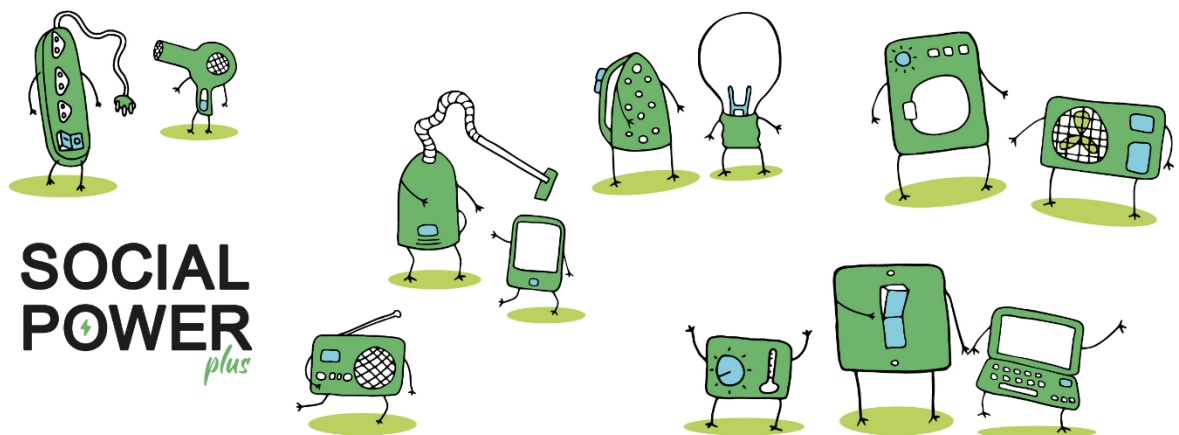




Final report dated 30 November 2023

SOCPOP - Social Power Plus

Empowering energy sufficiency and energy citizenship in households through co-designed app-based Community Energy Challenges



Source: ©SUPSI2020



University of Applied Sciences and Arts
of Southern Switzerland

SUPSI



Date: 30 November 2023

Location: Bern

Publisher:

Swiss Federal Office of Energy SFOE
Energy Research and Cleantech
CH-3003 Bern
www.bfe.admin.ch

Co-financing:

Stadtwerk Winterthur
Winterthur
www.stadtwerk.winterthur.ch

EKS AG
Schaffhausen
www.eks.ch

Technische Betriebe Wil
Will
www.tb-wil.ch

Subsidy recipients:

SUPSI
via F. Ruchat-Roncati 15, CH-6850 Mendrisio
www.supsi.ch

ZHAW
Technoparkstrasse 2, CH-8401 Winterthur
www.zhaw.ch

CLEMAP
Lavaterstrasse 66, CH-8002 Zürich
www.clemap.ch

Authors:

Francesca Cellina, SUPSI, francesca.cellina@supsi.ch
Tiziano Gerosa, SUPSI, tiziano.gerosa@supsi.ch
Pasquale Granato, SUPSI, pasquale.granato@supsi.ch
Evelyn Lobsiger, ZHAW, evelyn.lobisger-kaegi@zhaw.ch
Devon Wemyss, ZHAW, devon.wemyss@zhaw.ch
Pascal Kienast, CLEMAP, pascal@clemap.ch
Gino Abgomemewa, CLEMAP, gino@clemap.ch
Giovanni Profeta, SUPSI, giovanni.profeta@supsi.ch



SFOE project coordinators:

Nadège Vetterli, nadege.vetterli@anex.ch

SFOE contract number: SI/502062-01

The authors bear the entire responsibility for the content of this report and for the conclusions drawn therefrom.



Zusammenfassung

Individuelles und kollektives Handeln ist ein wichtiger Faktor für den Energiebedarf für Heizzwecke und andere Aktivitäten im Haushalt. Im Rahmen von Social Power Plus (SPP) haben wir Haushalte dazu motiviert, aktiv Suffizienz- und Effizienz-Maßnahmen umzusetzen, die auf ihren Strom- und Heizenergieverbrauch abzielen. Zunächst haben wir in einem Co-Design-Prozess die Community Energy Challenge und die dazugehörige App für nachhaltigen Energieverbrauch im Haushalt entwickelt. An diesem Prozess nahmen 54 Haushalte der Regionen Schaffhausen, Winterthur und Wil (Kunden der drei Energieversorgungsunternehmen EKS, SWW und TBW) teil. Die Einbindung der Teilnehmenden von Beginn des Designprozesses an ermöglichte es uns, theoriebasierte Erkenntnisse für wirksame Maßnahmen zur Verhaltensänderung mit den Bedürfnissen und Erwartungen der potenziellen Nutzenden zu ergänzen. Die daraus resultierende Community Energy Challenge macht den Energieverbrauch der täglichen Haushaltsroutinen bewusst. Die Interaktion zwischen den Teilnehmenden wurde durch die App-basierten Challenges zu diesen Haushaltsroutinen sowie durch die Pinnwand, die den direkten Erfahrungsaustausch zwischen den Teilnehmenden ermöglicht, gefördert. Die Annahme war, dass soziale Interaktionen, vermittelt durch die App und unterstützt durch Zielsetzungen, Challenges, Tipps und Verbrauchsrückmeldungen, zu einer bewussten Auseinandersetzung mit der Art und Weise, wie gewisse Aktivitäten im Haushalt ausgeführt werden, führen und so einen langfristigen Energiespareffekt erzielen.

Die SPP Community Energy Challenge wurde 2022 in den drei Regionen der teilnehmenden EVU's in einem quasi-experimentellen Umfeld getestet, an dem 221 Haushalte (die Versuchsgruppe) freiwillig teilnahmen. Die teilnehmenden Haushalte waren überwiegend männlich (73 %), mittleren Alters (durchschnittlich 53 Jahre) und hochgebildet (mehr als die Hälfte hat einen Hochschulabschluss). Die Intervention wurde in zwei Phasen organisiert: Während der ersten drei "intensiven" Monate (Februar bis April 2022) animierten die App-Funktionen zu einer hohen Interaktion zwischen den App-Nutzenden und zur Beteiligung an den App-basierten Challenges. Danach blieb die App für den Rest des Jahres (Mai - Dezember 2022) aktiv, um das Beobachten des eigenen Energieverbrauchs im Haushalt weiterhin zu ermöglichen. Die Teilnehmenden konnten ihre Erfahrungen weiterhin über die Pinnwand austauschen, es wurden jedoch keine neuen Challenges lanciert.

Um den Effekt der Teilnahme an der SPP Community Energy Challenge auf den Energieverbrauch zu ermitteln, wählten wir einen quasi-experimentellen Ansatz, indem wir eine vergleichbare Gruppe von Haushalten unter den Privatkunden desselben Versorgungsunternehmens (Kontrollgruppe) identifizierten und einen "Differenz-von-Differenzen"- Ansatz (DID) verwendeten. Unsere Ergebnisse zeigen, dass die Teilnahme an der SPP Community Energy Challenge im Durchschnitt zu Energieeinsparungen von fast 5 % führte, berechnet über einen Zeitraum von 11 Monaten nach dem Start der Intervention. Die Schätzung des Effekts über den DID-Schätzer berücksichtigt bereits die zeitliche Veränderung der Heizgradtage, die sich auf die Nachfrage nach Heizenergie auswirken. In den ersten Monaten nach dem Start der SPP Community Energy Challenge war der Effekt sogar noch höher, im Durchschnitt etwa 11 %. Der Effekt nahm also im Laufe der Zeit ab, was wahrscheinlich daran liegt, dass die Neuartigkeit und damit die Attraktivität der App mit der Zeit abnahm und es in der Phase, in der die App weniger intensiv bespielt wurde, an Anreizen und Erinnerungen mangelte. Die Auswirkungen der Teilnahme auf die Energieeinsparungen waren in allen drei Regionen und ihren unterschiedlichen Energieinfrastrukturen gleich.

Das selbstberichtete Verhalten in den Umfragen unter den Teilnehmenden deutet darauf hin, dass einige Haushaltsroutinen energiesuffizienter gestaltet wurden, insbesondere wurde eine Verringerung der Innenraumtemperatur berichtet. Dies steht teilweise im Gegensatz zu früheren Forschungsergebnissen, in denen argumentiert wurde, dass die Dekarbonisierung von Heizungssystemen eher durch energieeffiziente Maßnahmen (d. h. durch die Bereitstellung des von den Menschen erwarteten Komforts) als durch Energiesuffizienzlösungen (d. h. durch eine freiwillige Reduzierung des Energieverbrauchs) erreicht werden kann. Die Erfahrungen im SPP Projekt deuten



jedoch auch darauf hin, dass es ein Potenzial für den Austausch von technischem Wissen und Kompetenzen im Bereich der Energieeffizienz gibt, das explizit für Nachhaltigkeitsprozesse genutzt werden könnte, z. B. für persönliche Peer-to-Peer-Beratungsprozesse - entweder vermittelt durch eine App oder durch persönliche Interaktionen in Living-Lab-Prozessen. Andere Routinen wie Duschen und Wäschewaschen scheinen sich dagegen nicht verändert zu haben. Dies könnte darauf hindeuten, wie schwierig es ist, Normen und Konventionen in Bezug auf Sauberkeit zu verändern. Die Antworten in den Umfragen deuten darauf hin, dass das individuelle Feedback zum Energieverbrauch und die Zielsetzungsfunktion eine wichtige Rolle bei der Erzielung der Spareffekte gespielt haben. Die meisten Nutzenden wünschten sich sogar detailliertere Informationen über den Energieverbrauch, die jedoch unter den derzeitigen technischen und rechtlichen Rahmenbedingungen kaum bereitgestellt werden können. Im Rahmen der SPP Community Energy Challenge erwies sich der Versuch, durch Erfahrungsaustausch unter den Teilnehmenden eine Verhaltensänderung anzustossen, als weniger erfolgreich. Zukünftige Forschungsprojekte könnten versuchen, die Verbrauchsfeedback-App weniger in den Vordergrund zu stellen und sich stattdessen explizit auf das kollektive Hinterfragen aktueller Routinen zu konzentrieren, wie es typischerweise in Living-Lab-Ansätzen geschieht. Wir gehen davon aus, dass sich mittels solcher Ansätze auch «Nicht-Resource-Men-Haushalte» einzubeziehen lassen, da das quantitative Feedback der App nicht den Kern der Intervention darstellen würde. Die Einbettung der App-Nutzung in solche, umfassenderen Prozesse könnte Veränderungen in den Haushaltsroutinen unterstützen und dazu beitragen, dass die Energieeinsparungen langfristig erhalten bleiben.

Abschließend ist anzumerken, dass die begrenzte Stichprobengröße es uns nicht erlaubte, komplexere Interventionsdesigns zu entwickeln, die beispielsweise darauf abzielen, den Energiespareffekt bei Variation der soziodemografischen Merkmale der Haushalte abzuschätzen: Dies bleibt der zukünftigen Forschung vorbehalten. Darüber hinaus könnten weitere Forschungsarbeiten darauf abzielen, die Dynamik von Haushaltsroutinen und -verhalten zu untersuchen, um zu klären, wie verschiedene Personen auf die Teilnahme an ähnlichen App-basierten Community Energy Challenges reagieren, indem beispielsweise die Rolle von Alter, Geschlecht oder Bildung untersucht wird und wie sich diese auf das Energieverhaltensverhalten im Haushalt auswirken.

Riassunto

L'azione individuale e collettiva è un importante motore del consumo di energia che nelle economie domestiche viene effettuato sia per scopi di riscaldamento sia per altri scopi. In Social Power Plus (SPP) abbiamo coinvolto le economie domestiche nell'attuazione di misure attive di sufficienza ed efficienza energetica per ridurre il consumo di energia termica ed elettrica. Abbiamo avviato un processo di co-progettazione volto a sviluppare una "Sfida Energetica di Comunità" e un'app per smartphone che sfruttano contatori e sensori intelligenti e algoritmi non intrusivi per il monitoraggio del carico elettrico, sviluppati appositamente per il progetto. Il processo ha coinvolto 54 economie domestiche nelle regioni di Sciaffusa, Winterthur e Wil, presso le quali sono attive rispettivamente le aziende energetiche EKS, SWW e TBW. Il coinvolgimento delle economie domestiche fin dall'inizio del processo ha permesso di integrare elementi basati sulla teoria comportamentale con le esigenze e le aspettative dei potenziali utenti. L'intervento e app che ne risultati consentono di identificare le routine domestiche quotidiane che consumano energia e promuove l'interazione sociale tra i suoi utenti, attraverso sfide che affrontano direttamente tali routine. Una bacheca all'interno dell'app ("pinboard") facilita inoltre lo scambio diretto di esperienze tra i partecipanti, nell'ipotesi che le interazioni sociali, mediate dall'app e supportate dalla definizione di obiettivi, sfide, suggerimenti e feedback sui consumi energetici, potessero portare a ripensare il modo in cui in casa si consuma energia, per un effetto di risparmio energetico a lungo termine.

