

Exploring Persuasion in the Home: Results of a Long-Term Study on Energy Consumption Behavior

Patricia M. Kluckner¹, Astrid Weiss², Johann Schrammel³, and
Manfred Tscheligi^{1,3}

¹ ICT&S Center, HCI & Usability Unit, University of Salzburg, Austria

{Patricia.Kluckner,Manfred.Tscheligi}@sbg.ac.at

²ACIN - Automation and Control Institute, Vienna University of Technology, Austria
astrid.weiss@tuwien.ac.at

³CURE - Center for Usability Research & Engineering, Austria
schrammel@cure.at

Abstract. This paper presents a seven-months field study on a persuasive ambient display in private households. The *FORE-Watch* aims at adjusting the consumption behavior of energy users in 24 multi-person households and persuading them to change their timing of consumption activities; half of them were shown a forecast of the occupancy rate of the local energy grid (i.e., grid status), and the other half were shown how much energy will be delivered by green sources such as windmills (i.e., green energy). Our qualitative and quantitative survey and the energy consumption data revealed that the grid status group showed a more constant behavior than the green energy, indicating that the more dynamic forecast presentation did not lead to the same type of behavior change as the static forecast presentation. Overall, the *FORE-Watch* aroused awareness, attention and interest through permanent presence, clear and simple information, and changed the energy consumption behavior of our participants.

Keywords: Ambient persuasive display, eco-feedback and management technology, long-term in-situ study, changing energy consumption, home.

1 Introduction

The efficient use of energy-providing systems in private homes is an important step towards reaching the overall climate goals and the transformation of our society's sustainable energy consumption. A promising approach to address this problem is the utilization of ambient technologies informed by the field of persuasion. This choice is based on the insight that a change in individual and collective attitude towards the environment will result in a necessary behavioral change to reach the goal sustainability. Persuasion for sustainability can be challenging since people are not always intrinsically motivated to adopt a more environmentally friendly lifestyle. Therefore, information of individual benefits needs to be provided additionally to the collectivist goal of "saving the planet". Improving the efficiency and management of domestic energy consumption is an important area to save energy. Providing

customers with detailed feedback on their energy consumption supports them in changing their behavior towards CO₂ reduction. Neustaeder et al. [14] provide insights about how families understand the relationship between everyday activities and energy consumption. Ambient displays have been identified as being especially suitable to communicate this kind of information to the end-user. Various systems providing ambient feedback on the current energy consumption (e.g., Choreflect [20], POEM [13], eMeter [11], or the well-known Energy AWARE Clock¹) have been developed so far. However, up to date, very few long-term studies in real households have been conducted, which try to understand the influences of such ambient displays over time (e.g., Smeaton and Doherty [24], Midden et al. [12], Staats et al. [25]).

Studying persuasive strategies in the field is a challenging endeavor in determining how to measure the persuasiveness of a system and to ensure that an attitude or behavior change is sustainable. In the framework of the PEEM project (Persuasive End-User Energy Management), the *FORE-Watch* (Forecast Of Renewable Energy - Watch) was developed [22] based on a user-centered design approach [7]. It is an ambient persuasive display, which gives users feedback on their energy consumption, including forecast of energy availability. Our aim is to influence the energy consumption behavior with regard to the timing of consumption activities; in other words to consume energy when the forecast presents it. To investigate if the interface can achieve this change, we installed the *FORE-Watch* in 24 private multi-person households and accompanied and monitored them for a period of seven-months.

After an overview of related research, we describe persuasive systems and studies for energy consumption behavior change, in addition to presenting the link between persuasion and change with user studies. In order to prove the usefulness of the *FORE-Watch*, we show the experimental development of our long-term study in-situ of private households. We present the results for attitudinal changes and behavior modification measured over a period of time and related results.

1.1 Ambient Persuasive Systems for Change

Many studies have investigated the possibilities to persuade people (i.e., feedback, commitment, rewarding, etc.) within the context of sustainability. A study to understand consumers' awareness of energy consumption in the home and to determine their requirements for an interface led to a three-stage approach to support electricity conservation routines: raise awareness, inform complex changes, and maintain sustainable routines [21]. Awareness and reflection of household tasks on inhabitants' contribution was investigated [20] and increased engagement through the distributed ambient persuasive display, named Choreflect. While researching visual signals to induce interaction, Kukka et al. [9] identified a behavior termed display avoidance that people reveal with interactive public displays.

Mobile phone applications are a comfortable way to achieve awareness of saving electricity. Applications can be a suitable means of changing human behavior. An investigation of a mobile tool, UbiGreen, for tracking and supporting green

¹ <http://www.tii.se/node/5984>

transportation habits was developed [6] to explore the use of personal ambient displays and to give users feedback about sensed and self-reported transportation behaviors. Driven by the spread of smart meters (e.g., Strengers [26; 27]) and energy monitors, research on portable and stationary energy consumption displays resulted [28] that the comparison between these monitors showed that users took both displays to identify high-power devices in their home whereby conserve energy. But after the initial survey, participants used the portable displays as stationary ones. A current eco-feedback technology design, the eco-end user application prototype POEM (Personal Office Energy Monitor) was presented by Milenkovic et al. [13]. Petkov et al. [18] tried to understand what feedback individuals find relevant for energy savings. They designed mock-up screens that provide eco-feedback catering towards different pro-environmental values and concerns, and asked users to evaluate them. The most respondents ranked the egoistic screen as highest. They show how environmental psychology could play a role in informing the design of persuasive applications that motivate energy saving behavior. Also, Massung et al. [10] offered new insights and recommendations on the pro-environmental design of systems that target groups and communities.

