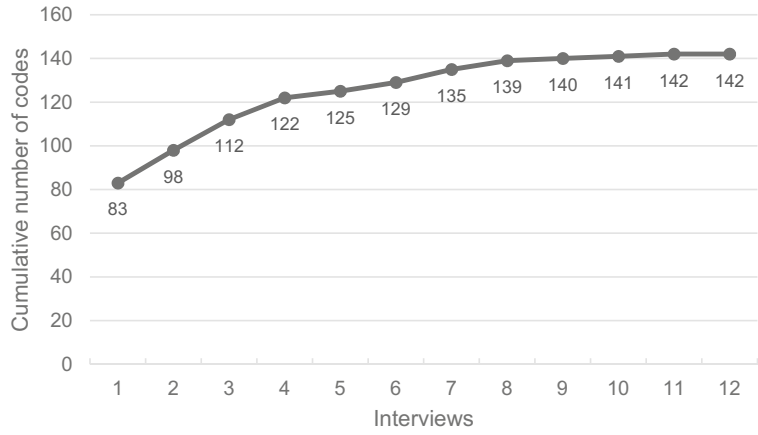

Fig. 4 The cumulative number of codes per interview

Results from questionnaires

Analysis of the questionnaire responses provided insight into the differences between the application user group and the reference group, between the users of app A and app B, as well as changes over the 6-month period between T1 and T2. We describe first the results concerning the use of the feedback system, second the effects on the respondent's insight in energy consumption, and third the effects of the feedback system on the energy-related behavior of the households. The mean scores for the variables discussed below are presented in Table 4. The distribution of the responses for the variables is included in the Appendix Table 5.

Use of the feedback system

As described above, the households had to apply for the feedback system and received it for free.

Table 3 Overview of demographics

Variable	Value
Mean age of respondent	55
% male	79%
Mean amount of household members	2.65 persons
Mean floor surface	141.7 m^2
Mean construction year of the house	1976
Mean gas usage (2013)	1737 m^3
Mean electricity usage (2013)	3843 kWh

The results of the questionnaires showed that the main reasons for applying were interest in gaining insight in one's energy consumption or in saving energy (64%). Most households who used the feedback system with one of the applications at T1 still used it at T2 (70%). Twenty-nine percent of the people had used more than one application (56 out of 196 respondents at T1). For these respondents, we asked which application they used most. For questions related to application use, we focused on the most used application, because it can be expected that this application has most influence on the application user's behavior. The majority of the application users mainly used only one of the smartphone apps, namely 55% used App A and 29% used App B at T1. The choice for the most used app was based on it being the first app people installed; this was either because it was the only one that—they thought—was available or for its functionality (respectively 23%, 26%, and 18% of the responses at T1).

At T2, 64 of 97 respondents (66%) were still using the initial app, while 4 had switched to another app and 29 had stopped using an application. The most mentioned reason that people had stopped using the feedback system between T1 and T2 was that they did not perceive added value (7 out of 19 answers). Other reasons mentioned more than once included that the participant did not make the effort to reconnect the P1-reader to the internet, to reinstall apps, or look up passwords (e.g., after changing phones) (3 out of 19), that he or she did not have time (2 out of 19), and that the system no longer functioned (2 out of 19).

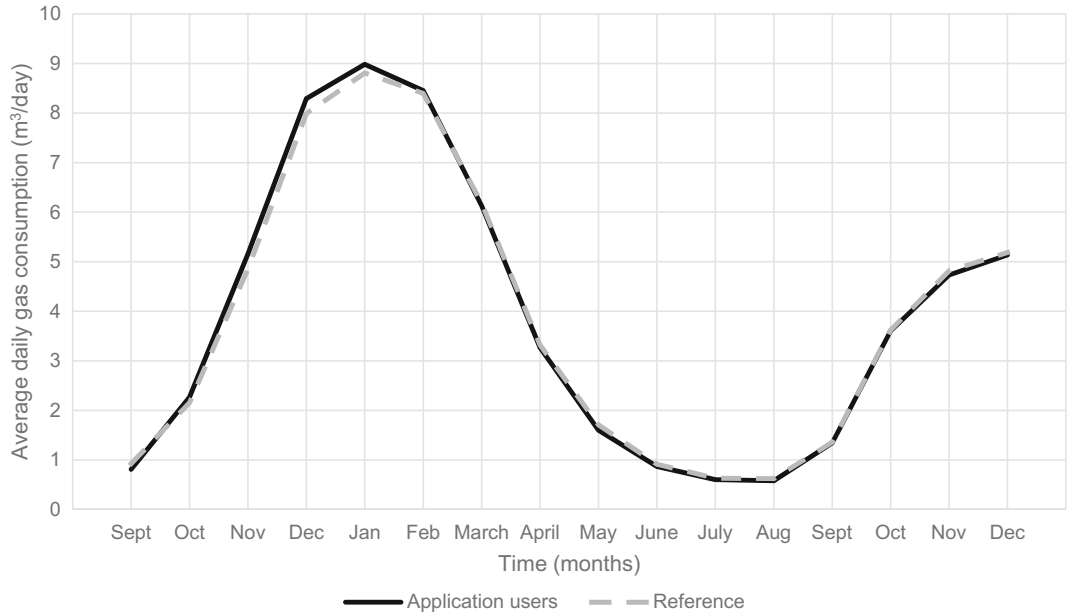

Fig. 5 Average daily gas consumption per month

At T1, the applications were used several times a week on average. The use frequency dropped significantly the longer the app had been in use. Between T1 and T2, the use frequency went from an average score of "several times a week" to an average score of "once a week" ($F(1,67) = 49{,}325$; $p < 0.001$). A Mann-Whitney U test revealed no significant difference in the use frequency of App A and App B.