La "Sfida Energetica di Comunità SPP" è stata testata nel 2022 nelle stesse tre regioni, mediante un disegno di ricerca quasi sperimentale che ha coinvolto 221 famiglie auto-selezionate (gruppo di trattamento). I rappresentanti di tali economie domestiche erano in prevalenza uomini (73%), di mezza



età (età media pari a 53 anni) e di elevato livello di educazione (più della metà possedeva una laurea). L'intervento è stato organizzato in due fasi: durante i primi tre mesi, le funzionalità dell'app hanno stimolato un'elevata interazione tra gli utenti e un elevato coinvolgimento in attività di risparmio energetico e sfide sulle routine domestiche. Successivamente, l'app SPP è rimasta attiva per l'auto-monitoraggio del consumo energetico dei partecipanti e per la condivisione delle esperienze tramite la pinboard, mentre non sono state rilasciate nuove sfide. Per stimare l'effetto medio sui consumi energetici prodotto dalla partecipazione alla "Sfida Energetica di Comunità SPP", abbiamo adottato un approccio quasi-sperimentale, individuando un gruppo comparabile di famiglie (gruppo di controllo) tra i clienti domestici delle stesse aziende energetiche e utilizzando lo stimatore "differenza-nelle-differenze" (DID). I risultati indicano che in media la partecipazione alla "Sfida Energetica di Comunità SPP" ha prodotto quasi il 5% di risparmio di energia, conteggiato su un periodo di 11 mesi dall'avvio e considerando la domanda totale di energia, per usi termici e non termici. Si noti che la stima dell'effetto attraverso lo stimatore "DID" tiene conto dell'evoluzione nel tempo nelle temperature esterne, che influisce sulla domanda di energia per scopi termici. Nei primi mesi, l'effetto è stato ancora più elevato, mediamente pari a circa il 11% di risparmio; successivamente è progressivamente diminuito nel tempo, probabilmente a causa della diminuzione dell'effetto novità e della riduzione degli stimoli offerti dall'app stessa. L'effetto della partecipazione alla "Sfida Energetica di Comunità SPP" è stato comparabile nelle tre regioni, caratterizzate da diverse infrastrutture energetiche.

Elementi auto-risportati che abbiamo raccolto tramite sondaggio indicano inoltre che alcune routine domestiche sono diventate più sufficienti dal punto di vista energetico, in particolare quelle relative alle impostazioni del termostato per il riscaldamento. Questo risultato è in parziale contrasto con ricerche precedenti, che hanno concluso che la decarbonizzazione dei sistemi di riscaldamento può essere ottenuta con maggior facilità attraverso misure di efficienza energetica (cioè garantendo agli individui il livello di comfort che si aspettano di avere), invece che di sufficienza energetica (cioè appoggiandosi a misure volontarie di riduzione dei consumi). Si noti comunque che elementi dell'esperienza di SPP suggeriscono anche che ci sia un potenziale per la condivisione di conoscenze e competenze tecniche su tema dell'efficienza energetica. Questo potenziale può essere esplicitamente sfruttato nell'ambito di processi di transizione verso la sostenibilità, ad esempio mediante attività di consulenza tra pari, mediate da una app o da interazioni e scambi di persona.

Altre routine legate alla doccia e al bucato, invece, non sono cambiate. Ciò può riferirsi alla difficoltà di spostare le convenzioni culturali su argomenti legati alla pulizia. Infine, le risposte al sondaggio suggeriscono che il feedback individuale sul consumo energetico e la possibilità di definire obiettivi personalizzati hanno svolto un ruolo importante nel guidare l'effetto. La maggior parte degli utenti era addirittura desiderosa di informazioni più dettagliate sul feedback sul consumo energetico, che tuttavia sono difficili da fornire nell'attuale quadro normativo e tecnologico. Per come è stata impostata la "Sfida Energetica di Comunità SPP", il tentativo di favorire il cambiamento attraverso la condivisione di esperienze si è rivelato meno vincente. Future ricerche potrebbero provare a dare meno importanza alla presenza e al ruolo delle app, concentrandosi invece esplicitamente sulla messa in discussione collettiva delle routine attuali, come tipicamente si fa nei "living lab". Ci aspettiamo che questi approcci favoriscano anche il coinvolgimento di soggetti diversi dagli "uomini risorsa" ("resource men") concettualizzati nella precedente letteratura. Infatti, l'app e il feedback numerico da essa fornito non sarebbe al centro dell'intervento. Inserire l'uso delle app in processi più ampi potrebbe favorire il cambiamento in tutte le routine domestiche, favorendo anche il mantenimento del risparmio energetico nel lungo termine.

In chiusura, segnaliamo che la dimensione limitata del campione non ci ha consentito di attuare più complessi disegni di ricerca, volti ad esempio a stimare l'effetto di risparmio energetico al variare delle caratteristiche socio-demografiche delle economie domestiche. Questo elementi potrebbero essere stimati nell'ambito di futuri progetti di ricerca. Inoltre, future ricerche potrebbero esplorare in modo specifico le dinamiche delle routine di consumo energetico nelle economie domestiche, con l'obiettivo di chiarire anche come differenti categorie di individui reagiscano alla partecipazione a simili "Sfide



Energetiche di Comunità” basate su app, ad esempio considerando l'età, il genere, il livello di educazione, per comprendere come questi fattori influiscano sulle routine che consumano energia nell'economia domestica.

Summary

Individual and collective action is an important driver of energy consumption in households, for both heating and non-heating purposes. In Social Power Plus (SPP) we engaged households towards active energy sufficiency and efficiency measures targeting their electricity and heating consumption. We first launched a co-design process aimed at developing a Community Energy Challenge and mobile app for sustainable energy consumption at home. The process involved 54 household participants in the regions of Schaffhausen, Winterthur, and Wil, where the EKS, SWW, and TBW utilities are respectively active. The involvement of participants from the start of the design process allowed us to complement theory-based insights for effective behaviour change interventions with the needs and expectations of potential users. The resulting Community Energy Challenge was designed to highlight daily energy-consuming household routines and promote social interaction between participants via app-based challenges directly addressing eight energy consuming household routines and pinboard features to facilitate the direct exchange of experiences between participants. The assumption was that social interactions, mediated by the app and supported by goal-setting, challenges, tips, and consumption feedback features, could lead to rethinking the way the routines are performed in the household, for a long-lasting energy saving effect.

The SPP Community Energy Challenge was tested in 2022 in the same three regions by means of a quasi-experimental setting that engaged 221 self-selected households (the treatment group). Applicants representing such households were mostly males (73%), middle-aged (53 years old on average) and highly educated (more than half has a university degree). The intervention was organised in two stages: during the first three “intensive” months (Feb – April 2022), app features stimulated high interaction between app users and engagement with energy saving activities in app-based challenges. Then the app remained active for continued self-monitoring of the participants' home energy consumption for the remainder of the year (May – December 2022). Participants could continue to share experiences via the app-internal pinboard, however no new challenges were released.

To estimate the average effect on energy consumption produced by participation in the SPP Community Energy Challenge, we adopted a quasi-experimental approach, by identifying a comparable group of households (control group) among the household customers of the same utilities and using a “difference-in-differences” (DID) estimator. Our findings indicate that, on average, participation in the SPP Community Energy Challenge produced nearly 5% savings of overall energy consumption (heating and non-heating purposes), computed over a period of 11 months from the launch of the challenge. The estimate of the effect via the DID estimator already accounts for the change over time in external temperature conditions, which affect the demand for energy for heating purposes. In the first months after the launch of the SPP Community Energy Challenge, the effect was even higher, on average about 11%; however it gradually decreased over time, probably due to the decrease of the novelty of the app and the lack of stimuli during the lower intensity period of app availability. The effect of participation on energy savings was consistent across the three regions and their different energy infrastructures.

Self-reported survey data indicate that some household routines became more energy sufficient, in particular a decrease in thermostat settings for heating. This is partially in contrast with previous research, that argued that the decarbonisation of heating systems can more easily be obtained via energy efficient measures (i.e. by providing individuals with the comfort level they expect), rather than energy sufficiency solutions (i.e. by relying on voluntary reductions of energy consumption). Elements from the SPP experience however also suggest that there is a potential for technical knowledge and competence sharing around energy efficiency, that might be explicitly leveraged for sustainability transition processes, for instance for personal peer-to-peer consultancy processes -either mediated by an app or by in-person interactions.



Other routines around showering and laundry, instead, appear not to have changed. This may hint at the difficulty of shifting cultural conventions on cleanliness-related topics. Finally, survey responses suggest that individual feedback on energy consumption and the goal setting feature have played a major role in driving the effect. Most users were even eager for more detailed information on energy consumption feedback, which however can hardly be provided under the current technology and regulatory framework. Within the set-up of the SPP Community Energy Challenge, the attempt to foster change through sharing experiences proved to be less successful. Future research might attempt to give less prominence to the presence and role of the app and rather to explicitly focus on collectively questioning and challenging current routines, as typically performed in living lab approaches. We expect that such approaches would allow to also engage “non resource men” which were conceptualised in previous literature. In fact, the app and its quantitative feedback would not be the core of the intervention. Framing app use within such broader processes might support changes in household routines and help to maintain the energy savings in the long-term.

We finally remark that the limited sample size did not allow us to develop more complex intervention designs, for instance aimed at estimating the energy saving effect on varying the socio-demographic characteristics of the households: these are left for future research. Furthermore, additional research might aim at specifically exploring the dynamics of household routines and behaviour, in order to also clarify how different categories of individuals react to participation to similar app-based Community Energy Challenges, for instance by considering the role of age, gender, or education, and how these factors affect performance of energy consuming routines within the household.

Main findings

- We co-designed and field tested an app-based “community energy challenge” to foster households’ energy efficient and sufficient behaviour with 200 households in three regions in Eastern Switzerland.
- Using a quasi-experimental research design, we found an average treatment effect equal to nearly 5% savings in the overall consumption of energy (heating and non-heating purposes) –up to nearly 11% in the first months immediately after the launch of the community energy challenge. Individual feedback on energy consumption and the goal setting features have likely played a major role in driving this effect.
- Self-reported survey data suggest that households changed their thermostat-setting routines to be more energy sufficient but maintained their showering and laundry routines. This hints at the difficulty of shifting cultural conventions on cleanliness-related topics.
- To promote changes in household routines and support long-term maintenance of the energy savings, future research should attempt to give less prominence to the energy data and role of the app, and rather explicitly focus on collectively questioning and challenging current routines, as typically performed in living lab approaches.