In [15] a theoretical foundation is offered for how to study Behavior Change Support System (BCSS) defining behavior change archetypes as well as design principles for such systems. A central cluster of work focused on Electricity Consumption Feedback and highlighted emerging energy system trends with strong relevance to HCI and interaction design, including smart grid, demand response, and distributed generation technologies [19]. They outlined a range of opportunities to engage emerging systems referring to experimental, behavioral, social, and cultural aspects; new areas move beyond, i.e., energy feedback displays, increase awareness, and motivate individual conservation behavior.

1.2 Studying Long-Term Change through Persuasive Systems

Summarizing the above-mentioned studies, a lot of work has already been done in order to influence consumption behavior with respect to energy. However, only little work exists on long-term persuasive effects of these persuasive systems. The idea to compare on-site social stairs vs. “Piano-Staircase” (each step plays a note) toward long-term behavior change was currently investigated [17]. The goal was to help users to activate intrinsic motivation and not to persuade or extrinsically motivate them to act or behave differently in that moment of time. However, their outcome could not present a solution of which trigger lead to intrinsic motivation.

Moreover, we want to mention three specific long-term studies in the energy sector similar to our research. We reference the current work of Smeaton and Doherty [24] as linkage between energy consumption and long-term data collection. In their experimental setup, 24 households were involved and they collected electricity consumption data over a period of 18 months. Four real-time intervention methods to impact the reduction of consumption were used. Due to the fact that we also used a display in our research, we focused on the results of their touchscreen display. Further interventions were an initial web page, e-mail rewards, and a web portal. First,

background information was gathered and built up into an energy consumption behavior profile. Second, participants got the touchscreen application, which noted recorded savings. Results showed that thirteen regular households saved 14% in one month on average after receiving access to the desktop display, 9% two months later, 14% three months later, and 16% four months later. They concluded their work with the statement that a variety of feedback mechanisms and interventions may be needed in order to sustain user interest. The EcoTeam Program [25] is an intervention package that aims to target considerable number of behaviors and durable change in pro-environmental behavior. A detailed analysis of the three-year longitudinal study of behavior suggests that the interplay between the factors behavioral intention and habitual performance can predict participation before the study and the degree of social influence experienced during participation. A focus on the appraisal of climate risks and interactive approaches to influence energy consumption in households are mentioned [12]. They state that sustainable consumption is conceptualized as the result of various types of interactions between users and systems.

1.3 Research Gap

The information needed for appropriate actions is not presented in the context where it is needed most, i.e., when interacting with the home appliances or environment. Therefore, the feedback lacks a direct and tangible link to the consumers' behavior. Beyond self-monitoring, ambient displays made it easier for the user to reflect their behavior or attitude by arising unconscious aspects, such as our electricity consumption, to our awareness. These aspects give us the possibility to understand and react accordingly. An ambient/situated display can support reflection and allow adjusting reactions. Current studies of such displays frequently have shortcomings with regard to long-term effectiveness; initial results tend to wear off once the novelty effect is over. We clearly acknowledge the value and the need of short-term studies in order to improve the design of a system, e.g. [3] and to identify major issues (e.g., usability). There are few studies that investigate ambient persuasive displays in private households over a long period to get insights about attitude and behavior change. We claim that in order to gain an understanding how our persuasive systems need to be designed to have a sustainable persuasive interface, long-term field studies with actual representative users and mature deployed systems are required. We contribute an essential part to comprehend persuasion and sustainability.

Our research attempts to focus at improving the communication of energy feedback by seamlessly integrating it in the environment of the user and providing it where and when it is most useful and efficient. Such an integration of feedback could increase the comfort of the users, as no abstract translation and explicit attention towards achieving the goals is needed. Additionally, we provide forecast information to set a link between timing of activity and feedback information for maximum efficiency and conclusive traceability for the feedback values. The objective of our study is to research and develop new strategies and tools for the home context that provide energy-saving related feedback in a persuasive and unobtrusive way and, thereby, have the potential to help to reduce energy consumption without loss of

comfort. We explore persuasive technologies to influence behavior towards optimized end-user energy management. Tailored persuasive approaches overcoming limitations of existing solutions have been developed, prototyped, and experimentally validated with real user in realistic long-term settings. Thus, we decided to conduct a seven-month study in 24 households, which is accompanied by several different questionnaires at various points of measurement to identify that aspects change after which period of time.

2 The *FORE-Watch*

The aim of our study was to explore if this eco-management technology, *FORE-Watch*, influences the energy consumption behavior, specifically with regard to the timing of consumption activities, either to consume energy when the grid status is high or when green energy is available, over a longer period of time. The *FORE-Watch* triggers energy consumption behavior through cues equipped with persuasive strategies. Our seven-months in-situ household study involved 24 running systems on tablets and installed smart meters in private homes. The study was accompanied by several different questionnaires at various points of measurement to identify if attitudes and behaviors are changing over time (t) and if so which aspects of the App caused the change. Moreover, we wanted to explore if the different forecast methods also impact how participants experience the *FORE-Watch* and its persuasive potential.