Effects of feedback system and time on insight in energy consumption

A repeated measures ANOVA was performed with the expected insight as dependent variable, the groups (application user group vs. reference group) as between-subjects factor, and time (T1–T2) as within-subject factor. Application users

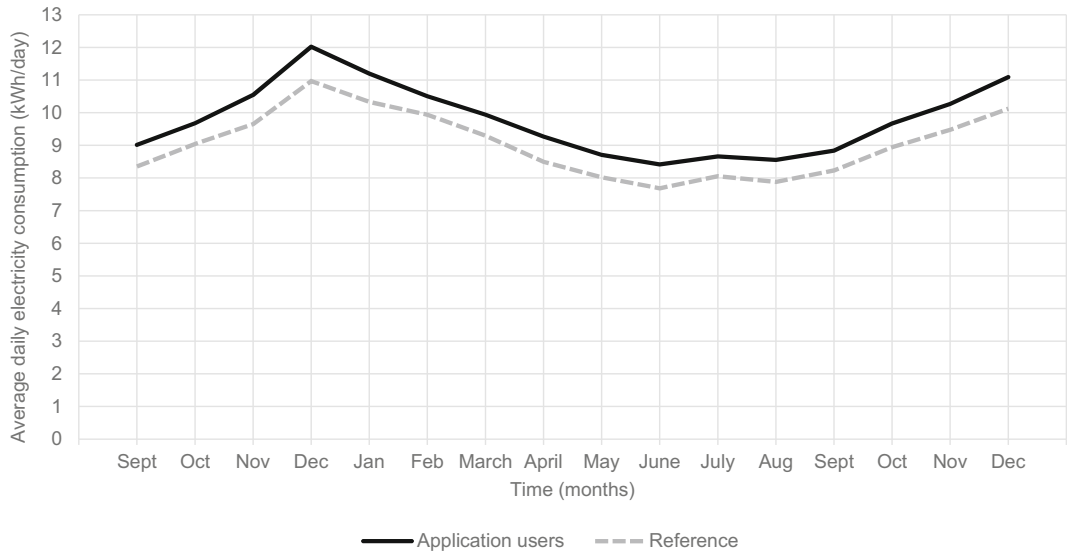

Fig. 6 Average daily electricity consumption per month

Table 4 Mean scores of the questionnaire results for the different groups. See Appendix Table 5 for more details about the possible scores for each topic. The superscripts in the 4th and 5th columns (a, b, and c) indicate significant differences between test results, e.g., the superscript a for the results on insight in energy consumption indicates a significant difference between the application user group and the reference group

Topic	Group	*n*	T1 M (SD)	T2 M (SD)
Use frequency	Application user group	68	6.19 (1,22)[a]	4.79 (1,70)[a]
	Reference group	N.A.		
	Total	N.A.		
Insight in energy consumption	Application user group	94	2.52 (0,65)[ac]	2.32 (0,75)[bc]
	Reference group	103	2.01 (0,77)[ac]	1.96 (0,85)[bc]
	Total	197	2.25 (0,76)[c]	2.13 (0,82)[c]
Help with energy-saving application	Application user group	68	3.34 (0,96)[a]	2.76 (0,88)[a]
	Reference group	N.A.		
	Total	N.A.		
Help with energy-saving bi-monthly overview	Application user group	60	N.A.	2.58 (0,85)[a]
	Reference group	89	N.A.	2.45 (1,14)[a]
	Total	149	N.A.	2.50 (1,03)
Behavior change–daily behavior	Application user group	97	1.90 (1,03)[ac]	2.29 (1,04)[bc]
	Reference group	127	1.54 (0,92)[ac]	1.86 (1,09)[bc]
	Total	224	1.70 (0,98)[c]	2.04 (1,09)[c]
Behavior change –changes to home	Application user group	97	No: 86 (89%)	No: 66 (68%)
			Yes: 11 (11%)[a]	Yes: 33 (32%)[ab]
	Reference group	127	No: 117 (92%)	No: 112 (88%)
			Yes: 10 (8%)	Yes: 15 (12%)[b]
	Total	224	No: 203 (91%)	No: 178 (79%)
			Yes: 21 (9%)	Yes: 46 (21%)
Behavior change–intention to take measures	Application user group	97	2.52 (1,05)[ac]	2.09 (1,03)[bc]
	Reference group	127	2.02 (1,07)[ac]	1.65 (0,87)[bc]
	Total	224	2.24 (1,09)[c]	1.84 (0,97)[c]

reported expecting more insight into their households' energy consumption levels than did the reference group ($F(1,195) = 20{,}269$; $p < 0.01$). For both groups, the expected insight diminished between T1 and T2 ($F(1,195) = 6278$; $p < 0.05$).

The application users also reported the extent to which the application helped them to save energy. There was a significant difference between T1 and T2 (Wilcoxon signed-rank test, $p < 0.05$). At T2, the scores were lower than at T1. Application users appreciated the application more than the bi-monthly overview in helping them to save energy (Wilcoxon signed-rank test, $p < 0.05$). A Mann-Whitney *U* test did not show a significant difference between App A and App B in terms of their tendencies to help save energy.

Effects of feedback system and time on the household's behavior

At T1 and T2, the questionnaire respondents were asked about the extent to which they had adjusted their daily behavior. A repeated measures ANOVA was performed with behavior change as dependent variable, the groups (application user group vs. reference group) as between-subjects factor, and time (T1–T2) as within-subject factor. The application users reported more behavior change than did the reference group at both T1 and T2 ($F(1,222) = 11{,}704$; $p < 0.001$). Furthermore, both groups indicated an increase in behavior change between the first and the second questionnaire ($F(1,222) = 21{,}584$; $p < 0.001$).

A repeated measures ANOVA with behavior change as the dependent variable, time as the within-subjects factor, and application (App A and App B) as the between-subjects factor indicated a main effect of time ($F(1,79) = 11,077$; $p < 0.01$), but did not reveal a significant effect of the apps. Thus, for both apps, more behavior change was reported at T2, but we did not find support for the possibility that one app performed better than the other.