Contents

1	Introduction	11
1.1	Background information and current situation	11
1.2	Purpose of the project.....	11
1.3	Objectives	12
2	Procedures and methodology	12
3	Co-design of the SPP app and Community Energy Challenge	14
3.1	Theoretical framework	14
3.2	Integrating co-design into theory-informed intervention design	16
3.3	Co-design methodology	17
3.4	Implementation of the SPP co-design process	18
3.5	Co-design outputs.....	21
3.6	The SPP Community Energy Challenge and Toolbox.....	25
3.7	Discussion.....	31
4	Analysis of the effects of the SPP Community Energy Challenge.....	34
4.1	Methodology	34
4.2	Surveys and in-app automatic data collection	37
4.3	Recruitment of the participants (treatment group households).....	39
4.4	Definition of the DID analytic sample	40
4.5	Heating and cooling degree-days	46
4.6	Impact evaluation.....	48
4.7	Estimate of treatment effects	51
4.8	Discussion.....	54
4.9	Recommendations for future app-based interventions	69
5	Conclusions	74
6	Outlook and next steps	76
7	National and international cooperation	77
8	Communication.....	77
9	Publications.....	79
10	References.....	79



Abbreviations

List of used abbreviations:

ATT	Average treatment effect on the treated
CATT	Conditional average treatment effect on the treated
CDD	Cooling degree-days
DID	Difference-in-differences
EKS	Elektrizitätswerk des Kantons Schaffhausen
HDD	Heating degree-days
ICT	Information and communication technologies
MAP	Model of action phases
NN	Nearest neighbor
OLS	Ordinary least square
PV	Photovoltaic
PSM	Propensity score matching
SPP	Social Power Plus
SWW	Stadtwerk Winterthur
TBW	Technische Betriebe Will
TDD	Total degree-days
TPB	Theory of planned behaviour
TWFE	Two ways fixed effect
WP	Working package



1 Introduction

1.1 Background information and current situation

Households are responsible for 31.4% of total energy consumption (Bundesamt für Energie, 2017) and are therefore an important intervention point for the Swiss energy transition. Yet, in the context of a strongly centralised energy supply system, until recently households have been considered passive market actors and marginal recipients of technology. Achieving the goals of the Energy Strategy 2050, however, calls upon households to adjust daily energy usage habits towards sufficiency and to invest in energy-efficient household equipment. In such a context, addressing the human factors underlying energy-relevant decisions and behaviour and the socio-cultural aspects that favour active energy citizenship, becomes essential.

Research is increasingly focusing on better understanding the underlying mechanisms (so-called drivers) that influence human behaviour change and energy savings, and their different effects (Delmas et al., 2013; Sovacool, 2014) and intervention approaches (Frederiks et al., 2015; Karlin et al., 2015). Thanks to recent progress in ICTs, with smart metering roll-outs by utility companies, ease of installation of sensors, and the wide availability and diffusion of smartphones by the consumers, energy-saving interventions are increasingly performed by means of applications (apps) for internet-enabled devices. This uniquely allows for customized, (nearly) real-time energy feedback and possibilities of interaction with and between the users.

Additionally, a significant reduction in energy consumption has been observed in socially embedded interventions (Breukers and Mourik, 2013). In particular, also from the behaviour change perspective, there is a growing tendency to approach consumers no longer as individual agents for change, but rather as socially situated individuals that are part of a wider community (Mengolini et al. 2016). Here, social norms, collective actions and community welfare become important drivers: informative contents typically involve social comparison as an engagement trigger, i.e. they report a household's energy use compared to that of similar neighbours, as well as playful challenges, competitions, collaborations (e.g. self-consumption communities), community rewards and events. In order to avoid rebound effects, social comparison and descriptive social norms can be accompanied by the activation of injunctive messages coupled with energy-saving tips (Cialdini et al. 1990; Schultz et al., 2007; Allcott, 2011).

1.2 Purpose of the project

Using the insights of the latest research, stated above, new behavioural intervention design should explore effectiveness of community-oriented strategies, which adopt a participatory approach to the question of energy-related behaviours with the aim of building a sense of community, shared goals and values. Smart tools exploiting the ongoing large-scale smart meter roll-out, that are aimed at favouring awareness on individual and community-level energy demand, as well as at stimulating an overall demand reduction and optimisation of production and consumption patterns, could support cohesion and enhance feelings of membership to a community. Next to triggering more effectively energy sufficiency behaviour within the residential building sector, they may contribute to the empowerment of energy citizens in general, thus laying some possible groundwork also with respect to the future configuration of the electricity grid into a system of self-consumption communities.

Against this backdrop, the Social Power Plus (SPP) project developed and tested a behaviour change app persuading a reduction in household energy consumptions, by directly engaging its target users in a living lab framework. For this purpose, SPP activated a co-design, participatory process to design features of the app and accompanying "Community Energy Challenge". The chosen features were designed into the SPP app and tested in a quasi-experimental setting in three Swiss communities. Quantitative and qualitative analyses on the impact of the SPP Community Energy Challenge provide insights on the SPP overall energy saving effectiveness. The lessons learnt from the whole SPP process



allow us to draw recommendations on how to support the energy transition in households by leveraging ICT-based technologies such as smart meters, sensors, and persuasive apps.

Overall, the Social Power Plus project complements the SFOE project “Living the 2000 Watt Society: Reduction of Energy Consumption through Behaviour Change” (Sandmeier et al. 2020), that designed and tested behaviour change interventions in the field of mobility, warm water, space heating and electricity consumption in the Hüttengraben area in Küsnacht, with the aim of achieving a 2000 Watt society within their building projects. Social Power Plus complements this project by introducing an app-based intervention, specific behavioural suggestions through the Community Energy Challenge, and change in energy consumption habits stemming from persuasive features.

1.3 Objectives

In previous research, our team developed an app-based energy savings challenge and app, called Social Power (Wemyss et al., 2018), that allowed households to monitor their electricity consumption in real time through a gamified, lay-person visualisation, which connected actions to energy use, without the need for a more complex understanding of the energy system (Hermann et al., 2018; Marek et al., 2020). Households were placed in teams, within which they were invited to collaborate to collectively save a given amount of energy or to save more energy than a rival team (compared to their historical average consumption). While the real-world test of the app-based energy savings challenge successfully resulted in approx. 8% electricity savings in two Swiss cities, the savings were not maintained one year after the intervention ended – at least, in terms of statistical significance. Furthermore, high drop-out rates were observed during the intervention (Wemyss et al., 2019).

We hypothesize that such drop-out and relapse to previous behaviour are due to a lack of explicit incorporation of user knowledge, practices, and preferences, into the design of the Social Power challenge. SPP therefore specifically aims at exploring such hypothesis, by overcoming the expert-based approach and actively engaging potential target users in the design of the behaviour change intervention itself (the Community Energy Challenge), in a “living lab” approach.

Furthermore, SPP also aims at improving the scope and quality of the feedback offered to households in terms of their energy consumption, by including energy consumption for heating purposes, which represent the largest share of household energy consumption. This supports developing the added-value from the current smart meter roll-out for the demand side. For this purpose, the SPP project developed an app to provide households with feedback on their own disaggregated energy consumption data, at least in terms of heating and non-heating purposes.

The SPP app (including supporting ICT sensors) and the SPP Community Energy Challenge have been designed and the related software and hardware tools have been deployed (which we overall refer to as the “SPP Toolbox”) are tested for their impact on energy consumption. This is done via a quasi-experimental approach, by experimentation in real-life conditions that involves a group of SPP treated households and a matched comparable group of untreated households.

The SPP project has therefore two key objectives, summarized as follows:

- Objective 1. Co-design the SPP app and Community Energy Challenge with potential users and stakeholders;
- Objective 2. Assess the energy savings produced by use of the SPP app and participation in the Community Energy Challenge, in comparison to control groups.

2 Procedures and methodology

Activities are performed in three different regions in the German-speaking part of Switzerland, integrating three different energy providers (utility companies and project co-funders) and the available



technology and energy consumption metering infrastructure, specifically the regions of Wil in Canton St. Gallen with Technische Betriebe Wil (TBW) as partner utility, Winterthur in Canton Zurich with Stadtwerk Winterthur (SWW) as partner utility, and Schaffhausen in Canton Schaffhausen with Elektrizitätswerk des Kantons Schaffhausen AG (EKS) as partner utility.

Having three parallel implementations in the three regions allows to assess the effectiveness and real-life feasibility of the SPP Community Energy Challenge under different socio-technical contexts, thus providing insights also on the possibilities for exploitation at a larger scale. The SPP app and Community Energy Challenge are thus sufficiently flexible to adapt to different technological set-ups of a local energy system, particularly accounting for the electricity and gas grids.

Project activities are organised around six work packages (WPs), that are schematically represented in Figure 1. Project activities start with WP2, whose focus is the creation of the Living Lab process to co-design the app-based SPP app and Community Energy Challenge with their potential target users. The SPP app is fed by data from gateways able to collect energy consumption data from gas and/or electricity smart meters or sensors, that are developed in WP3 by project partner CLEMAP. Also in WP3, CLEMAP develops non-intrusive load monitoring algorithms (NILM) aimed at disaggregating heating energy consumption from non-heating energy consumption data. The resulting “SPP Digital Toolbox”, which is the output of WP3, thus includes the SPP app, the NIALM algorithms, and the sensors to collect energy consumption data where no smart meters are available. The toolbox is sufficiently flexible as to manage energy consumption data produced in different technological settings and under different metering infrastructures.

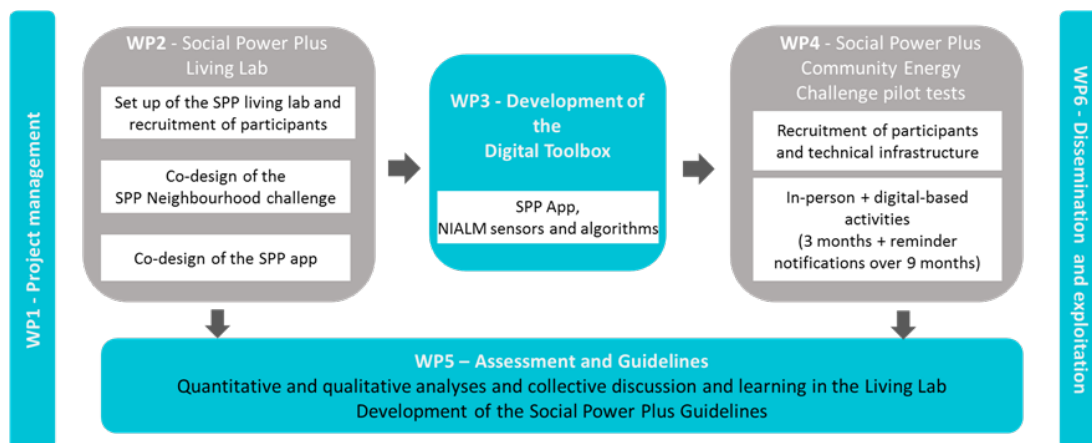


Figure 1: The organisation of project activities in Work Packages (WPs).

The effectiveness of the SPP app and Community Energy Challenge is tested in three quasi-experimental, one year-long real-life pilots involving around 200 voluntary households in the three regions (WP4). The impact on energy consumption, reported energy consuming habits and routines, as well as motivation, social norms, and perceived behavioural control is estimated by means of both energy consumption data provided by smart meters and sensors and surveys administered to the participating households before, after 3 months (end of the SPP Community Energy Challenge) and 9 months later. We also assess the households' perceptions on the level of engagement produced by the different features of the SPP app and Community Energy Challenge, by means of both quantitative and qualitative analysis of survey items (WP5).

The whole process offers opportunities for learning from the real-life experimentation from the perspective of the different actors that are involved in project activities: the households, the utilities, the technology development company deploying the consumption breakdown algorithms and sensors (CLEMAP), and the ZHAW and SUPSI research team running co-design activities and the quasi-experimental estimates of the treatment effect. Analysis of materials collected during the interaction with participating households and final interviews and workshops with all such actors allow to draw



recommendations for future activities aimed at supporting the energy transition by leveraging smart meters, energy consumption data, and persuasive app-based interventions (WP5).

Throughout the project, active dissemination activities (WP6) targeting both the general public, practitioners and the academic community are performed, by means of conference presentations, online blog posts, and manuscripts submitted to scientific journals.

The following chapters focus on the two main SPP objectives: the co-design of the SPP app, Toolbox and Community Energy Challenge (Chapter 3), and the assessment of their effectiveness via a quasi-experimental research intervention (Chapter 4). These chapters describe the specific methodologies we used to achieve each objective, present the results we obtained, and discuss them in light of our research hypotheses and the relevant scientific literature. In Chapter 5 we provide our recommendations for future app-based activities aimed at fostering energy saving in households, from the perspective of both researchers and practitioners, while in Chapter 6 we provide our conclusions about the whole experience within the SPP project. Finally, Chapter 7 shortly reports on the collaborations we have activated during the SPP project, and Chapter 8 provides a summary of the dissemination activities we performed during the project.

3 Co-design of the SPP app and Community Energy Challenge

The co-design of the SPP app and Community Energy Challenge was informed by a literature review, grounding the app features in behavioural theories and ensuring that co-design activities properly integrate and expand on -rather than repeat- evidence from previous empirical research. In the next sections we summarise key activities and outputs we performed to achieve project Objective 1.