2.1 Design and Implementation

We want users to align their energy consumption with requirements from production and distribution within the grid. The design process was based on various sources, both theoretical (literature review) and empirical-based (user requirements and design experts). We decided to focus on the familiar and well-known system of the traffic light system with red, yellow, green, to adjust action. Due to the fact that this color code was recognized unconsciously and intuitively, there is little cognitive demand. Furthermore, the device is based on a kitchen clock and shows additional information to the time of day. The *FORE-Watch* consists of two views: the forecast and the feedback. Each view presents three cues equipped with persuasive strategies (see Fig. 1). The instruction of activities for the participants is as described. There are three colors: green represents a request of action (e.g., turning on the washing machine), yellow an action realization is not recommended, and red an action realization should be avoided. Darker shades represented unsuitable moments for action, while brighter shades depict a suitable moment for action to obtain a smooth transition. This color-triad also represents good, bad, medium performances based on the adjustment at the forecast and the KW/h-consumption. In the following we present each cue.

Clock. The clock displays the forecast of green energy/grid status profile of the next hour in detail per minute. The background information about the different forecast views is more complex. The colors indicate activities, but the composition of (1) static, dependent of the grid load patterns and (2) dynamic, dependent of the

availability of the eco-electricity (windmills) are different. (1) Green means a lot of energy is available and convenient to consume the stream and perform action; yellow means there is a moderate amount and red reflects limited energy available. (2) Green means a lot of electricity is produced, unfavorable to consume the stream and perform action; yellow, medium level of energy; and red, there is little electricity produced.

Timeline. This timeline indicates the forecast of the next 24 hours. The kink of the axis indicates the current hour and divides the areas of “*today*” and “*tomorrow*”.



Fig. 1. Forecast view: clock, timeline “*today*” and timeline “*tomorrow*” (left), and the feedback view: average consumption, consumption per phase, and reward icons (right).

Average Consumption. The display shows the average electricity consumption (in KW/h units) over a week, the day before is labeled as “*yesterday*”. The dot represents the average from baseline start.

Consumption per Phase. This implies: How much energy to a certain red, yellow, or green zone of the clock was consumed in proportion to the KW/h consumption. It uses two modalities of feedback electricity consumption (1) bar chart, which is divided into a week (the day before is labeled as “*yesterday*”) and the average consumption from baseline start, and (2) reward icon for honor that are three levels of performances/adjustment of the forecast (gray-shades for good, very good, excellent).

2.2 Integration in the Home

In the following section, we present how the *FORE-Watch* was integrated in the private homes of our participants, explaining the interaction loops between the smart meter, the tablet on which the *FORE-Watch* was running, and the servers where data was stored. Furthermore, we explain how the color and feedback visualizations for seven days were calculated.

Forecast Data. The grid status data (*static forecast*) was provided daily by the energy consumption company. The data was computed according to a mathematical model by the usability research and engineering company. For the green energy group, the wind forecast (*dynamic forecast*) was used from the Norwegian

Meteorological Service. For our purpose, the data of Salzburg/Austria were retrieved. This forecast is only available in a six-hour rhythm; the values were interpolated to get a fifteen-minute interval. For both conditions, data were sorted and transformed into color (top third assigned green, second third yellow, and final third red).

Feedback Data. The smart meters sent the consumption data of the customer to a database, which runs on a server, via IP gateway and cable modem (wireless). The consumption data from our cooperating energy consumption company, which was stored in a database on a server, sent this data to one of our servers and categorized it into phases. Thus, the consumption was immediately updated in that current color which is then displayed in the bar chart. At the end of the day, the totals for the various colors were compiled. In a similar procedure the trophies were developed.

3 The Long-Term Field Study

Our long-term study was set-up as a low-intrusive study with little researcher involvement. Therefore, we decided to start the field phase with kick-off workshops in which the participants were informed about the *FORE-Watch*, followed by the gathering of all attitudinal measures by the means of online surveys in which participants could conveniently fill in at home. At the end of the study, final reflection workshops were conducted with the participants in order to gather qualitative data, as well as give participants the opportunity to provide feedback. Moreover, the integration of the *FORE-Watch* in the smart meter architecture allowed us to constantly gather KW/h-data on the energy consumption behavior.

The following set of research questions synthesizing attitudinal change drives this exploratory household study. (1) Can an attitudinal (behavior) change be observed in the perception of the *FORE-Watch* depending on the various forecasts (static vs. dynamic)? (2) Can an attitudinal (behavior) change be observed in the perception of the *FORE-Watch* depending on the various points of measure?

3.1 Participants

A total of 28 users (24 energy consultants and four of their partners) participated in our seven-month study ($M=44.85$, $SD=8.20$ years; educational levels: 75.0% teaching/professional school, 10.7% high school/ a-levels, 10.7% college/university). The recruitment was done with consulting experts of the energy company. This procedure was efficient on the basis of the ongoing time and the effort to install the smart meter and the IP-Gateways; the predefined count of smart meters reduces the households that participate at the study. As a result, we got a more homogenous group (men, between 29-55 years), where we could calculate tendencies. We separated this group of 24 households into two parts with random distribution of the living space in square meter, the type of house (one-family house, apartments, and townhouse), and the number of household members. We made this separation because we used the possibility to integrate two forecast conditions: the dynamic one, which is the green energy profile, and the static one, which represents the grid status group.

3.2 Measurement

We used online surveys as evaluation of the *FORE-Watch* in various trials during the study to gather self-reporting data. These questionnaires gave us results about attitude and awareness change. The various types of questionnaires can be distinguished in two categories: (1) the *FORE-Watch* questionnaires includes the Single-Cue Questionnaire (SCQ), and the Grouped-Cue Questionnaire (GCQ), and (2) the User Characteristic questionnaires (UCQs) with the General Questionnaires (GQs) the Profile Questionnaires (PQs), and the User Experience Questionnaire (UXQs). We used the UXQs to get an overview of the users' experience interacting with *FORE-Watch*, their motivation in different phases during seven-months, and their trust into the application. Moreover, quantitative data was collected through the integration of the *FORE-Watch* in the smart meter architecture that allowed us to constantly gather data on the energy consumption behavior (KWh units). The UCQs consisted of a compilation of various sub-questionnaires. The advantage of these questionnaires was especially in their standardization and, thus, the repeated use monitored changes in attitude and self-evaluation to determine attitude changes. The frequency of measurement time points enabled us to include smaller setting changes within the study period, which would otherwise have remained hidden. This supports the measurement of long-term patterns of behavior and attitude in real environments.