To investigate whether the application use frequency was positively related to changes in daily behavior, we performed a correlation analysis. The extent of daily behavior change correlated significantly with use frequency (at T1: Spearman's rho = 0.279, $p < 0.01$; at T2: Spearman's rho = 0.473, $p < 0.01$). The same applies for the intention to make changes to the house (at T1: Spearman's rho = 0.320, $p < 0.01$; at T2: Spearman's rho = 0.365, $p < 0.01$). Thus, the more frequently people used the app, the more likely they were to report behavior change.

The respondents were also asked whether they had actually replaced appliances or made adjustments to their home (answering with yes or no). At T1, 9% of all respondents reported to have taken measures. A chi-square test did not show a significant difference between the application users and the reference group at T1. For the response at T2, a chi-square test showed that the application user group had taken significantly more measures than the reference group, namely 32% vs. 12% ($\chi^2(1, n = 224) = 13,680$ $p < 0.01$). Furthermore, at T2, more application users reported to have taken measures than at T1 (McNemar's test, $p < 0.05$).

When comparing the users of App A with those of App B, chi-square tests did not show significant effects between the two apps at T1 or at T2. More users of App A reported energy-saving measures at T2 than at T1 (McNemar's test, $p < 0.05$). For users of App B, this effect was marginally significant ($p = 0.065$).

In an open question, the respondents stated what adjustments they had made. The most mentioned measures at T2 were as follows: changing the lights, changing insulation, and buying energy efficient appliances, at respectively 47%, 14%, and 12%.

We also asked about the extent to which respondents intended to make adjustments to their home in the future. A repeated measures ANOVA with intention as dependent variable, the groups (application users and reference group) as between-subjects factor, and time (T1–T2) as within-subject factor was performed. There were main effects for group and time. The intention to make adjustments was higher in the application users group than in the reference group ($F(1,222) = 17,792$; $p < 0.001$) and the intention dropped in both groups between T1 and T2 ($F(1,222) = 25,468$; $p < 0.001$).

To assess differences in the effects of App A vs. App B, a repeated measures ANOVA was performed, with the extent to which adjustments were intended as the dependent variable, time as the within-subjects factor, and the app (App A and App B) as the between-subjects factor. This showed a main effect of time ($F(1,79) = 13,835$; $p < 0.01$) but no significant effect between the apps. Thus, for both apps, higher intention to make adjustments was reported at T2, but the specific app used caused no effect.

The decrease in reported intention may be explained by the fact that households took measures between T1 and T2, and as a result had less intention to make more adjustments. Alternatively, it may be that over time households had clearer ideas about whether they would implement the initially intended changes. If they had become less inclined to, their intention would lessen.

When respondents indicated an intention to make adjustments, they were asked to clarify what they intended to do. Figure 7 shows the distribution of responses over the mentioned measures. Most mentioned were insulating (25%), installing solar panels (20%), and replacing lights (16%). It is remarkable that 18% of the respondents stated that they did not have a concrete idea yet.

Interview results

The analysis of the consumption data did not demonstrate energy saving in the group of application users compared to the reference group. The questionnaire results, on the other hand, indicated that the application users perceived more insight

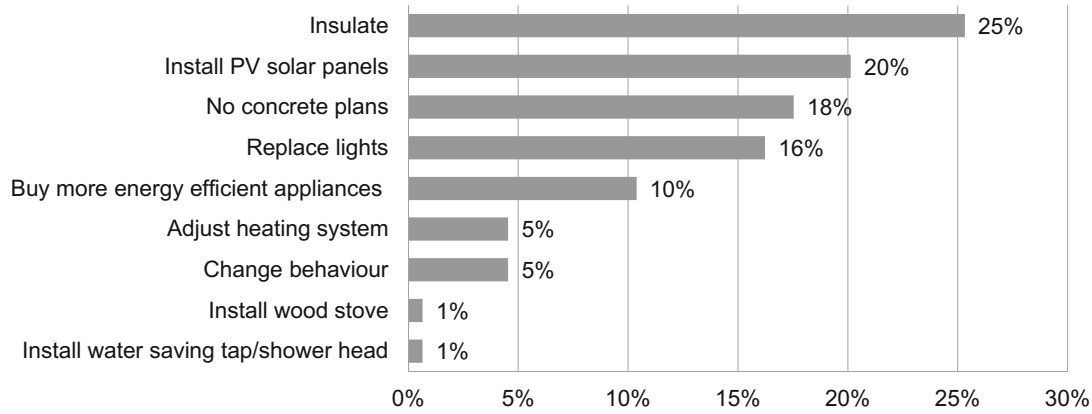

Fig. 7 Intended adjustments on T2 as a percentage of the total number of responses (154) for a particular measure ($n = 136$)

into their consumption and reported more energy-saving behavior. The results from the interviews provided more insight into this apparent contradiction by shedding light on how the application users utilized the feedback system. The results presented here are based on statements by multiple persons.

Use of the feedback system

The interview results indicate that the application users used the applications initially to gain insight into their consumption and then to monitor their consumption levels over time. The decrease in use frequency between T1 and T2 can be explained by this transition. In the beginning, the users would use the app more frequently to learn how their energy patterns were built up. Some of the respondents actively checked the consumption of specific appliances by switching them on and off or by trying to explain the total consumption ("What appliances are currently on?"). After a while, when they had sufficient insight in their consumption patterns ("know how it works"), they mainly used the app for monitoring purposes, which was done on a less frequent basis. When deviations from the usual known pattern were seen, a user would look for an explanation and, if possible and desired, take measures. Different respondents discovered in this way that, for example, the oven was not functioning well, that the door of a built-in refrigerator did not close anymore, or that the heating element of the dishwasher was not functioning.