3.1 Theoretical framework

In the last decade, smartphone apps have grown in use, both in scientific and commercial contexts, to encourage more sustainable energy consumption through novel, timely and personalised information feedback (Baptista and Oliveira, 2019; Chatzigeorgiou and Andreou, 2021; Wood et al., 2014). For example, there are apps which estimate energy consumed when showering by measuring water volume and temperature using a Bluetooth-enabled in-line shower meter (Tiefenbeck et al., 2019) or provide energy savings recommendations and incentives based on real-time feedback from household energy consumption, measured by smart meter (Fraternali et al., 2019). Overall, meta-reviews of intervention studies have found smartphone apps to be successful in facilitating behaviour change (Chatzigeorgiou and Andreou, 2021).

Broadly, smartphone apps are a medium to intervene in the user's everyday reality and leverage this engagement to provide persuasive feedback that influences user behaviours (Hekler et al., 2016). Thus, a specific theory of behaviour change can be operationalised as an intervention in an app using different design features. For energy consumption, the aim is to trigger, motivate, and maintain new patterns of energy use which last beyond the intervention period (Ohnmacht et al., 2017). Herein, the Model of Action Phases (MAP) highlights the cognitive processes before, during and after the behaviour is enacted, and thus has inspired this specific methodology for intervening in daily sustainability-relevant behaviours (Bamberg, 2013).

The MAP framework denotes behaviour in four phases: predecision, preaction, action and postaction (Figure 2). The framework associates relevant socio-psychological factors (e.g. attitudes, intentions, or social norms) which help move the behaviour forward through the phases (Ohnmacht et al., 2017). The framework combines various theories of behaviour change, with the most important being (Bamberg, 2013): 1) the activation of personal norms in the Norm Activation Model (Schwartz and Howard, 1981) to elicit reflections and feelings about one's actions and 2) increasing perceived behavioural control from



the Theory of Planned Behaviour (TPB) (Ajzen, 1991) to support aligning one's actions with one's environmental attitude.

For a behavioural intervention, the MAP framework can address a diverse group of individuals, who may be at different phases, using phase-specific factors. SPP focuses on the preaction, action, and postaction phases, and thus five socio-psychological factors are relevant to activate to progress behaviour forward (Figure 2) (Ohnmacht et al., 2017). In the preaction phase an individual is motivated to set an intention to do the new behaviour. The first factor in this phase is *perceived behavioural control* and describes the ease someone has to perform the desired behaviour. Originating from TPB, perceived behavioural control implies that someone has the required opportunities and resources, and holds a belief of self-efficacy to fulfil the desired behaviour (Ajzen, 1991). The second factor is the individual's *attitude* toward the behaviour, which necessarily should be positive to support the setting of the intention, as shown in the TPB (Ajzen, 1991). In the subsequent action phase, the behavioural intention is operationalised into a precise and contextual implementation intention which involves activating *action planning*, as well as *coping planning* when obstacles arise. Improving action and coping planning involves skill building and addressing habitual patterns (Bamberg, 2013). The final phase, postaction, necessitates increasing an individual's *ability to recover from setbacks* to continue to implement and maintain the new behaviour despite obstacles or occasional failure (Ohnmacht et al., 2017).

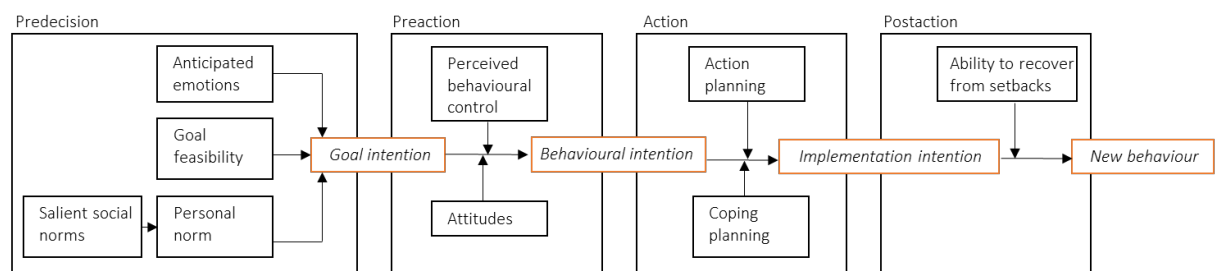


Figure 2: Conceptual schematic of the Model of Action Phases (schematic after Ohnmacht et al. (2017)).

Considering the MAP framework, three specific user-related challenges exist which impede the progression through preaction, action and postaction. At the preaction phase, the so-called intention-behaviour gap (Brown and Sovacool, 2018; Allcott and Mullainathan, 2010) is most challenging. Here, awareness of the impacts of energy consumption on climate change or even the benefits of energy savings may be present, yet the subsequent energy saving behaviour does not take place. Herein, theories purporting rational behaviour change fall short, thus interventions could consider what is particularly relevant at this phase to trigger a new behaviour (Frederiks et al., 2015). Many sociologists argue that the environmental context in which a behaviour occurs (e.g. considering the social relations, available infrastructure, institutional processes, etc.) impacts the execution of behaviour, as well as the person's ability (self-efficacy and perceived behavioural control) to change their behaviour (Wilhite et al., 2006). Further, targeting intention alone will not necessarily lead to a long-term behaviour change, and thus interventions should consider behaviour change through to the postaction phase (Bamberg, 2013).

During the action phase, when an intervention may be motivating a new behaviour, app users can lose interest and drop-out, highlighting the general problem of participant attrition in behavioural interventions (Allcott and Rogers, 2014; Georges et al., 2016). A lack of awareness regarding the potential benefits of the behaviour change can lead to disengagement (Georges et al., 2016). When these aspects are left out, an intervention may be ill-fitting or hard to engage with for the user or fail to support continuation of the behaviour and thus contribute to high intervention attrition and loss of effect.

Finally, the new sustainable behaviour must be maintained in the postaction phase, but rarely are interventions designed for this phase (Chatzigeorgiou and Andreou, 2021). Thus users can relapse back to the previous behaviour while in the postaction phase (Wemyss et al., 2019). It cannot be expected that a behaviour becomes an engrained habit during a short-time intervention and thus maintenance of the desired behaviour requires continued support after the intervention (Ohnmacht et al., 2017). Frey



and Rogers (2014) suggest pathways to support maintaining behaviour in the long-term, including providing reminders connected to environmental cues to create associations to the new behaviour or changing people's beliefs, attitudes and interpretations about the impacts of their behaviours.

Considering these challenges, realizing the potential of app-based behaviour change may require more user involvement in the design (Budgen and Stedman, 2019; Darby, 2010; Honebein et al., 2009). Thus, co-designing behavioural interventions may be a promising approach to tailor an intervention to the potential user's attitudes, motivations and realities, and thus result in greater energy saving effectiveness (Itten et al., 2021; Cockbill et al., 2020; Strengers, 2014).

Also, it may be more effective to not look at consumption patterns in isolation, but rather within the context of the daily behaviours which household members are already engaged in, which need to be derived from the household members themselves (Chatzigeorgiou and Andreou, 2021; Darby, 2020). These challenges align with the earlier call from Buchanan et al. (2015) to design for the human components of information feedback, that is considering context, motivation, knowledge, and engagement. Overall, we believe co-design approaches are well-suited to address these types of challenges.

3.2 Integrating co-design into theory-informed intervention design

We propose that co-design can be used in the design of a behavioural intervention using an abductive approach, i.e., an iterative process of exploring, developing and evaluating ideas for a problem with researchers and the potential behaviour changers (Steen, 2013). Hurley et al. (2021) provide guidance on an abductive approach to adapt the co-design process to effectively integrate research theory and experiential knowledge. Through a reflexive facilitation process, they have found that intervention design was enhanced, particularly during the initial "fuzzy" stages of design. Therefore, use this approach to develop a co-design process for behavioural interventions following the MAP.

A behaviour change intervention is defined as "a coordinated sets of activities designed to change specified behaviour patterns" (Michie et al., 2011). Thus, in this context activities are intervention techniques and design features with an app, such as social comparison or feedback on consumption impacts. Thus, the choice of the activities which induce a behaviour change can benefit from knowledge of temporal and spatial factors which might enable or impede the user.

Additionally, co-design can support designing to address existing challenges. For example, engagement with an app can be related to the frequency, novelty and relevance of the information provided. Thus, intervention design should consider how to maintain interest in consumption impacts, considering that the novelty of electricity consumption data feedback decreases with time since it is quite repetitive (Hargreaves et al., 2010). Importantly, the users involved should be strategically chosen in order to represent the full range of the target group(s) for the product or service to be designed, as opposed to aiming for representative or complete participation (Reed et al., 2018).

Co-design is increasingly found in "living labs", namely processes aimed at co-creating and validating innovation within collaborative, real-world environments (Pallot et al., 2010; Bergvall-Kåreborn et al., 2010; Dell'Era and Landoni, 2014). Such processes open up to "participatory mindsets" and help users to become active partners of the value creation process (Schuler and Namioka, 1993; Sanders, 2003): beyond "designing for the users", living labs support "designing with the users". The approach involves users during the design process (e.g. through interviews, surveys, focus groups, or pilot testing). In this way, the product is designed for its intended use and is argued to be ultimately more effective and efficient (Abrás et al., 2004).

There is no standard assessment of the value of the new design inputs resulting from the abductive approach. However, relevant criteria can be defined in advance or developed with the research team and the users as part of the co-design process (Itten et al., 2020; Cockbill et al., 2019). While some studies focus on the value added to the product, users, or other actors involved (Steen et al., 2011), most studies assess design inputs based on their originality, user value, and feasibility (Itten et al., 2020;



Cockbill et al., 2020; Magnusson, 2003; Mitchell et al., 2016). Within the few quantitative empirical assessments, inputs from the users in a co-design process were significantly more original than inputs solely from the research team (Magnusson, 2003) or those from users not in a co-design process (e.g., those asked for inputs by survey) (Mitchell et al., 2016). However originality and feasibility tended to be inversely related, as more novel inputs may be less feasible to actually implement (i.e. out of scope, more expensive to execute, less practical) (Magnusson, 2003).

Designing with the users has been previously explored in energy transition research, frequently in combination with smart meter-based data collection procedures, to improve the effectiveness of energy saving interventions. For example, consumption data has been used as feedback to provide support for energy efficient purchase decisions based on household appliance use (Dalén and Krämer, 2017), improve energy efficient appliance use behaviour (Wever et al., 2008; Mourik and Breukers, 2015), or capture multi-faceted benefits including increasing comfort, energy savings, transparency and overall consumer awareness (Böhm and Szwec, 2013). Recently, an SFOE-funded research project in Canton Vaud investigated how to engage local stakeholders to identify effective ways to reduce the “energy performance gap” between the energy consumptions of highly energy-efficient buildings according to the project design, and their energy consumption in real life (Mastelic et al., 2018). Differently than such cases that rely on technology to reduce consumption (e.g., smart shower metering systems, more efficient appliances, automated shutdown, etc.), in SPP we focus on behavior change measures only, delivered via the Social Power Plus app’s persuasive features and the related Community Energy Challenge. Distinctively, we see the behavioural component of the energy transition as additional and complementary to the accompanying technical and regulatory changes planned in the Energy Strategy 2050.

3.3 Co-design methodology

The co-design process we developed aims to gather concrete proposals for the SPP app and Community Energy Challenge intervention design, through iterations between the participants and the research team. The process is based on other co-design processes in the field of sustainable energy behaviour change (Itten et al., 2021; Webb et al., 2016; Cockbill et al., 2020; Preston et al., 2020). It is initiated at the start of the intervention design process and involves the research team, volunteers of the target population (i.e., the participants), and domain experts that are relevant for the specific behaviour in focus.

The time schedule of activities of the co-design process we adopt is presented in Figure 3. We opted for co-design activities to begin with an online survey to characterize the co-design participants, followed by two workshops facilitated by the research team, to *Discover* and *Design* the behaviour change intervention. These activities follow the context mapping approach of Visser et al. (2005), wherein sensitization to the topic is followed by the participants generating their own ideas, and then analysis by the project team.