3.3 Procedure

We started the field phase with kick-off workshops to inform participants about the *FORE-Watch* and to complete the following parts of the UCQs. The GQs: *technic affinity questionnaire (TA-Q)*², *general motivation questionnaire (GMQs)*³, the PQs: *Meyers-Briggs Type Indicator (MBTI)*⁴, and *NEO Personality Inventory (NEO-FFI)*⁵. The study closed with final reflection workshops in order to gather qualitative data (recording) as well as to give participants the opportunity to give feedback. During the study, all measures were taken by the means of online surveys sent to fill in via web link (see Fig. 2). The UXQs was measured three times at twelve-week intervals over seven-months as well as the SCQ; the GCQ ten times at two-week intervals.

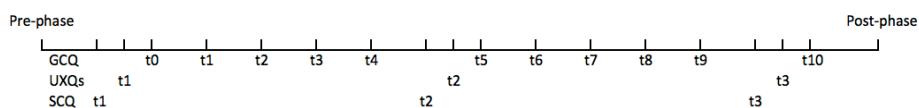


Fig. 2. Long-term data collection.

2 Karrer, K., Glaser, C., Clemens, C., Bruder, C.: Technikaffinität erfassen – der Fragebogen TA-EG. In A. Lichtenstein, C. Stöbel und C. Clemens (Hrsg.): *Der Mensch im Mittelpunkt technischer Systeme*. 8. Berliner ZMMS Spektrum, Reihe 22, Nr. 29, S.196-201 (2009).

3 <http://www.psych.uni-potsdam.de/people/rheinberg/messverfahren/index-d.html>

4 Briggs, K., Myers, I. (1989) developed; German version: Bents, R. and Blank, R. (1995).

5 Borkenau, P., Ostendorf, F.: *NEO-FFI*, 2nd version personality questionnaire (2008).

In the following, we present in detail the UXQs. The *Intuitive Interaction (INTUI)*⁶ is a tool for detecting various components of intuitive use and experience while their product usage. The *Positive and Negative Affect Scale (PANAS)*⁷ detects the emotional state. The *Short Scale of Intrinsic Motivation (SIM)*⁸ estimates subjective experience in terms of an activity. The *Unified Theory of Acceptance and Use of Technology (UTAUT)*⁹ is a tool that provide a unified theoretical basis from which to facilitate research on information system and technology adoption and diffusion. The *Technology Acceptance Model2 (TAM2)*¹⁰ used the approach of behavioral intention and extend it with social influence and cognitive process variables. The *Attractiveness makes the Difference 2 (AttrakDiff2)*¹¹ rates the user experience product perceptions and evaluations. We used the *TRUST*¹² questionnaire, the *Need Scale (BDS)*¹³, and the Motivation Questionnaires (*Incentive Focus Scale (AF-S)*¹⁴, the *Perceived Locus of Causality (PLOCQ)*¹⁵, the *Situational Intrinsic Motivation Scale (SIMS)*¹⁵).

4 Results and Interpretation

As mentioned before, the results of our long-term study (March-September 2012) are the online survey results of the UXQ, the procedure of the data analysis, and the results and implications of the attitudinal change of participants over time. With a 76% response rate, we achieved an excellent result, as 20% is an average common

6 Ullrich, D., Diefenbach, S.: Exploring the Facets of Intuitive Interaction. In Ziegler J., Schmidt, S.: Mensch & Computer, 251-260 (2010).

7 Watson, D., Clark, L. A., Tellegen, A.: Development and validation of brief measures of Positive and Negative Affect: The PANAS scales. In J of Personality and Social Psychology, 54, 1063-1070; German translation: Krohne, H.W., Egloff, B., Kohlmann, C.-W., Tausch, A. (1996). Experimental study with the German version of „Positive and Negative Affect Schedule“ Diagnostica 42, 139-156 (1998).

8 Wilde, M., Bätz, K., Kovaleva, A., Urhahne, D.: Überprüfung einer Kurzskala intrinsischer Motivation, 31-45 (2009).

9 Venkatesh, V., Morris, M.G., Davis G.B., Davis, F.D.: User Acceptance of Information Technology: Toward a Unified View, MIS Quarterly, 27 (3), 425-478 (2003).

10 Venkatesh, V., Davis, F.D.: A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. In J Management Science, 46 (2), 186-204 (2000).

11 Hassenzahl, M., Burmester, M., Koller, F.: AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In: Ziegler, J., Szwilus, G. (Hrsg.), Mensch & Computer 2003. Interaktion in Bewegung, 187-196 (2003).

12 McKnight D.H., Carter M., Thatcher J.B., Clay P.F.: Trust in a specific technology: An investigation of its components and measures. In Trans. Manag. Inform. Syst. 2, 2, Article 12, 25 pages (2011).

13 Hassenzahl, M., Diefenbach, S., Göritz, A.: Needs, affect, and interactive products – Facets of user experience. Interacting with Computers, 22 (5), 353-362 (2010).