With respect to the main functions of the apps, interviewees explained the different uses of the current consumption screen and the historical consumption screen. The current consumption information was used to understand where and when energy was consumed, how much energy-specific appliances used, and to find out where energy was wasted. The historic consumption was used to monitor consumption over time, to explain changes, and to check for deviations from the usual pattern. In general, historic consumption per day/week/month was perceived to be more relevant than the current consumption. We expect that the comparison over time made the information more meaningful to the application users, because they could interpret the differences between high and low consumption during a specific day, for example by knowing that the washing machine was used or that the heating was off because no one was at home. The historic information was also used to monitor and control the time of day that appliances are used. Someone who had consciously taken energy efficiency measures would look at the historic consumption graphs to see if the measures were actually resulting in lower consumption levels compared to before. In this respect, some of the respondents stated that they were waiting to complete a year of historical consumption data so that they could compare particular months with those of the year before.

The development of habits around the use of the application seems to be related to the use frequency. One user checked the daily

consumption pattern before going to bed and would regularly discuss with his wife how to decrease their consumption. Others would consult the app as part of their regular "playing" with their smartphone and checking of social media apps. Others would open the app "just sometimes," or consult the app at a moment when they "just wondered" how much the current energy consumption was.

In each case, the interviewee was the only person in the household who had installed and used the application. This person, who could be called the "energy manager," was often the person in the household who stimulated and executed the energy-saving activities in the household and who dealt with the contract with an energy provider. Most respondents explained that they discussed insights from the app with their partner (10 of 11 who lived with a partner). The involvement of household members ranged from being more or less actively involved in maintaining or finding new energy-saving behavior, to leaving energy management to the "manager," and following his/her suggestions. During the interviews, we met one household where both partners were active as "energy managers." One of them used the app and reviewed bi-monthly overviews, while the other dealt with the energy provider—"she knows the password"—and had noted down the meter readings on a weekly basis before they had a smart meter.

In the questionnaires, we found that application users perceived improved insight in their energy consumption. Given the additional insights from the interviews, we can say that this improved insight would only have translated to an increased awareness in the whole household if the energy manager shared insights from the app with household members and they were also (becoming) involved in energy-saving activities.

Insight in energy consumption and help for energy saving

From the interviews, we learned that the applications can lead to insight into household energy consumption expressed on several levels: (1) Insight into the amount of energy the household consumes and the pattern of consumption, (2) Insight into the consumption of specific appliances and how those figure into overall consumption, (3) Insight into how one's consumption compares to other households or to what is "normal," (4) Insight into possibilities for energy saving, namely by knowing which appliances, technologies, or behaviors contribute to the desired energy consumption levels. While the first two levels were mentioned by most respondents, the third and fourth were mentioned by only a few of the interviewees.

There were various opinions on the extent to which the apps actually helped to save energy. Some people were content with the provided insight and stated that it helped them to save energy, while others required more concrete, detailed information. These latter were not sufficiently motivated by the available information because it did not provide them with the tools to take action. This may also explain why 18% of the respondents with intentions to take energy-saving measures were not able to mention concrete plans (see Fig. 7). One respondent, for example, said the household had already saved a lot (insulation, efficient appliances, energy-saving routines) and now needed more detailed information in order to optimize appliance use. A common remark when asked about the use of the feedback was that they wondered how to utilize the information the app provided. They wanted more actionable insights that are applicable to their own situations.

Changes to the household's energy-related behavior

Some interviewees stated that they had become more aware of their consumption and thus were looking for—and taking advantage of—energy-saving opportunities for their homes and daily routines. For example, one respondent had structurally lowered the set temperature of the central heating system since he had installed the feedback system. Others indicated that they had replaced light bulbs or purchased more efficient appliances like dishwashers, washing machines, and refrigerators. Several respondents suggested that the feedback itself did not cause the energy-saving action, but that it helped to inform or incentivize decision making. Intentions thus appear to be reinforced by the feedback system.

The interviewees also shed light on factors that may have limited the effects of the feedback on energy saving. One important reason for little behavior change that the interviewees mentioned was that they already had energy-saving habits and had already made several investments to increase (or maintain) the household's energy efficiency. Another reason was that respondents did not want to lessen their general comfort, even though they expected that there were possibilities to save more energy.

In cases in which the feedback information was used to evaluate the current situation and the household was satisfied with the consumption levels, e.g., because it was below the average for similar households or within one's budget—no action was taken. Additionally, energy efficiency investments or changes in behavior were not likely when the household did not have concrete ideas about what action(s) to take. In order to actually take action, the household must take the extra step of making the effort to find out *how* to save more energy. A user's motivation to save energy should be high enough to take this extra step.

The questionnaires indicated that more energy-saving measures were taken in the long term at T2. This may be explained from the interview results, which suggest that households first must take action to understand what they can and want to change, and then make changes at moments that suit them. Two respondents stated that they had asked for advice on the best option for their specific situation. One of them explained that he was getting advice and quotations for a new air conditioner, but had not decided yet. Furthermore, respondents would often wait for the right time to make adjustments, for example because an appliance had not yet reached its end of life or because the household could not yet afford a certain investment. This was mentioned in connection to changing light bulbs and appliances such as refrigerators and washing machines, and insulation.

Discussion

As noted in the introduction of this article, a premise for the promotion of energy feedback systems is that people become more aware of their energy consumption patterns and undertake activities to save energy, such as changing their behavior or implementing energy efficient technology. From other studies, we know that the way in which the feedback is provided—the design of the feedback system—plays a role in how people engage with that system and whether or not energy savings are achieved (e.g., Kobus et al. 2015b; Buchanan et al. 2015). From the questionnaires, we found that the users of the smartphone/tablet applications reported higher awareness and more energy-saving activities than the reference group. However, we did not find an actual effect in terms of measured gas and electricity consumption levels.

How can these differences in results be explained? First, the results are based on different samples that may differ in their energy behaviors. Yet considering that no energy savings were revealed by an additional analysis on the more restricted sample of the households who participated in both the energy measurement and questionnaire studies, this explanation does not seem likely. Another more likely explanation would be that participants may have given socially desirable answers in the questionnaire study, meaning that participants provided an over-optimistic picture of their energy behavior. Finally, the interview results indicate that the applications tended to be used to monitor consumption levels and thereby gain more awareness, rather than to achieve lower consumption levels.