The *Discover* workshop first introduces the participants to each other and the research team and establishes a code of conduct to encourage a comfortable workshop atmosphere. The workshop content can begin with background information on the behaviour in focus and continue with a detailed presentation of similar interventions already performed, to provide inspiration. This sensitization step is important to introduce the topic, as well as the theories being applied by the research team (Hurley et al., 2021). Live survey tools are also used to enable quick and anonymous feedback during the workshop and to stimulate short discussions in order to explore the results in more detail.

The second workshop *Design* is held within a few weeks of the first workshop. Here, the goal is to use generative methods (33) to facilitate exchange between the participants and the researchers and collect design suggestions for the intervention. Outputs from the *Design* workshop are compiled and reviewed by the research team and the experts from the field, to develop a proposal for the intervention design.



Based on these initial workshops, the research team and domain experts prepare an intervention prototype to present to participants in the *Evaluate* workshop. The resulting prototype is discussed and evaluated by the participants to get their feedback and suggestions to further enhance the design.

Overall, the workshops aim to reveal participant’s needs, motivations, and challenges to save energy. The research team then identifies how to best integrate the participant’s inputs into the MAP process to increase chances for progress between preaction and action phases and maintain the behaviour in the postaction phase.

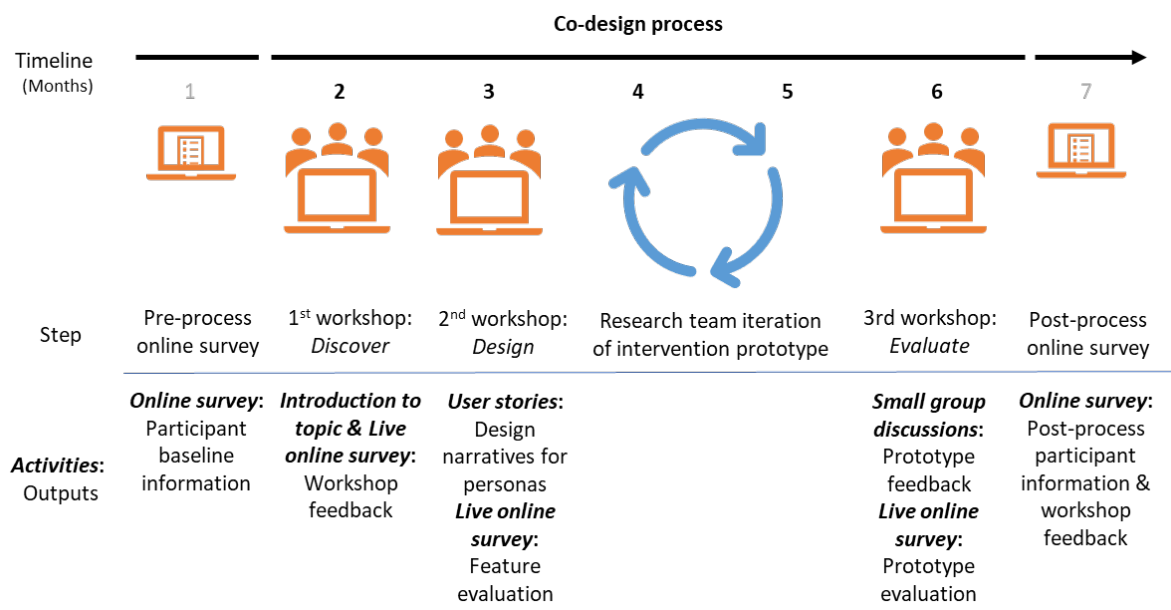


Figure 3: The co-design process: all activities involve participants and researchers, except for the prototype iteration by the research team.

3.4 Implementation of the SPP co-design process

Activities aimed at setting up the campaign for recruiting of participants to co-design activities started in 2020 with the invitation of participants over channels provided by the three energy utilities. Letters and flyers were sent with, and separately from, energy bills, as well as through online communication channels. Potential participants had the technical setup of the target group of the app, i.e. stand-alone house with a smart meter and heat pump or gas heating (see Section 4.1.2). For the Schaffhausen region, additional sensors were installed as smart meters were not available. In the region of Wil, potential participants were invited who had already been involved in an innovative heating renovation programme. Overall, approximately 1200 people were contacted. The utilities also offered a small incentive for participation, e.g. a voucher for local businesses.

The recruitment campaign concluded in early January 2021, with the recruitment of 54 household members willing to participate in the co-creation process. Their socio-demographic descriptions are outlined in Figure 4. The average age of the participants was 53 years old, and they were mostly male homeowners, living in an adult-only household.

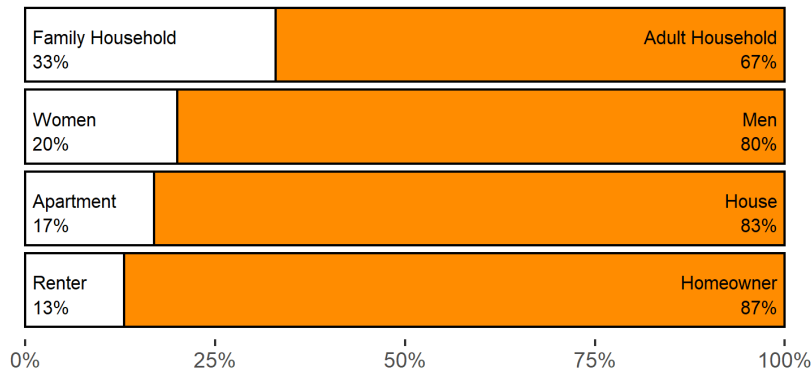


Figure 4: Descriptive information on characteristics of co-design participants (N=54).

The three workshops envisioned by our methodology were held in parallel in each of the three regions, giving nine workshops in total, performed between February and June 2021. Separation by region helped build relationships between the participants and allowed a focus on the local context, as the energy delivery and measurement system differed in each region. Also, the small number of participants per workshop (on average, 15 people) allowed more time for each participant to voice their opinions. Of the 54 participants, 45 (83%) attended at least two of the three workshops. Due to COVID-19 restrictions, all the workshops were held online using a video conferencing platform (Zoom).

We specifically aimed to use co-design to address the challenges outlined in Section 3.1: overcoming the intention-behaviour gap between the MAP phases; ensuring participant engagement (i.e., reducing attrition); and fostering long-term behaviour change. Considering the MAP framework for energy savings (Ohnmacht et al., 2017), the household members were expected to provide insights on what worked for them to save energy at home and what they imagined they would need to progress further. Table 1 outlines exemplary inputs from participants engaged in other co-design processes specifically for energy savings (Itten et al., 2020; Webb et al., 2016; Cockbill et al., 2020), as an indication of the types of inputs we expected to collect. These inputs are mapped onto the socio-psychological factors associated at each MAP phase following (Ohnmacht et al., 2017).

Table 1: Exemplary participant feedback we expected to collect during the co-design process, with respect to socio-psychological factors at each MAP phase.

Behaviour change MAP phase	Socio-psychological factors	Exemplary participant feedback on...
Preaction	Attitudes	<ul style="list-style-type: none"> Motivations to save energy and protect the environment/ climate Concerns to achieving savings in their own homes Knowledge/perception about their own impact Influence of social comparison
	Perceived behaviour control	<ul style="list-style-type: none"> Their own energy and technology literacy Reflection on their patterns of consumption
Action	Action planning (Planning skills)	<ul style="list-style-type: none"> Knowledge about how to save in their own homes Feedback on specific saving potentials in their homes
	Coping planning (Solving implementation problems)	<ul style="list-style-type: none"> Problems to save energy (real or imagined) Social interactions (sharing of suggestions and experiences) with other peer households
Postaction	Resisting relapses & dealing with setbacks	<ul style="list-style-type: none"> Interactions with other household members Needs for reminders and additional information



The co-design process was implemented as follows. The *Discover* workshop ran empathising activities to get the participants familiar with each other and supported the research team in getting to know the energy awareness of the participants. Further, the workshop critically explored other energy savings apps with the aim to identify their strengths and weaknesses. At the end of the workshop, a live online survey tool was used to collect feedback on what could be improved for the next workshops.

The main generative exercise was performed in the *Design* workshop: small groups of five people used generative prototyping (Sanders and Stappers, 2014) to develop stories of how a fictitious user (a persona) in a preaction phase would interact with a new energy savings app. The personas had a specific energy savings goal (i.e. heating, washing or cooking, see Figure 5) and should move from preaction to postaction through the behavioural intervention. The personas explicitly reflected the characteristics of the participants, based on the pre-process survey and *Discover* workshop, to elicit what the participants' experienced, what helped them, or what they would still need to reach the energy savings goal.

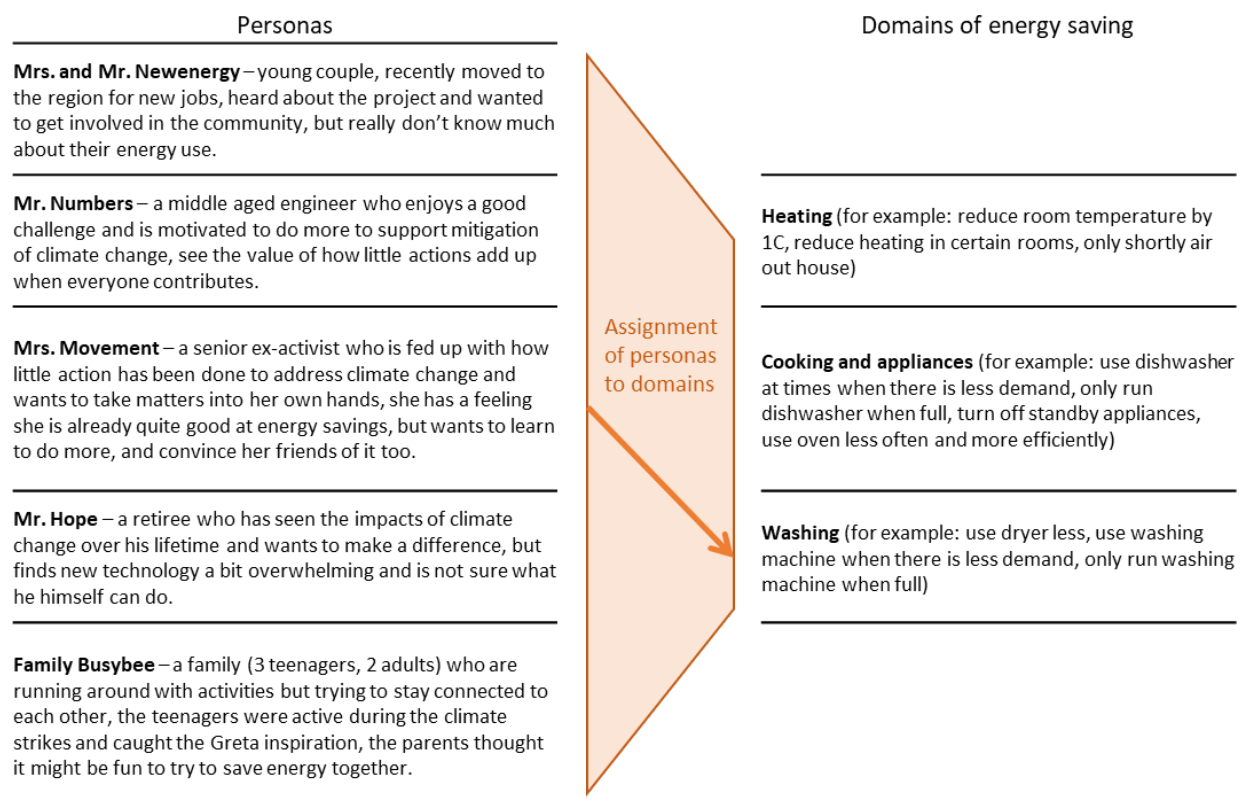


Figure 5: User personas in a preaction phase and domains of energy saving at home used for the development of user stories.

During the *Design* workshop, each small group was invited to create a specific user story about: when and how often the persona would open the app, what information she/he would like to see in the app, which data is interesting and how it should be visualized and contextualized, what the persona would like to do with the data, and how she/he would like to interact with other users of the app. Particularly, when developing a story for the persona trying to reach the particular energy saving goal, each group answered the following questions:

- Which specific energy behaviour shall be addressed in your user story?
- How can the app support behaviour change in this area?
- How is the app used by the persona?
- What features and information can be found in the app to support the behaviour change?
- What does interaction with other app users look like?