14 <http://www.psych.uni-potsdam.de/people/rheinberg/messverfahren/index-d.html>

15 Lonsdale, C. Sabiston, C.M., Taylor, I.M., Ntoumanis, N.: Measuring Student Motivation for Physical Education: Examining the Psychometric Properties of the Perceived Locus of Causality Questionnaire and the Situational Motivation Scale, Psychology of Sport and Exercise 12, 284-292 (2011).

result for email based online surveys [1]. The recruitment of energy consultants as participant's maybe a reason for the high commitment and the fact that the *FORE-Watch* as intervention in the home was a constant reminder for study participation.

In the data analysis, our external project partner screened the KW/h-data of the forecast view and the feedback view. They analyzed the consumption data from smart meter with the access to the fifteen-minute consumption values of the users. All smart meter failures were adjusted. An assessment of displacement effects could be enabled per user by means of time differentiated in fifteen minutes intervals and were calculated by working days, Saturdays, or Sundays/holidays. For each of these intervals, a difference between current consumption and mean over seven-months was calculated. The data transfer of the forecast always ran stable, as it had a separate exchange of information cycle. Due to indirect monitoring, we could calculate that on average, the *FORE-Watch* was activated over 60% of the time. However, based on the missing consumption cycle or interruptions of the feedback data, we had to eliminate several item sets of the attitudinal questionnaire belonging to the feedback view. We used the following criterion for data exclusion. If the consumption signal was unstable less than five out of fourteen days, the data for this period of time was excluded as we defined that five consecutively days are a decent reference score. In order to identify whether or not a questionnaire had to be excluded, we compared the timestamp of the missing data with the questionnaire schedule. Then, we searched for the missing consumption data and checked if the person was on vacation or if response patterns could be identified in the data.

4.1 Findings on Attitudinal Change and Implications

We present selected findings of the UXQs, which give insight into the attitudinal change of our two user groups at three times of measurement. Our data was not normally distributed, therefore, we depict in Table 1 the bivariate correlation (Spearman) stratified by total group (TG): static and dynamic group (Mann-Whitney *U*-Test) and all significant effects (*p*) for all times of measurements (*t1-t3*).

No significant effects (all *p*'s = >.05) during *t1* to *t3* in the total group and between the static vs. dynamic group were revealed at the Need Scale¹² and the Motivation Questionnaires^{13,14,15}. Therefore, we conclude that participants responded similar to these questionnaires during the run of the always-on *FORE-Watch*.

Significant results could be identified for the **INTUI** questionnaire. The group answer differs in the scale *magical experience* at *t3* (*p*=.047), whereas static has higher scores (*Mdn*=4.50; dynamic *Mdn*=4.00), i.e., the interaction with the product is seen as extraordinary and fascinating. The scale goes beyond the fulfillment of tasks. **Implication:** Eco-feedback and management products achieve their value not at the beginning of their usage; their interaction value rises after a period of time.

The analysis of the **PANAS** revealed at the scale "How did you feel in the last days? (LD)" and the scale "How did you feel in general? (G)" on *t1* significant (*p*_{LD}=.009; *p*_G=.012), whereby the static reached higher scores (*Mdn*_{LD}=30.40; *Mdn*_G=34.30; dynamic *Mdn*_{LD}=19.10; *Mdn*_G=26.20). In both scales, the positive affect scale achieved higher scores as the negative affect scale. Therefore, the emotional

state during the usage of the display was positive with the qualities: active, excited, attentive, enthusiastic, committed, excited, interested, strong, proud, and alert. **Implication:** Ambient persuasive displays obtain a positive effect over time.

For the *interest/enjoyment* scale of the **SIM** at *t1*, a significant difference could be identified ($p=.034$), whereby dynamic reached higher scores ($Mdn=2.67$; static $Mdn=1.33$). At the beginning of the study, the interaction with the *FORE-Watch* was more entertaining and interesting for the dynamic forecast group. **Implication:** The factors fun and enjoyment seemed to be a short-term phenomenon. Due to the fact that enjoyment and interest are very flexible, individual depended variables, and nearly impossible to sustain, flexibility in information presentation through updates could be the sustainable solution to maximize enjoyment and interest for a product.

The analysis of the **UTAUT** questionnaire revealed that at *t1* the scale *effort expectancy* calculate a significant difference ($p=.033$), whereby dynamic reached higher scores ($Mdn=19.25$; static $Mdn=16.50$), meaning that they considered more effort to use the application than the static. This result was expected at the beginning of our study, as the forecast view of green energy is more dynamic than that of the grid status, which was clearer, more comprehensible, and easier to learn and to use at the beginning. **Implication:** Displays who present information in a dynamic way are more complex, require more attention, and take a longer time to pass to routine/habit. At *t1*, the scale *behavioral intention to use the system* the groups differed significantly ($p=.019$), in which the static ($Mdn=14.00$) has higher scores (dynamic $Mdn=10.67$). That indicates that they presented a higher intention of long-term use. **Implication:** After short-term use, the static group recognizes a higher potential for long-term use of our ambient persuasive display. This may have caused due to the fact that they reached their routine behavior earlier, as their presented information is more periodic.

Table 1. User Experience Questionnaire: significances.