Given that the designs of the apps in this study are similar to many of the apps available in the Netherlands, it will be important in the future to look into ways to improve their designs, so that they not only facilitate energy monitoring but also encourage energy saving. The insights from the interviews into how people used and evaluated the feedback system provide some explanations and allow for the formulation of recommendations for the design of an app-based feedback system.

The study results suggest that the applications were used to gain insight in one's energy consumption levels and to monitor consumption over time (i.e., to check whether consumption levels are as expected), rather than as a tool to save energy. The applications did not seem to offer sufficiently

concrete and actionable information to assist in energy saving and, in the case of low motivation, to motivate undertaking energy-saving activities. These findings are in line with a study by Nilsson et al. (2014) in which insufficient understanding of the provided information and lack of interest in energy saving were shown to be important barriers to achieving energy savings. The application should thus provide information that is relevant, meaningful, and actionable to users who do not have a strong interest in understanding their consumption.

Hargreaves et al. (2013) described how a householder may think that it is not possible to save energy while maintaining a comfortable lifestyle because he or she is not aware of possibilities or methods for achieving energy savings without decreasing one's comfort. We also found this in the interviews in this research. We also found that no energy-saving actions were taken in households where the feedback information was used to evaluate the current situations and the householders thought that the consumption levels were acceptable, for example, because a consumption level was lower than that of similar households, or was within the household's budget. These findings lead us to recommend that an application designed to achieve change in a household's energy-related behavior must guide users towards becoming aware of how and where to save energy in ways that match the household's needs and abilities.

The finding that the apps were used for monitoring purposes after an initial learning phase was also observed in the evaluation of the smart meter roll out in the UK. The UK researchers observed that most households that received an in-home display after the installation of the smart meter were still using it 2 years later to monitor their consumption. There were also energy-savings effects of about 2–3% (Darby et al. 2015). In our research, we found that most of the application users were still using their app after more than 6 months (74% in the questionnaire at T2). So perhaps the fact that people can continue to use an app to monitor their consumption may, in time, pay off as people gradually change their behavior and implement energy-saving measures. To gain more insight into this potential, we suggest looking into the effects of apps that have been available in the market (paid or for free) for several years.

The feedback system was easily accessible with the smartphone/tablet application, but this did not facilitate frequent use of the app in the long term. We found a significant drop in use frequency between T1 and T2. Given that the application users were self-selected to the research by applying for the feedback system, we would expect a basic interest in energy consumption and therefore recurring/frequent use of the app. If a high-use frequency could not be sustained within this sample during the research, households that are not interested in applying for a free feedback system may be even less likely to use an app regularly. As we uncovered in the interviews, one reason for low-use frequency can be the limited relevance of the information to the user and declining perceived relevance of the information after an initial learning period. Additionally, an app can easily be "out of sight, out of mind." People have to deliberately look up and open the app on their smartphone/tablet. When they are not committed to do so, or triggered by, e.g., an icon on their home screen or notifications, this provides a barrier to seeing and using the feedback.

A higher use frequency was found to be related to more energy saving by Kobus et al. (2015b) and is suggested as well by the correlation found in this study with respect to reported use frequency and behavior change. Therefore, we recommend exploring which design elements of an app draw more attention from users, for example, via notifications with tips or reminders related to one's actual consumption pattern, or ambient cues indicating via colors, or icons showing when energy consumption is higher than expected from one's normal consumption. Sustaining a user's interest in the feedback information can be a challenge for energy saving via a feedback system (Buchanan et al. 2015; Darby et al. 2015; Kobus et al. 2015b). App designers should thus also consider how an app can continue to be relevant as the household's consumption patterns and interests change over time. Given that apps are device-specific—unlike in-home displays—and easily updated, they offer a valuable opportunity in the sense that they can grow along with their users.

Another factor that may have influenced the effectiveness of the feedback system was that the apps were only used by the "energy manager" of the household. The information about a household's energy consumption was therefore only handled by one person. This person would have to initiate conversations about the household's energy use and to get cooperation from family members for investing in energy efficiency measures or to make changes in daily routines. This barrier was also observed by Van Dam (2013) and Hargreaves et al. (2010). The disadvantage of an app, compared to an in-home display, is that it is not a shared object in the household. We have two suggestions for this issue: first, to make it easy for the energy manager to share insights and ideas with his household members, and second, to make the use of the app more attractive and relevant for all household members. This also means that the feedback system has to respond to the dynamics of daily practices in homes. In the words of Strengers, the design has to move beyond the interest of the rational and individual "Resource Man" (Strengers 2014).

We found that the implementation of energy-saving measures takes time. People have to find out what energy-saving measures they want to take, and this often involves waiting for the right moment. A necessary replacement of appliances and the making of home improvement plans are often moments when energy efficiency considerations are more easily taken into account and executed (Verplanken and Wood 2006; Stieß and Dunkelberg 2013). For the effects of the intervention in this study, this means that the feedback may contribute to energy saving that occurs after the intervention. For the design of a feedback system, we recommend that it supports the decision-making process for the implementation of energy-saving measures by complementing energy consumption information with more contextual information and advice, e.g., about how to best get advice tailored to your home from local experts, about cost-benefit trade-offs, and about available subsidies.

Finally, it is important to reflect on what the energy-saving potential actually was during the study. The energy-saving actions of the households may not have had a noticeable impact on the overall consumption levels. Several households suggested that their house and behavior were already quite energy efficient, which leads to the question of how much more energy could have been saved with the measures that were primarily mentioned: changing light bulbs and purchasing efficient appliances. For interventions like the one in this study, it is recommendable to estimate the potential savings beforehand. Knowing this, you could also provide your users with more relevant and actionable information tailored to their own situations.

Limitations of the study

We structured our research to obtain a representative sample of Dutch households. The demographics of the sample do however differ to some extent from those of the general Dutch population and those of other countries. In comparison to the Dutch population[4], the sample for this research is older and more highly educated. More of the houses are privately owned than rented. The homes are relatively new (later built), and they also have larger floor surfaces, more occupants, and higher energy consumption on average. Complementary research with samples of different compositions could provide insight herein.