The groups guided themselves in developing a user story. This was intended to reduce the influence of researchers in the process and enabled more participant-led outcomes. This generative technique helps researchers access deeper knowledge of the participants as the participants imagine scenarios for a potential user and place themselves in their own created stories (Visser et al., 2005; Sanders and Stappers, 2014). Following a presentation of the user stories, the workshop ended with a live online survey asking about preferences for design features, such as notification frequency, team competitions, goal settings, weekly challenges and gamification elements like points.

A qualitative content analysis (Baxter, 2020) of the user stories and live online survey responses was performed by the research team and reviewed by the domain experts to identify design inputs for the app-based intervention prototype. A preliminary intervention design and non-functional app prototype were presented at the final *Evaluate* workshop, during which participants were invited to provide feedback, via small group discussions (max. 7 participants) and a live online survey.

3.5 Co-design outputs

Participants developed 12 user stories in the *Design* workshop, based on the five different preaction user personas and three energy saving goals provided. The stories depict different ways a persona uses an energy feedback app and engages with the accompanying behaviour change intervention to reach an energy savings goal. This uncovers the participants' perceived needs, motivations, and challenges to save energy. From these stories, five design feature categories, containing 16 unique inputs, were extracted (Table 2).

Some inputs can impact multiple MAP phases, thus ultimately eight inputs emerged related to the preaction phase, eight for the action phase, and three for the postaction phase. This is a valuable outcome considering the heterogeneity of future app users. To address the phase of energy saving behaviour in different household domains, the app could incorporate various features to meet people where they are in the process (Bamberg, 2013).

In the preaction phase (see Table 1 for reference to socio-psychological factors), the participants' emphasis was clearly on increasing perceived behavioural control through detailed energy feedback for more impactful energy saving, and less so on increasing pro-environmental attitudes. Throughout the co-design process, the participants repeatedly requested more individual-level energy use details, energy savings tips, and integration across multiple energy technologies (e.g., data on PV panel production, the load status of an electric vehicle, and interactions between these). This explicit interest in individual, data-driven feedback was further seen in the live online surveys where participants were asked to rank different app features based on their interest to use them: the predominant interest was in seeing one's own energy saving data, and not in the group-level achievements. Interestingly, the motivational gamification elements (e.g., competing with a team, earning points) were often ranked lowest, especially those related to between-household interaction in teams of participants.

For the action phase, the design inputs strongly aligned with solving implementation problems and reducing the risk of dropouts, as participants suggested to set savings goals and competitions to promote active engagement. As well, they suggested an interaction element, a pinboard, where users could interact with each other. The pinboard could be used to both ask questions and exchange experiences, which is particularly relevant for the postaction phase to provide continued encouragement and address newly arising challenges (e.g., because of seasonal changes). The pinboard could also be used to announce results of a between-regions energy saving challenge, thus reinforcing social norms.

Finally, participants proposed to reinforce the behaviour in the postaction phase through notifications coming from the app to help with resisting relapses. Additionally, they suggested energy consumption data should be available at various disaggregation levels over a longer period to review one's own progress.

The user stories also highlighted the implicit motivation and barriers from the participants' perspectives, which go beyond categorization in the MAP, specifically: how the user should interact with the app, what



is the best way to display feedback, what does a team energy savings challenge look like, how can energy savings be achieved, what information should be presented in the app. Importantly, the participants highlighted where potential barriers exist, in terms of digital literacy, convenience, and motivation.

Table 2: Design inputs from user stories for the SPP app.

Design feature	Inputs from participants' user stories	MAP phase	Features included in intervention and app prototype presented in <i>Evaluate</i> workshop	Decision by research team
Energy data feedback	Feedback on environmental impact of own energy consumption	Preaction	Not included	This feature was not implemented due to the diversity of environmental impacts associated to different energy sources.
	Comparison of savings to similar households	Preaction	Heating and non-heating energy consumption of similar households.	Households grouped (families, adults < 65 years old, adults > 65 years old) and average consumption data is shown for their category.
	Energy efficiency rating of household and appliances	Preaction	Not included	Outside of project scope because of technical complexity and resource intensity to get sufficient accuracy on heterogeneous appliances.
	Energy consumption data can be viewed at different disaggregation levels and time scales	Preaction	One section is dedicated to visualizing energy consumption data (bar charts) by hour, week and over the whole project, as well as for heating, non-heating uses, and total energy consumption.	Participants envisioned a more interactive interface where charts could be dynamic (zoomed-in for additional time scales, etc.). However, this was outside the technical resources of the project.
		Postaction		
	Individual goal setting (energy saving) and feedback on progress towards goal	Action	Individual goal setting for both heating and non-heating energy consumption. Charts and feedback on the daily and weekly progress.	Goal setting performed at start and can be modified at any time.
	Integration of other household energy data, such as self-production (e.g., installed PV) or electric vehicle charging	Action	Not included	Outside of project scope because of technical complexity to integrate additional energy data interfaces, which may be proprietary, or more advanced disaggregation algorithms.
Attractive and easy-to-understand visuals for sharing with others	Action	On home screen: intuitive visual representations of energy consumption and goal achievement using stoplight colouring. As well as comparison with similar households.	Features for sharing over social media were not included.	



Design feature	Inputs from participants' user stories	MAP phase	Features included in intervention and app prototype presented in <i>Evaluate</i> workshop	Decision by research team
Energy saving challenges	In-app competition with others, including within the household	Action	Monthly energy saving competition between the three regions using the app.	The competition between regions compares the average energy savings of each region for one week to see which region saves the most. No explicit within-household challenges were introduced, since assessing an individual person's contribution was not technically possible.
	Prompt energy savings actions in different areas of the home	Action	Overall, 21 individual energy saving challenges, focusing on 8 household energy consuming areas (heating, showering, washing, cleaning, cooking, dishwashing, studying and working, recreation).	Challenges incentivize actions to be taken at home, and completion is not quantitatively assessed by the amount of energy saved, as this is technically feasible with the disaggregation possibilities. Instead, to fulfil a challenge the user must share about their experience, problems, and solutions in the Pinboard.
User inputs	Customised notification frequency and timing (e.g., lunchtime savings tips every day at 11 a.m.)	Action Postaction	Automatic notifications sent by the app at pre-defined times.	In the context of the limited resources available for development of the app, implementation of the suggested feature was not possible.
	Track energy use-relevant activities (e.g., doing laundry or having visitors for dinner) to see impact on energy consumption	Preaction	On a daily calendar, users can manually add an activity at a specific time. Such memos can then be compared with the chart showing hourly consumption data.	The energy diary feature was included in the app prototype, but it could not be included in the final app version, since other features were given priority in software development and resources were not sufficient to include all the prototype features.
Customised recommendations	Energy saving recommendations based on analysis of appliance consumption data	Preaction	Overall, 90 general energy saving tips, covering the 8 challenge topics (heating, showering, washing, cleaning, cooking, dishwashing, studying and working, recreation) were added in the app and rated on three point scale along their energy saving potential. However, no personalised tips were implemented.	Offering customised energy saving recommendations, also accounting for external conditions such as the weather or using location features, would require a dedicated project aimed at automatically analysing, disaggregating and comparing energy consumption data against a benchmark and/or other automatically collected data (mobility, weather). This was outside the scope of the project.
	Location tracking to recommend heating regime (e.g., less heating when people not at home)	Preaction	Not included	
	Tips based on weather (e.g., plan laundry tomorrow to use sun for drying)	Action	Not included	



Design feature	Inputs from participants' user stories	MAP phase	Features included in intervention and app prototype presented in <i>Evaluate</i> workshop	Decision by research team
Exchange with other participants	Pinboard for user posts with thread function	Preaction	Pinboards for each of the challenge topics (resulting in over 20 unique pinboards) were created.	An internal "Pinboard" for posting messages (e.g., questions, feedback, suggestions) to all users and the research team was included. This is supposed to favour greater interaction within the app users. Thread function was not technically feasible within the resources of the project.
		Postaction		
	Pinboard with entertaining functions like emoticons or pictures	Action	The pinboard allows for picture uploads.	Emoticons could not be provided due to technical restrictions of the app design.

The knowledge transfer from the co-design workshops to our research team was planned directly after the *Design* workshop and included two months for development (as shown in Figure 3). The new app prototype and intervention design were presented to the participants as exemplary smartphone mock-ups, which provided a clear visual representation of the main app screens and features (Figure 6:). However, it was not possible for the participants to directly experience the functionality by themselves in an interactive app. Nevertheless, in the *Evaluate* workshop participants could directly provide feedback on usability and additional development possibilities.

The prototype of the app-based intervention presented in the final *Evaluate* workshop integrated the design inputs that were technically feasible and within the resources and scope of the project as identified by the research team and experts (see the last two columns of Table 2). For example, the app prototype included a feature to set an energy saving goal with daily and weekly feedback on progress towards the goal. Also, the prototype incorporated an approach for users to self-report appliance use in the app to match it to electricity consumption peaks. Many of the ideas were original and had never been implemented in previous apps developed by our teams (e.g. to have competitions within a household or to display the environmental impact of energy consumption) while others improved ideas we had already explored in a similar app-based behavioural intervention (Wemyss et al., 2019; wemyss et al., 2018) (e.g. using stoplight colouring to quickly and easily indicate progress).

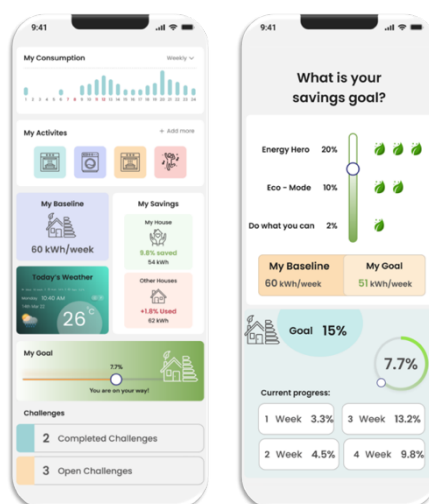


Figure 6: Exemplary non-functional prototype of the co-designed SPP app (left: home screen, right: goal setting screen).



3.6 The SPP Community Energy Challenge and Toolbox

Overall, the design of the app and the related behaviour change intervention directly incorporated many of the participants' ideas. However, the final design was a compromise of all the inputs from the participants and technical and resource limitations in the project.

The SPP Community Energy Challenge and app were finally designed around the key elements summarized in Table 3, with the overall aim of fostering both energy efficiency, sufficiency and flexibility in consumption. The individual-level features, which need to be as customized as possible, refer to goal setting and individual challenges, besides the provision of detailed feedback on energy consumption (disaggregated at least between consumption for heating and for non-heating purposes) for the individual's home. Also, comparison with the consumption of similar households is identified among the relevant features to be offered at the individual level.

The collective-level features, instead, refer to the possibility to share experiences, including successes and failures when dealing with the individual challenges, with the other participating households. For this purpose, an internal forum or "pinboard" for posting messages was created to reinforce the feeling of being part of a community of people engaged towards the same common goal of saving energy. Furthermore, to improve engagement and the feeling of belonging to one's community, monthly energy savings competitions invited participants in the same region to compete against the other regions to save the most energy.

The SPP Community Energy Challenge was designed to last 12 weeks, from February, 1 2022 to April, 30 2022. It was then supported by eight additional months for follow-up, until December, 30 2022, during which energy consumption feedback, reminders and energy-relevant information were to be provided to the participants to avoid relapse to previous energy consumption behaviour. Thus, the SPP Community Energy Challenge is composed of a three-month "high-intensity" app interaction period, followed by an eight-month "low-intensity" interaction period. During the "high-intensity" period, monthly live meetings were offered to all participants, as an opportunity to strengthen social exchange and answer any questions by the research team. These meetings were arranged on a monthly basis and, due to the COVID restrictions, they were all performed online via Zoom, and included break-out sessions for different topics and regions to support the close interactions that would have occurred via in-person meetings.

Table 3: Summary of the features of the SPP Community Energy Challenge and app.