Questionnaire	Time	Scale	Total Group		Static & Dynamic			
			<i>t</i>	r_s	p^*	<i>U</i>	<i>Z</i>	p^*
INTUI	3	Magical Experience	.42	.040	33.5	-1.99	.047	-.41
PANAS	1	Last Days	-.52	.006	32.5	-2.60	.009	-.51
	1	In General	-.50	.009	34.0	-2.52	.012	-.49
SIM	1	Interest/Enjoyable	-.41	.031	52.0	-2.12	.034	-.40
UTAUT	1	Effort Expectancy	-.43	.030	41.5	-2.13	.033	-.42
	1	Behav. Intention ¹	-.47	.015	37.5	-2.35	.019	-.46
TAM2	1	Perceiv. Usefulness ²	-.41	.038	43.5	-2.04	.041	-.40
	3	Perceiv. Usefulness ²	-.45	.025	38.0	-2.19	.029	-.44
	1	Output-Quality	-.51	.008	34	-2.54	.011	-.50
	2	Output-Quality	-.58	.002	27	-2.91	.004	-.57
	1	Result Demonstr. ³	-.44	.026	40.5	-2.18	.029	-.43
AttrakDiff2	1	Pragmatic Quality	-.59	.001	25.5	-2.97	.003	-.58
	1	Attractiveness	-.60	.001	25.0	-3.00	.003	-.59
	1	Hedonic Qu.-S. ⁴	-.70	.000	15.5	-3.49	.000**	-.68

* $p = 2$ -tailed

** Highly significant

¹Behavioral Intention to Use the System

³Result Demonstrability

²Perceived Usefulness

⁴Hedonic Quality - Stimulation

The rating for the scale *perceived usefulness* of the **TAM2** on *t1* and *t3* differed significantly ($p_{t1}=.041$; $p_{t3}=.029$). The static ($Mdn_{t1}=9.67$; $Mdn_{t3}=11.17$) had in both times of measurement higher scores (dynamic: $Mdn_{t1}=8.67$; $Mdn_{t3}=7.00$). The results depict that the static view shows an easier handling with the application, which indicated a higher perceived usefulness at *t1* and at *t3*. In the mid-term, no differences were detected. At the beginning of the usage, the interaction with the application is for both groups necessary to understand how it supports their decision and manage behavior. **Implication:** Perceived usefulness is an important aspect that underpins the success of our watch over time, as it is rated important not only short-term but also long-term. But, we also recognize that it plays no role any longer after three months. At the scale *output-quality* on *t1* and *t2*, the groups were significantly different ($p_{t1}=.011$; $p_{t2}=.004$); the static ($Mdn_{t1}=5.50$; $Mdn_{t2}=6.00$) had in both times of measurement higher scores (dynamic $Mdn_{t1}=4.50$; $Mdn_{t2}=4.50$). At the beginning and in mid-term, the grid status group stated that they are satisfied with the quality of data and that the quality is generally very high. **Implication:** The data output should be presented in a high quality, balanced way. There were group differences in the positive sense about their opinion and attitude in short-term and mid-term. At the scale *result demonstrability*, the groups at *t1* differed significantly ($p=.029$), whereby the static ($Mdn_{t1}=12.50$) scored higher (dynamic $Mdn_{t1}=10.75$). This scale shows that the user has no difficulty to tell others about the results; it is easy to do or to explain the impact of the application, as well as that fact that the usage of the application is obvious. At the beginning of the study, the groups gain positive insights about the meaning of the *FORE-Watch*. They know about the background of the display and feel well informed about it. **Implication:** Knowledge about a product itself, background information, and ‘sense-attributing’ application as reduction or conservation of energy can be an indication for long-term usage.

For the scale *pragmatic quality* of the **AttrakDiff2**, the groups differed significantly in the rating ($p=.003$) at *t1*. The static ($Mdn=29.86$) had higher scores (dynamic: $Mdn=25.71$). This predicts that certain users may use it to achieve the specified goal, as to adjust action with effectiveness, efficiency, and satisfaction. The ratings of the product as simple, practically, and clearly is an indication that the *FORE-Watch* was very well accepted; there were group difference at short-term, which disappeared afterwards. **Implication:** Simplification and practical applications are well-known criteria for the success and efficiency of products. Significant results ($p=.003$) were presented for the scale *attractiveness*, whereby the static ($Mdn=27.86$) had higher scores (dynamic: $Mdn=24.57$) at *t1*; the application is among others considered nice, good, attractive, and inviting. **Implication:** Attractiveness is a strong component of the advertisement like “attractive products sell better.” The success of an ambient persuasive display seems also to be depended on its attractiveness. For the subscale *hedonic quality - stimulation* significant differences between groups ($p<.001$) could be found at *t1*, for which the static ($Mdn=28.57$) had higher scores (dynamic $Mdn=24.57$). The group output presents that the App stimulated the human need of personal development, i.e., the impact of knowledge and skills is different by groups. The *FORE-Watch* is innovative. **Implication:** Ambient persuasive displays,

which are convincingly innovative, have a higher potential of usage, as they show their actuality and novelty as well as their future view of our society.

After the presentation of the results in granularity, we can answer the research questions. In the perception of the *FORE-Watch*, we observed an attitudinal change between the various forecast groups and points of measurement. The green energy group rated the forecast presentation (and the feedback presentation) more positively as the grid status group. At the three times of measurement, the grid status group had higher scores, especially at $t1$, than the green energy group.