This research was executed in a natural setting within the regular process of smart meter installations. The recruitment of participants for questionnaires and meter readings had to be done separately from the meter installation and the participants' application for the feedback system, in order to avoid influencing the network operator's usual way of approaching households for smart meter installations. As a result, a self-selection bias was introduced for the application user group and we were not able to work with one large sample for which both meter readings and questionnaire results were available. Our research approach did however provide a unique opportunity to study the effects of a feedback system in a normal situation. By setting up the

[4] Based on data from the Central Bureau of Statistics of The Netherlands. Accessed online via statline.cbs.nl in 30 March 2016.

three studies, we were able to gain complementary insights into the effects of the apps on household energy consumption, both quantitatively and qualitatively.

In this research, we chose to let one respondent represent the household, yet one respondent cannot fully represent a multi-person household. This study collected and examined the viewpoints of household energy managers. For a more complete picture of the role of a feedback system, future research would have to address the complex dynamics of households and the (energy-consuming) products and services they use. This is particularly relevant for finding out how products and services can facilitate energy efficiency within a household's daily practices and as related to its wider social context, as described, e.g., by Shove (2010), Gram-Hanssen (2010), and Schwartz et al. (2015).

Conclusion

This research contributes to the existing literature about feedback systems with insights into the use and effectiveness of a smartphone/tablet app. The application users in the sample with measurement of energy consumption levels did not show a decrease in their energy consumption during the research period, compared to the reference group. Yet in the sample who responded to questionnaires, application users reported increased awareness and energy-saving activities compared to the reference group. Further insight from interviews with application users indicated that people used the apps mainly to learn how their energy consumption levels are built up and to monitor the consumption levels over time, rather than to decrease one's consumption levels. In line with other research into feedback, the interview results suggest that an app could be more effective with information that is more actionable and meaningful with respect one's own specific situation. Furthermore, more effectiveness can be expected when a higher use frequency is stimulated and insights are provided that relate to the goals the end-users want to achieve in their household.

Based on this research, we cannot yet conclude whether apps do or do not lead to energy saving. Further exploration is recommended with respect to how the design of such apps can encourage a wide audience to monitor their consumption and guide them in taking action to change their consumption levels. In light of the implementation of smart meters and feedback systems based on meter data, policy makers should be aware that feedback systems do not necessarily lead to energy savings because their effectiveness depends on their designs and the contexts in which they are implemented. Feedback systems play a valuable part by making energy consumption visible, but a comprehensive approach with complementary products, services, and/or policies is important in facilitating a household's process of learning and decision making about energy efficiency. Policy makers should take this into account when defining goals, approaches, and guidelines for stimulating energy saving.

In view of the current energy transition, with increasingly decentralized production of renewable energy and the growing uptake of electric transport and heat pumps, we suggest broadening the discourse on smart meter implementation for households. The insight in energy consumption and production that smart meters can provide should go beyond encouraging households to solely save energy, enabling them to adjust their electricity consumption patterns to the demand and supply of (local and renewable) energy.

Data statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to them containing information that could compromise research participant privacy.

Conflict of interest The authors declare the following interests:

- Daphne Geelen and Annemieke Bulters are employed by Enexis B.V., a distribution network operator implementing smart meters according to national policy in the Netherlands.
- Delft University of Technology received financial support from Enexis B.V. for Ruth Mugge and Sacha Silvester to be able to contribute to the research reported in the submitted work.

Appendix

Table 5 Measures from questionnaire

Topic	Questions [A]: for application user group [R]: for reference group	T1 or T2	Answer options	Responses Total	Reference group	Application users	App A	App B
Requesting P1-reader	Why did you apply for the P1-reader? [A]	T1	1. The P1-reader is for free.	–	–	10%	–	–
			2. Curious about my energy consumption / would like to save energy / want to have more insight.			64%		
			3. Nice to try a new device/gadget.			12%		
			4. Other, namely …			14%		
			More than one answer possible			*% of answers N = 129*		
Use of application	Which of the applications do you use most? [A]	T1	1. App A	–	–	55%	–	–
			2. App B			29%		
			3. Desktop application			12%		
			4. Online application			4%		
						N = 97		
		T2	1. App A	–	–	54%	–	–
			2. App B			34%		
			3. Desktop application			12%		
			4. Online application			0%		
						N = 68		
	Why are you not using or have you stopped using the application(s)? [A]	T2	… (open answer)	–	–	–	–	–
	What is the main reason that you are using this application most? [A]	T1	1. Design	–	–	17%	12%	33%
			2. Functionality			18%	14%	27%
			3. For free			3%	2%	3%
			4. Independent provider			0%	0%	0%
			5. Known provider			3%	0%	9%
			6. First one I installed			23%	27%	3%
			7. The only one available / that I found			26%	31%	15%
			8. Other, namely …			11%	15%	9%
			More than one answer possible			*% of answers (N = 111)*	*% of answers (N = 59)*	*% of answers (N = 33)*

Table 5 (continued)