Overall aim of the Social Power Plus Energy Challenge		<ul style="list-style-type: none"> Sufficiency and efficiency: save energy (heating and electricity purposes); Increase flexibility of energy use to help integrate renewables.
Individual app features	"My home", "My consumption", "Challenges"	<ul style="list-style-type: none"> Provide opportunities to set customized energy saving goals (energy saving percentage against a baseline consumption); Provide opportunities for engagement in individual energy saving and flexibility challenges (with customized commitments); Provide feedback on disaggregated energy consumption; Provide comparison with similar households.
Collective app features	"Pinboard", "Tips"	<ul style="list-style-type: none"> Allow possibilities to share experiences with the other households, via a simple online forum/chat; Provide (approx. 90) tips for energy saving in the respective challenge areas of household energy consumption
	"Regional energy saving competitions"	<ul style="list-style-type: none"> Exploit both collaborative (within the regions EKS, TBW, SWW) and competitive (between the regions) motivational factors; Provide as a "booster feature", that is not always active, but is activated on the last week of every month of the SPP Community Energy Challenge; Engage all the households in one region to collaborate to save more energy than the households in the other regions.



	“Monthly live meetings”	<ul style="list-style-type: none">• One meeting a month during the high-intensity period to:<ul style="list-style-type: none">○ to create and strengthen ties between the participating households;○ and address any questions by participants to challenge managers;• ideally offered in-person; due to COVID-19 restrictions, they were offered online via Zoom.
--	-------------------------	--

3.6.1 Energy consumption feedback and goal setting

The app provides daily feedback in the landing screen (“My Home”). A direct connection with energy smart meters (electricity and gas) provides the user with daily feedback on the change in consumption compared to their historical baseline (consumption measured over a comparable past period). If consumption is higher than the baseline, the app background is red; if it is lower, the app background is green (Figure 7.a). This provides an immediate evaluation metrics, letting app users quickly understand whether their level of consumption is good (green) or bad (red). Doing so, SPP leverages an injunctive social norm, suggesting that users should reduce their consumption in order to only see a green background.

“My Home” also provides a comparison with the average energy savings of similar households in the SPP community (families with kids, adults younger than 65 years old, adults older than 65 years old). Here SPP exploits a descriptive norm emphasizing how others are performing. SPP also presents bar plots reporting the household's hourly consumption of the previous 24 hours or seven days (Figure 7. b), or the weekly consumption since the beginning of app use. This allows users to discover their own daily and weekly consumption patterns by intuitively correlating the periods in which they perform energy consuming activities at home with the periods when the bars are high. It is also possible to choose to visualise the energy consumed for heating and non-heating purposes, which are estimated by means of the non-intrusive load monitoring algorithms (NILM, Huber et al., 2021) developed for the project.

The choice to use the background color in the main screen (“My Home”) in order to provide the users with an immediate feedback on their energy consumption is consistent with the outcomes of the recent SFOE-funded “EVISU” research project. EVISU specifically aimed at identifying appealing, simple and intuitive visualizations of energy consumption data, based on the interaction with both experts and lay people. Also in their case, the background color of the smartphone emerged as the most effective way to convey immediate and intuitive information about the amount of energy consumption (Marek et al., 2020). More detailed representations, based on bar charts, emerged instead as the most effective way to provide more detailed and disaggregated energy consumption data, such as the ones provided in “My Consumption” in the SPP app.

The element of social comparison with similar households concerning energy consumption and achieved savings proved to be effective in the warm water use intervention made in the Hüttengraben area (Sandmeier et al., 2020), consequently we introduced this element in the SPP app as well. As the visualization of their own historical energy consumption was very much appreciated also by the inhabitants in the Hüttengraben area, we decided to keep this information as one of the key elements regarding energy consumption feedback in the SPP app.

At the start of app use, every user is requested to specify two energy saving targets (as % saved) for heating and non-heating purposes (Figure 7.c) and then received weekly feedback on the achievement of the goal. If households reach their savings target, a congratulatory message appears on the home screen; otherwise, the message motivates to keep trying to save energy. If the target is too easy (or too difficult), it can be modified by the user at any time. To help goal-setting, the app provides the households with a slider ranging from 0 to 30% savings and shows the baseline energy consumption of each household and the achievable energy savings stemming from different behaviour and practices regarding household energy consumption depending on where the user sets the slider.

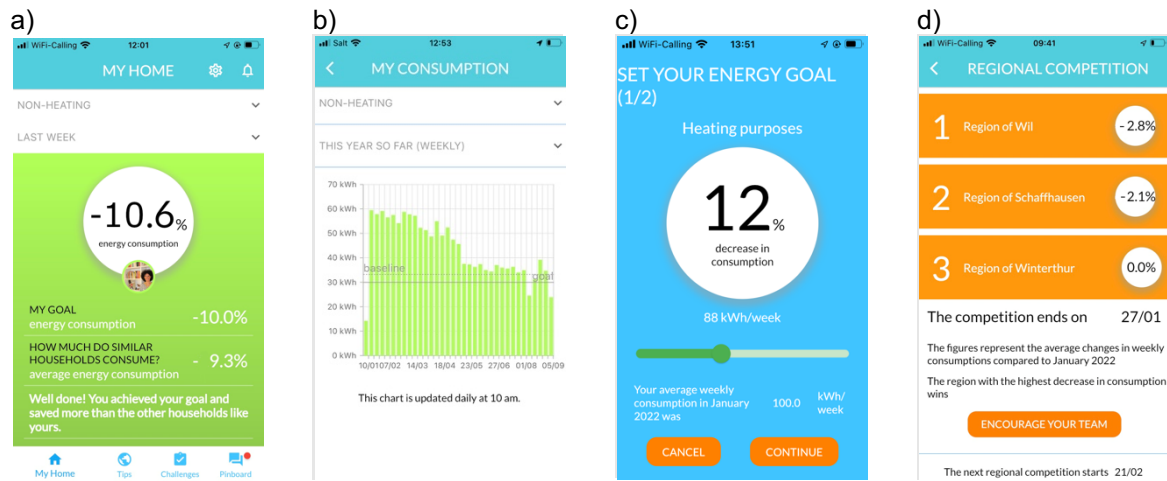


Figure 7: App screenshots providing: a) feedback on energy consumption change; b) feedback on hourly energy consumption; c) goal setting; d) the results of the regional energy competition.

3.6.2 Challenges and tips

To support households towards the achievement of their goal and target, SPP invites users to join hands-on challenges, which aim at modifying the dominant ways household routines are performed to use less energy. Challenges deal with eight household routines (heating, showering, washing, cleaning, cooking, dishwashing, studying and working, enjoying recreation time): for example reduce room thermostat settings by a few degrees, decrease the number of weekly washing machine cycles, or enjoy digital-detox free-time at home during which electronic devices are not used (Figure 8.a and Figure 8.b).

Overall, twenty-one challenges are offered (1-3 challenges per household routine) and are released every two weeks during the Community Energy Challenge. Each challenge lasts for two weeks, during which all household members can potentially engage to save energy by recrafting routines. Households are free to ignore the challenges or to engage in more of them in parallel. The completion of challenges is self-regulated: household members can perform them at the times that best fit their lifestyle and weekly schedule, and are compatible with the constraints affecting their lives. Challenges can also be repeated over time, to support habit creation. To help household members to start taking action, SPP accompanies each challenge with a list of (non-customised) tips, which suggest simple energy-saving actions or low-demand ways to perform the related routine (Figure 8.c). Overall, ninety tips are offered across the different routines.

Challenges and tips are offered to the whole community of SPP users, and thus may stimulate an evolution of social norms and conventions. Through the challenges, the app's focus moves from energy consumption data per se, to what energy is used for (Butler et al., 2018), namely to routines that are performed in the household. Challenges attempt to question collective conventions about the functions and needs that energy consuming good and services are expected to meet (Royston et al., 2018, page 129). Importantly, co-design discussion suggested that challenge achievements do not need to be quantitatively assessed and verified by the energy savings shown in the app. Not only would such an assessment be difficult to perform given the current low accuracy in disaggregation of energy consumptions between different appliances, it would be far from the reason why the challenges were introduced. Challenges are aimed to stimulate individual reflection on current routines and to motivate sharing of experiences on how to complete them. The aim is to get the households' commitment, not to measure their quantitative results, as we expect the energy saving to happen synergistically through all the different mechanisms in the app. Challenge achievement is therefore simply performed by asking the users to share a picture, short text, or tip, summarizing the households' experience when engaged in the challenge, including what worked and what did not, and is directly posted in the "pinboard".



3.6.3 Pinboard and regional energy competition

Besides the support by the project team via challenges and tips, SPP aims at leveraging social interactions and at activating a social learning process between households of the SPP community. For this purpose, an app-section is dedicated to the “pinboard” wall, similar to a virtual forum, where app-users are invited to post comments (Figure 8.d). All users receive a notification when a new pinboard message is posted and every post is visible for all, thus enabling “passive” social interactions (via the observation of others) for those who do not actively engage in posting messages. Users can also “like” messages and can follow-up with a post to engage in dialogue.

There are pinboard sections for each challenge, plus two additional ones regarding “Anything about Social Power Plus” and the “Regional Energy Competition” (see further in this section). To incentivise app users to interact with each other via the pinboard and to fully exploit it as an enabler of social interactions, and hence learning, it is directly connected with the challenges and the comments about the participants’ experiences. To support the creation of a community feeling and the sense of belonging, additional monthly online meetings are open to all SPP households. The pinboard and monthly meetings are key to enable social interactions that potentially activate social learning towards energy sufficiency.

Finally, an additional “booster” social feature is activated once a month: leveraging the sense of belonging to one’s region, a week-long regional energy saving competition is launched between the participants living in different regions (Figure 7.d). No real-life prizes are won, though notifications in the pinboard congratulate the winning region, thus providing a virtual reward, which is expected to help to keep the interest high. The regional energy competition is activated in the last week of every month, thus three times during the “high-intensity” period of the SPP Community Energy Challenge. All the households are automatically included in the competition, namely they do not need to enroll, and all the households are put in the team of their region, no matter what their characteristics are. They are given the goal to save as much as energy as they can, compared to the other two regions, and results of the winning region are revealed at the end of the week. Also in this case, communication and sharing of experiences between participants is encouraged, via the pinboard feature. To remind households of the upcoming competition, dedicated notifications are sent before, during and at the end of the competition.

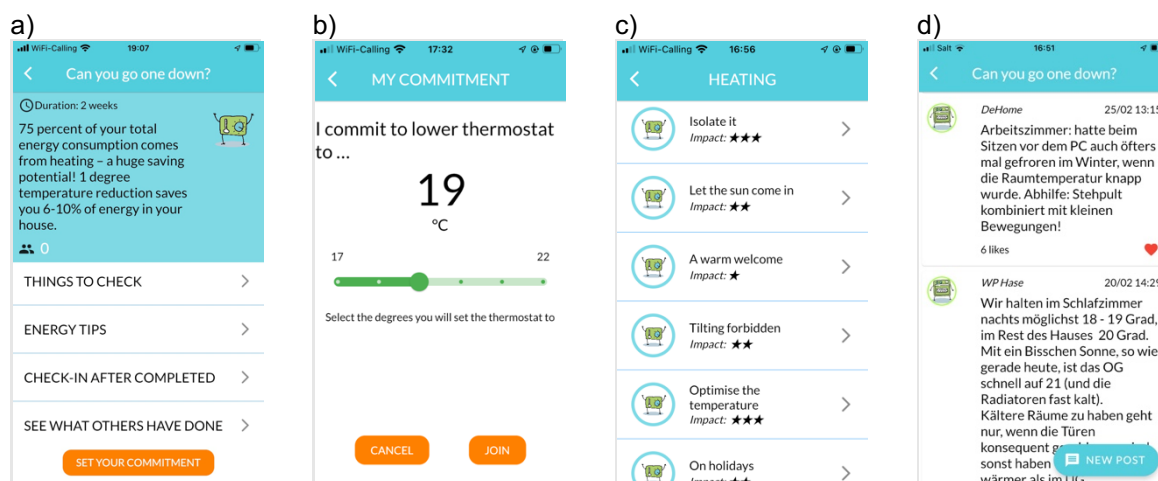


Figure 8: App screenshots providing: a) description of one of the 21 challenges; b) setting a commitment in the challenge; c) energy-saving tips related with the challenge; d) challenge-related comments by app users in the pinboard section.