4.2 Behavioral Change

The analysis of the KW/h-consumption of the two experimental groups resulted that a low correlation between consumption behavior and the dynamic group exists; a weak trend is presented for the static group. This difference was expected since the dynamic group due to the different temporal dynamics of the forecast had more room for a maneuver. Moreover, the calculation of “consumption per phase” points out that behavioral change appeared to occur particularly in red phases. Shifting of activities occurs also into yellow and not only in green phases, which can be explained by the proximity in time. This shifting into yellow (adjacent) and non-green area maybe resulted in the fact that the time horizon for activity shifting is very short and depends on type of activity. This can possibly be explained by increased consumption in the yellow zone with a simultaneous reduction in the red zone. Furthermore, the absence of real-time consumption data, the non-presented data from the pre-phase of the study and the temporal periodicity of the forecast, the evidence of effects for the static group is particularly difficult. Over the period of seven-months, we analyzed the temporal periodicity (24 hours) of the forecast and, contrary to our expectations, we find out that the static forecast presented information more dynamic (s-curve with the peak around 11:00 a.m.) than the green energy/dynamic forecast itself. Belonging to the weekly forecast presentation, we saw recurring patters in the grid status group and slowly shifting and atypical timely accumulations.

However, there are some indicators for the coherence of the KW/h-data: (1) data is consistent with the qualitative information of the subjects, (2) data shows expected differences between dynamic and static group, and (3) there is a magnification of the effect during the operating hours of the devices.

5 Discussion

In sum, six out of ten questionnaires used over three different points of measurement revealed significant results, i.e., 12 out of 32 scales show answer differences by groups. Five out of ten questionnaires present significant results at the beginning of the study. All questionnaires showed significant differences between the groups at $t1$ excluding the *INTUI*, which revealed significant results at the scale *magic experience* at $t3$. At these three times of measurement, the grid status group had in eleven out of fourteen scales higher scores than the green energy group. Interestingly, the results of

the *TAM2* confirmed differences for two scales at two times of measurement, *perceived usefulness*: presented significant differences between $t1$ and $t3$, and *output-quality* depicted significant differences between $t1$ and $t2$.

In the next section, we discuss our results for further research in the field of ambient displays and for the metering infrastructure and smart meters. Our qualitative survey revealed that the grid status group showed a more constant behavior on the basis of the considered answers than the green energy group, indicating that the more dynamic forecast presentation of the green energy did not lead to the same type of attitude change as in the grid status group. This result of the groups cohere with our assumption that the green energy group needed more time to be in touch with the *FORE-Watch*, i.e., to understand the presented information. Due to our energy consultants, we assume that other population groups need far longer to adjust their activity rhythm at the *FORE-Watch*. We suppose that the dynamic forecast group status caused a steeper learning curve and that the forecast is needed as a supportive tool in long-term use, as the availability of green energy does not follow a constant pattern. The grid status group could quickly learn the grid status forecast with the watch and, therefore, change their behavior sustainably (even without a forecast on the long run). Especially for the first time of measurement, the grid status group had higher scores in four out of five questionnaires than the green energy group. The green energy group presented various results regarding to the answers of the *INTUI* at the third time of measurement. In the mid-term and at the end of the study, the grid status group showed higher scores at the *AttrakDiff2*. These results indicated that the grid status group at the beginning of the study was more satisfied in using the App than the green energy group. At the end of the study, the acceptance increased for the grid status group and the intuitive use of the application for the green power group. Based on the questionnaires of motivation and the need scale, both seemed to be factors which are independent and not connected to influence participants during the use of the *FORE-Watch*. At the issue of users' behavior change, we can emphasize a magnitude of the effect size over the seven-month in-situ study; there is small significant correlation by the dynamic group and the static group presents a tendency.

In summary, attitude change and behavior change took place: the grid status group had more attitude changes by the *FORE-Watch* at the beginning of the study as the green energy group. However, this fact changed, and disappeared at the end of the study. Overall, the green energy group rated the forecast presentation and the feedback presentation more positive than the grid status group and showed more behavior change. Altogether, the behavior change arose in the red phases as well as modifications occurred in the yellow and also green phases.

In comparison to similar approaches (e.g. Peeters et al. [17], Smeaton and Doherty [24], Staats et al. [25]), we also advocate that intervention tools lead to changes of attitude or behavior. We understand the *FORE-Watch* as a trigger to target durable change in pro-environmental behavior. We agree with the argumentation of Midden et al. [12] that sustainable consumption is conceptualized as the result of various types of interactions between users and systems, and argue that our App equipped with various strategies supports the user-system interaction every time. Our study showed that the overall concept of the *FORE-Watch* as ambient persuasive display was

feasible for our participating households in a seven-month period of time, which leads us to the conclusion that it could sustainably change energy consumption behavior.

6 Conclusion

In general, our results presented that a lot of effects on user experience factors are relevant at the beginning of a study (short-term), but this effect decreases after seven-months. In long-term usage of the *FORE-Watch*, only the *TAM2* presents significant group difference on two scales at two times of measurement and the *INTUI* calculate significant differences between groups. Our results led us to the implication that the users showed relevant experiences at the first usage of the eco-technology and, if the usage were ongoing, the application would be integrated in daily life whereby other kinds of experiences occur and become important/essential for long-term sustainable change. As to answer our research questions, we can summarize that the user's attitude and behavior was modified during seven-months (three times of measurement) and between the groups justified in various forecasts.

With our work we contribute to the first long-term studies of user-centered developed persuasive ambient displays for home energy management. We can review that the variation of questionnaires was very fruitful in terms of methodology and give an extensive view of user's motivation, emotion, need, technology acceptance, and trust. The future step for this system include an in-depth investigation of persuasive cues that support the usage and promise to convey the desired behavior modification, without the interventions are experienced as intrusive or restrictive.

Acknowledgement. This work is part of the PEEM-Project (Persuasive End-user Energy Management) and funded by the Austrian Research Promotion Agency (FFG). It was supported by project partners Salzburg AG, and Center for Usability Research and Engineering (CURE).