Topic	Questions [A]: for application user group [R]: for reference group	T1 or T2	Answer options	Responses				
				Total	Reference group	Application users	App A	App B
Use frequency	How often do you look at the application? [A]	T1	1. Never	–	–	1%	0%	0%
			2. Less than once a month			5%	2%	4%
			3. About once a month			4%	4%	7%
			4.About once every two weeks			31%	36%	25%
			5. About once a week			27%	30%	25%
			6. A few times a week (not every day)			14%	11%	11%
			7. Once a day			18%	17%	29%
			8. Several times a day					
						N = 97	*N = 53*	*N = 28*
		T2	1. Never	–	–	9%	3%	13%
			2. Less than once a month			22%	32%	9%
			3. About once a month			9%	11%	9%
			4.About once every two weeks			22%	11%	26%
			5. About once a week			24%	24%	30%
			6. A few times a week (not every day)			9%	14%	4%
			7. Once a day			6%	5%	9%
			8. Several times a day					
						N = 68	*N = 37*	*N = 23*
Insight in energy consumption	To what extent do you think the smart meter (application users: in combination with the P1-reader) provides you better insight in your energy consumption? [A] [R]	T1	1. Not	20%	31%	6%	6%	4%
			2. A little	41%	44%	37%	35%	43%
			3. A lot	38%	24%	54%	58%	50%
			4. Quite a lot	2%	2%	2%	2%	4%
				N = 204	*N = 110*	*N = 94*	*N = 52*	*N = 28*
		T2	1. Not	24%	32%	14%	12%	14%
			2. A little	41%	38%	46%	40%	46%
			3. A lot	32%	26%	39%	42%	39%
			4. Quite a lot	3%	3%	3%	6%	0%
				N = 213	*N = 117*	*N = 96*	*N = 52*	*N = 28*
Help with energy saving	To what extent does the application help you to save energy? [A]	T1	1. Not at all	–	–	2%	2%	0%
			2. Very little			16%	17%	14%

Table 5 (continued)

Topic	Questions [A]: for application user group [R]: for reference group	T1 or T2	Answer options	Responses				
				Total	Reference group	Application users	App A	App B
			3. A little			37%	36%	36%
			4. A lot			31%	30%	39%
			5. Quite a lot			13%	15%	11%
						N = 97	*N = 53*	*N = 28*
		T2	1. Not at all	–	–	9%	11%	5%
			2. Very little			26%	26%	29%
			3. A little			44%	37%	52%
			4. A lot			21%	26%	14%
			5. Quite a lot			0%	0%	0%
						N = 68	*N = 38*	*N = 21*
	To what extent does the bi-monthly overview help you to save energy? [A] [R]	T2	1. Not at all	21%	28%	12%	15%	6%
			2. Very little	25%	21%	30%	35%	25%
			3. A little	36%	29%	47%	35%	63%
			4. A lot	17%	20%	12%	15%	6%
			5. Quite a lot	1%	1%	0%	0%	0%
				N = 149	*N = 89*	*N = 60*	*N = 34*	*N = 16*
Behaviour change	You may have taken measures to save energy in your home after the installation of the smart meter (application users: in combination with the P1-reader).	T1	1. Not at all	59%	68%	48%	53%	36%
			2. Very little	20%	17%	23%	15%	32%
			3. A little	13%	9%	20%	23%	25%
			4. A lot	7%	6%	9%	9%	7%
			5. Quite a lot	0%	1%	0%	0%	0%
				N = 224	*N = 127*	*N = 97*	*N = 53*	*N = 28*
	To what extent have you changed your daily behaviour? [A] [R]	T2	1. Not at all	43%	54%	29%	28%	21%
			2. Very little	21%	17%	27%	30%	29%
			3. A little	24%	17%	32%	28%	36%
			4. A lot	11%	10%	11%	11%	14%
			5. Quite a lot	1%	1%	1%	2%	0%
				N = 224	*N = 127*	*N = 97*	*N = 53*	*N = 28*
	Have you replaced appliances or made changes	T1	1. No	91%	92%	89%	87%	89%
			2. Yes, namely …	9%	8%	11%	13%	11%

Table 5 (continued)

Topic	Questions [A]: for application user group [R]: for reference group	T1 or T2	Answer options	Responses				
				Total	Reference group	Application users	App A	App B
	to your home based on the energy consumption information? [A] [R]			*N = 224*	*N = 127*	*N = 97*	*N = 53*	*N = 28*
		T2	1. No	79%	88%	68%	66%	64%
			2. Yes, namely …	21%	12%	32%	34%	36%
				N = 224	*N = 127*	*N = 97*	*N = 53*	*N = 28*
	To what extent do you intend to invest in your home to lower your energy consumption levels? [A] [R]	T1	1. Not at all	30%	39%	19%	13%	21%
			2. Very little	34%	35%	33%	32%	36%
			3. A little	21%	16%	29%	34%	29%
			4. A lot	12%	8%	18%	19%	11%
			5. Quite a lot	3%	3%	2%	2%	4%
				N = 224	*N = 127*	*N = 97*	*N = 53*	*N = 28*
		T2	1. Not at all	47%	56%	36%	32%	43%
			2. Very little	28%	27%	30%	30%	39%
			3. A little	18%	14%	24%	26%	11%
			4. A lot	5%	2%	9%	11%	4%
			5. Quite a lot	1%	1%	1%	0%	4%
				N = 224	*N = 127*	*N = 97*	*N = 53*	*N = 28*
	(if intending to invest) What kind of investments do you want to make? [A] [R]		… (answers were coded)	–	–	–	–	–

OpenAccessThis article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

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

References

Abrahamse, W., Steg, L., Vlek, C., & Rothengatter, T. (2005). A review of intervention studies aimed at household energy conservation. *Journal of Environmental Psychology, 25*, 273–291.

Anderson, W., & White, V. (2009). Exploring consumer preferences for home energy display functionality. *Report to the Energy Saving Trust, 123*. www.cse.org.uk/app/webroot/downloads/file/consumer_preferences_for_home_energy_display.pdf. Accessed 28 Jan 2019.

Bonino, D., Corno, F., & De Russis, L. (2012). Home energy consumption feedback: a user survey. *Energy and Buildings, 47*, 383–393.

Buchanan, K., Russo, R., & Anderson, B. (2015). The question of energy reduction: the problem(s) with feedback. *Energy Policy, 77*, 89–96.

Darby, S. (2006). The effectiveness of feedback on energy consumption. *A review for DEFRA of the literature on metering, billing and direct displays*. Environmental Change Institute, University of Oxford. A Review for DEFRA of the Literature on Metering Billing and direct Displays, 22nd. www.volksmeter.de/Abhandlungen/Darby2006.pdf. Accessed 28 Jan 2019.

Darby, S., Liddell, C., Hills, D., & Drabble, D. (2015). *Smart metering early learning project: synthesis report*. United Kingdom: Department of Energy and Climate Change (DECC).

Ehrhardt-Martinez, K., Donnelly, K. A., & Laitner, J. A. (2010). *Advanced metering initiatives and residential feedback programs: a meta-review for household electricity-saving opportunities*. Washington: Report for ACEEE.

European Commission (2017). *Energy Efficiency Directive*. https://ec.europa.eu/energy/en/topics/energy-efficiency/energy-efficiency-directive. Accessed 2 Jan 2017.

European Union (2012). Directive 2012/27/EU of the European Parliament and of the Council of 25 October 2012 on energy efficiency, amending Directives 2009/125/EC and 2010/30/EU and repealing Directives 2004/8/EC and 2006/32. *Official Journal, L, 315*, 1–56.

European Union. (2016). *EU energy in figures – Statistical pocketbook 2016*. Luxembourg: Publications Office of the European Union ISBN 978-92-79-58247-9.

Fischer, C. (2008). Feedback on household electricity consumption: a tool for saving energy? *Energy Efficiency, 1*, 79–104.

Geelen, D., Reinders, A., & Keyson, D. (2013). Empowering the end-user in smart grids: recommendations for the design of products and services. *Energy Policy, 61*, 151–161.

Giordano, V., Gangale, F., Fulli, G., & Sánchez Jiménez, M. (2011). *Smart grid projects in Europe: lessons learned and current developments*. Petten: Joint Research Centre of the European Comission.

Gram-Hanssen, K. (2010). Standby consumption in households analyzed with a practice theory approach. *Journal of Industrial Ecology, 14*(1), 150–165.

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods, 18*(1), 59–82.

Hagaman, A., & Wutich, A. (2017). How many interviews are enough to identify Metathemes in multisited and cross-cultural research? Another perspective on Guest, Bunce, and Johnson's (2006) landmark study. *Field Methods, 29*(1), 23–41.

Hargreaves, T., Nye, M., & Burgess, J. (2010). Making energy visible: a qualitative field study of how householders interact with feedback from smart energy monitors. *Energy Policy, 38*(10), 6111–6119.

Hargreaves, T., Nye, M., & Burgess, J. (2013). Keeping energy visible? Exploring how householders interact with feedback from smart energy monitors in the longer term. *Energy Policy, 52*, 126–134.

Karjalainen, S. (2011). Consumer preferences for feedback on household electricity consumption. *Energy and Buildings, 43*(2), 458–467.

Karlin, B., Zinger, J. F., & Ford, R. (2015). The effects of feedback on energy conservation: a meta-analysis. *Psychological Bulletin, 141*(6), 1205.

Kobus, C., Klaassen, E., Mugge, R., & Schoormans, J. (2015a). A real-life assessment on the effect of smart appliances for shifting households' electricity demand. *Applied Energy, 147*(June), 335–343.

Kobus, C., Mugge, R., & Schoormans, J. (2015b). Long-term influence of the design of energy management systems on lowering household energy consumption. *International Journal of Sustainable Engineering, 8*(3), 173–185.

McKerracher, C., & Torriti, J. (2013). Energy consumption feedback in perspective: integrating Australian data to meta-analyses on in-home displays. *Energy Efficiency, 6*(2), 387–405.

Nilsson, A., Bergstad, C. J., Thuvander, L., Andersson, D., Andersson, K., & Meiling, P. (2014). Effects of continuous feedback on households' electricity consumption: potentials and barriers. *Applied Energy, 122*, 17–23.

Ofgem & DECC (2011). *Smart Metering Implementation Programme - Response to Prospectus Consultation, Overview Document*. Available from: https://www.ofgem.gov.uk/sites/default/files/docs/2011/03/sm-prospectus-reponse-overview_0.pdf. Accessed 28 Jan 2019.

Schwartz, T., Stevens, G., Jakobi, T., Denef, S., Ramirez, L., Wulf, V., & Randall, D. (2015). What people do with consumption feedback: a long-term living lab study of a home energy management system. *Interacting with Computers, 27*(6), 551–576.

Shove, E. (2010). Beyond the ABC: climate change policy and theories of social change. *Environment and Planning A, 42*(6), 1273–1285.

Stankovic, L., Stankovic, V., Liao, J., & Wilson, C. (2016). Measuring the energy intensity of domestic activities from smart meter data. *Applied Energy, 183*, 1565–1580.

Stieß, I., & Dunkelberg, E. (2013). Objectives, barriers and occasions for energy efficient refurbishment by private homeowners. *Journal of Cleaner Production, 48*, 250–259.

Strengers, Y. (2014). Smart energy in everyday life are you designing for resource man? *Interactions, 21*(4), 24–31.

Van Dam, S. (2013). *Smart energy management for households*. Doctoral thesis. Delft: Delft University of Technology.

Van Dam, S. S., Bakker, C. A., & Van Hal, J. D. M. (2010). Home energy monitors: impact over the medium-term. *Building Research & Information, 38*(5), 458–469.

Van Dam, S. S., Bakker, C. A., & Van Hal, J. D. M. (2012). Insights into the design, use and implementation of home energy management systems. *Journal of Design Research, 14, 10*(1–2), 86–101.

Van Elburg, H. (2014). *Monitor Energiebesparing Slimme Meters (Besparingsmonitor)*. Report Rijksdienst voor Ondernemend Nederland. https://www.google.nl/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwjn_Lqj65DgAhUSLFAKHQm7C9IQFjAAegQICRAC&url=https%3A%2F%2Fzoek.officielebekendmakingen.nl%2Fblg-301727.pdf&usg=AOvVaw1oYvtvmsPYtgXEY_tlMng3. Accessed 28 Jan 2019.

Verplanken, B., & Wood, W. (2006). Interventions to make and break consumer habits. *Journal of Public Policy & Marketing, 25*(1), 90–103.