References

1. Andrews, B., Nonnecke, D., Preece, J.: Electronic survey methodology: A case study in reaching hard to involve internet users. In Int. J of HCI, 16(2): 185-210 (2003).
2. Beale, R., Courage, C., Hammontree, M., Jain, J., Rosenbaum, S., Vaughan, M., Welsh, D. Longitudinal Usability Data Collection: Art versus Science? In Proc. CHI EA'08
3. Costanza, E., Ramchurn, S.D., Jennings R.J.: Understanding Domestic Energy Consumption through Interactive Visualisation: a Field Study. In Proc. UbiComp'12.
4. Fogg, B.J., Eckles, D. Mobile Persuasion: 20 Perspectives of the Future of Behavior Change. Stanford Captology Media (2007).
5. Froehlich, J., Dillahunt, T., Klasnja, P., Mankoff, J., Consolvo, S., Harrison, B., Landay, J.A.: UbiGreen: Investigating a mobile tool for tracking and supporting green transportation habits. In Proc. CHI '09, 1043-1052 (2009).
6. Froehlich, J.: Promoting Energy Efficient Behaviors in the Home through Feedback: The Role of Human-Computer Interaction. In Proc. HCIC'09 Winter Workshop Boaster Paper.
7. Kluckner, P.M., Weiss, A., Sundström, P., Tscheligi, M.: Two Actors: Providers and Consumers inform the Design of an Ambient Energy Saving Display with Persuasive

Strategies. First Int. Conf. on Behavior Change Support Systems (BCSS), Adjunct PERSUASIVE'13, Position paper, 33-44 (2013).

8. Kluckner, P.M.: Improving long-term Persuasion for Energy Consumption Behavior: User-centered Development of an Ambient Persuasive Display for private Households. In Adjunct PERSUASIVE'13, Ext. Abstracts (2013).
9. Kukka, H., Oja, H. Kostakos, V. Goncalves, J., Ojala, T.: What makes you click: exploring visual signals to entice interaction on public displays. In Proc. CHI'13, 1699-1708 (2013).
10. Massung, E., Coyle, D., Carter, K.Jay, M., Preist, C.: Using Crowdsourcing to support Pro-environmental Community Activism. In Proc. CHI EA'13, 371-380 (2013).
11. Mattern, F., Staake, T., Weiss, M.: ICT for green: how computers can help us to conserve energy. In e-Energy'10: Int. Conf. on Energy-Efficient Computing and Networking, Interdisciplinary Quarterly, 16 (1) 98-105 (2010).
12. Midden, C., McCalley, T., Ham, J., Zaalberg, R.: Using persuasive technology to encourage sustainable behavior. Sustainability WS PervasivE'08
13. Milenkovic, M., Hanebutte, U., Huang, Y., Predergast, D., Pham H.: Improving user comfort and office energy efficiency with POEM (personal office energy monitor). In Proc. CHI EA'13, 1455-1460 (2013).
14. Neustaeder, C., Bartram, L., Mah, A.: Everyday Activities and Energy Consumption: How Families Understand the Relationship. In Proc. CHI'13, 1183-1192 (2013).
15. Oinas-Kukkonen H.: A foundation for the study of behavior change support systems. In J Personal and Ubiquitous Computing (2012)
16. Pasch, M. and Landoni, M. Longitudinal assessment of a user experience evaluation tool for children. In Proc. CHI EA'12.
17. Peeters, M., Megens, C., van den Hoven, E., Hummels, C., Brombacher, A.: Social Stairs: Taking the Piano Staircase towards Long-Term Behavioral Change. In Proc. PERSUASIVE'13, 174-179 (2013).
18. Petkov, P., Köbler, F., Goswami, S., Krcmar, H.: Personalised Eco-Feedback as A Design Technique for Motivating Energy Saving Behaviour at Home. In Proc. NordiCHI'12.
19. Pierce, J., Paulos, E.: Beyond energy monitors: Interaction, energy, and emerging energy systems. In Proc. CHI'12, ACM Press, 665-674 (2012).
20. Reitberger, W., Kastenmiller, M., Fitzpatrick, G.: Invisible Work: An Ambient System for Awareness and Reflection of Household Tasks. In Proc. PERSUASIVE'13, 180-191.
21. Riche, Y., Dodge, J., Metoyer, R.A.: Studying always-on electricity feedback in the home. In Proc. CHI'10, ACM, NY, USA, 1995-1998 (2010).
22. Schrammel, J., Gerdensitsch, C., Weiss, A., Kluckner, P.M., Tscheligi, M.: FORE-Watch - The clock that tells you when to use: Persuading users to align their energy consumption with green energy availability. In Proc. AMI'11, 157-166 (2011).
23. Shneiderman, B., Plaisant, C.: Strategies for Evaluating Information Visualization Tools: Multi-dimensional In-depth Long-term Case Studies. In WS BELIV'06.
24. Smeaton, A.F., Doherty, A.R.: Persuading Consumers to Reduce their Consumption of Electricity in the Home. In Proc. PERSUASIVE'13, 204-215 (2013).
25. Staats, H., Harland, P., Wilke, H.A.M.: Effecting Durable Change. In Environment and Behavior, 36 (3), 341-367 (2004).
26. Strengers, Y.A.A.: Designing eco-feedback systems for everyday life. In Proc. CHI'12
27. Strengers, Y.A.A.: Smart metering demand management programs: challenging the comfort and cleanliness habitus of households. In Proc. OZCHI'08, 9-16 (2008).
28. Yun, T.-J.: Investigating the impact of a minimalist in-home energy consumption display. In Proc. CHI EA'09, 4417-4422 (2009).