3.6.4 The SPP Toolbox: app software development, sensors, disaggregation algorithms

The SPP app was developed for both iOS and Android operating systems and released for free download from both [Play Store](#) and [App Store](#) in January 2022. It was also accompanied by a backend system, to allow the managers of the SPP Community Energy Challenge (the research team) to independently modify the content of challenges, tips, and push notification messages (both automatic



and manual ones) whenever needed. Finally, the app was built in English and German. Such a flexibility was introduced to facilitate the later diffusion of the SPP Community Energy Challenge and app in other language contexts.

The SPP app will be removed from the Play Store and App Store after December, 31 2023. After the end of the project, the research team would in fact no longer be able to ensure the technical support and maintenance required by the evolution of the operating systems and backend platforms. A short introductory video and a detailed tutorial of the app will however remain available on the SPP website at <http://www.socialpower.ch/index.php/tutorial>. This ensures a presentation of its key features, structure, and user interface will always be available for dissemination purposes. The full software code of the SPP app will however be stored in SUPSI data-bases and thus is available for further projects.

Besides the SPP app, the other two components of the SPP toolbox are the sensors and the disaggregation algorithms for energy consumption feedback. Their features are strictly inter-twined with the technological characteristics of the energy infrastructures where the SPP Community Energy Challenge will take place. Therefore, here we present them by directly referring to the specific implementation of the SPP Community Energy Challenge in the three regions of Schaffhausen, Wil, and Winterthur, that is the focus of next Chapter 4.

The sensors were developed for SPP based on previous prototypes developed by CLEMAP under the name of CLEMAP One. CLEMAP One is a 3-phase energy meter that is able to fully monitor the electricity consumption of a property. It is installed in the distribution board or fuse box of the property and serves as an electricity monitoring device without the need for additional measurements of any kind. The meter consists of an electrical 3-phase sensor equipped with an integrated Wi-Fi receiver. This connects to the CLEMAP server via the wireless network or LAN cable, where the encrypted data is processed and then forwarded to SUPSI Servers for visualization on the SPP app. CLEMAP One enables the installer or customer to analyze and extract the electricity consumption of the property in depth thanks to 3-phase details on active, reactive power, voltage and current.

Considering the project budget allowed to only develop 100 CLEMAP sensors, a plan on how to allocate them across the regions of the three utilities to be involved in the SPP Community Energy Challenge was developed, by considering availability and characteristics of their metering networks. A decision was made to attribute all CLEMAP sensors to participants from EKS customers in the Schaffhausen region where currently no smart meters are available. As the sensors only read electricity consumption, EKS participants had to live in a house with an electricity-based heating system (heat pumps or direct electric heating).

For the region of SWW customers in Winterthur, instead, certified Landis+Gyr smart meters for electricity were in deployment by the utility. In that case, we opted for directly relying on the SWW smart meters for electricity data collection and made an agreement with the SWW utility to install and activate them with high priority to the SPP households, in order to guarantee they were available for the start of the SPP Community Energy Challenge. Also in this case, therefore, the requirement follows, that households participating in SPP in the SWW region have an electricity heating system.

For the case of the TBW customers in Wil, instead, certified smart meters for gas and electricity were already widely installed and operated by ESOLVA AG. In that case, the decision was to directly exploit such smart meters for data collection during the SPP Community Energy Challenge.



Additionally, households in the SWW and TBW regions could not have an installed photovoltaic plant (PV), as the self-produced electricity could not be accounted for in the household consumption read by the (bi-directional) smart meter and thus the SPP app could not provide accurate information. Bidirectional smart meters only track the exchange of electricity with the grid, therefore in case of self-consumption of electricity produced by household-related PV installation, lower consumption values than the actual ones would be registered, making the provision of feedback unrealistic. For the region of EKS, instead, the presence of PV plants was not critical, since the CLEMAP sensors measure the whole amount of electricity consumption associated to the households, independently of its source (the grid or a possible PV plant). The installation scheme of the CLEMAP sensors in the EKS region is represented in Figure 9.

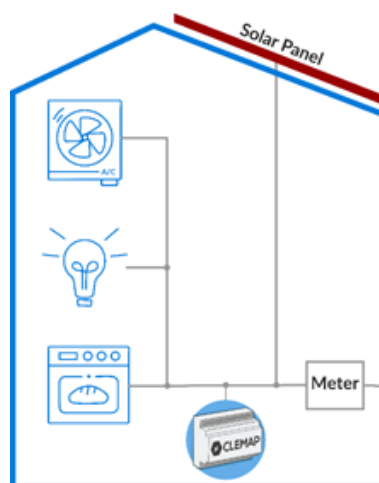


Figure 9: Typical installation of a sensor device with a EKS household with PV-installation.

An additional requirement for participation was to have an independent heating systems, thus participants had to live in stand-alone/single family houses (as households with a dedicated heating system in multi-apartment buildings are not common in Switzerland).

Based on the characteristics of the metering network and available energy data, we identified different objectives for energy disaggregation algorithms in each region. In the EKS and SWW regions the differentiation of heating, non-heating and standby electricity consumption was completed by disaggregation algorithm based on the reactive and active power (for more information, see for instance a blog-post on CLEMAP website at <https://www.clemap.com/blog/lastaufschlussung-wp>). For the EKS region participants, the CLEMAP sensors have much higher frequency than the smart meters (twelve data points per second for a CLEMAP sensor vs. one data point per fifteen minutes being sent by smart meter), which ensures higher accuracy of the disaggregation estimates.

Table 3 summarizes the requirements for households to participate, the data collection sources, and the goals of energy disaggregation algorithms. The disaggregation accuracy of the developed algorithms, the characteristics of which are summarised in Table 3, was estimated by CLEMAP to be equal to 80-85% for the case of EKS and thanks to reference measurements calculated to be equal to 85-97% for the case of SWW. For the case of TBW, disaggregation between heating and non-heating consumption has a 100% accuracy, since the related data comes from two different metering sources.

Table 4: Requirements for participants of the SPP Community Energy Challenge, data collection sources, and energy consumption feedback provided by the SPP Toolbox, for each region.

Utility	Requirements for SPP participating households			SPP collects energy consumption data via...		SPP provides feedback on energy consumption for...
	Type of heating system	Energy source for heating	Presence of PV plant	Heating purposes	Non-heating purposes	
TBW	Individual (single family house)	Gas	Only accepted if energy production is	Gas smart meters	Electricity smart meters	<ul style="list-style-type: none"> • Heating • Non-heating



SWW	Individual (single family house)	Electricity (heat pump)	entirely fed into the grid, without self-consumption	Electricity smart meters	<ul style="list-style-type: none"> • Heating • Non-Heating • Standby
EKS	Individual (single family house)	Electricity (heat pump/ direct electric)	Accepted, as it does not affect consumption data collected by the sensors	CLEMAP sensors	<ul style="list-style-type: none"> • Heating • Non-Heating • Standby

In the case of EKS, data acquisition and synchronisation with the server on which data is collected was directly ensured by CLEMAP, as they both developed the sensors and managed the related server. For TBW and SWW, data was coordinated between the utilities' smart meter infrastructure and SPP. In SWW the synchronisation via XML data synchronisation was operative since September 2021, while the synchronisation with TBW server infrastructure (managed by the company ESOLVA) and based on a dedicated FTP server, was deployed in December 2021. In all cases, the related energy consumption data was first collected on CLEMAP servers, where data was processed to get the needed consumption disaggregation feedback. The disaggregated data was then passed by API to SUPSI servers for visualisation on iOS and Android SPP app.

3.7 Discussion

Co-design uncovers different underlying factors, including unconscious habits (Lee et al., 2020) and considers the broader context of behaviour change beyond intentions alone (Bamberg, 2013). In the specific SPP case, co-design activities allowed space for the participants to reflect, imagine, and contribute to our understanding of what stands between the preaction, action and postaction phase in terms of practical and real-world barriers. These barriers may be missed by a strictly theory-driven approach. Further, the user stories and workshop could directly address the challenges of participant attrition and behavioural maintenance which are specifically context dependent (Gerorges et al., 2016; Frey and Rogers, 2014). For example, co-design participants suggested to send smartphone notifications at relevant times before a behaviour takes place or to use the pinboard to enhance the feeling of a supportive community towards a common goal.

Overall, the co-design process informed decisions beyond best-practices from literature and incorporated local needs and current interests in the topic of energy saving. Through the user stories and feedback of participants, the design addressed the persisting issues of low engagement in behavioural interventions, lack of embeddedness in the user's real-world context, as well as developed approaches to support long term maintenance of the desired behaviour.

3.7.1 Reflections on the co-design process

The participants served as novel sources of inspiration for our experienced app development team. However, we had to balance the limited feasibility of some of the inputs within the pre-defined scope of the project, as well as acknowledge the participants' limited experience with designing for behaviour change. While the participants felt that their opinions mattered and were heard, it was important to keep expectations realistic that not all feedback could be incorporated into the final app and intervention design. For example, during the workshops, it became apparent that the participants were less interested in the motivational gamification elements and preferred instead to have more data and direct feedback on their own behaviour. This input primarily enhances perceived behavioural control at the preaction phase. However, to further support the preaction phase following the MAP framework, we additionally chose to address attitude in the prototype design. Thus, taking a wider audience into consideration, gamified motivational features were included in the app (e.g., individual challenges and a between-groups energy saving competition), along with the requested individual-level behavioural and energy use feedback.



Additional to the single design inputs, the user stories also provided a rich picture of participant priorities. We used the motivations and barriers mentioned by the participants to guide decision-making during the design. For example, the intervention resulted in less focus on competition, earning points, or reaching a predefined goal, than we initially envisioned. This was a critical decision for us considering the hypothesis that individual behaviour change may be motivated by the awareness of social norms (Bergquist et al., 2019), e.g. knowledge that others are also acting in alignment with the same goal. However, this type of reflexivity on research is intrinsic to co-design and achieving more transformational impact (Nikas et al., 2020). Thus, the prototype and the intervention now focus more on exchange and support between users, and the ability to set their own individual saving goal.

Overall, this abductive process trained reflexivity in the research team, i.e., an increased awareness of assumptions and roles in decision-making, which is an expected outcome of co-design (Steen, 2013).

3.7.2 Reflections on co-design participants

The socio-demographic indicators reported in Figure 4 show that the participants are not representative of the general Swiss population in terms of their age (Swiss average 43 years old vs. 53 in sample), gender (Swiss average 50% men vs. 80% men in sample), house type (Swiss average 18% (semi-) detached house vs. 83% in sample), and house ownership (Swiss average 36% owner vs. 87% in sample) (SFSO, 2022a, 2022b). However, representativeness was neither expected nor the aim. Instead, participants should reflect the intended behaviour change population, and thus in the implementation case we recruited household energy consumers with a similar energy infrastructure as would be needed for participation in the future app-based intervention: that is, they live in a (semi-) detached house with their own electrical or gas heating system.

The requirement for a (semi-) detached house is related to house ownership, which is rare in Switzerland at only 36% of homes being privately owned, and thus the lowest in Europe (compared, for example, to 65% in the UK and 51% in Germany (Statista, 2021)). Thus, it is not surprising, that the average age of participants is quite high, as many people in Switzerland can only afford a house later in life. The predominance of men in the participant group is not surprising either, as affinity and self-efficacy for energy and technology-related topics tend to be more prevalent in men (Zeldin et al., 2008; Chan, 2022), thus more men will participate in such processes if recruitment does not explicitly aim for a gender balance (Bear and Woolley, 2011).

Nevertheless, the participants are likely not representative of the general population from the perspective of their phase of behaviour change to save energy. The app is intended for people who are at the preaction stage of energy savings, i.e., they want to reduce the impacts of their energy consumption but are missing the necessary information, knowledge and competencies (perceived behavioural control), or the last push to act (attitudes). It was apparent that the incentives to participate in the co-design process aligned with the participants' relatively high energy awareness and knowledge, which qualitatively emerged throughout the process. Such awareness and knowledge also quantitatively emerge the pre-process survey, that, among the others, measured the following item: "I believe that I can contribute to solving the energy problems by making appropriate energy-related choices and action", by using a 5-point Likert scale (1- I do not agree; 5- I fully agree). The mean response is in fact equal to 4.13 (SD = 0.91; n = 38), showing high perceived awareness and knowledge.

In the same pre-process survey, we also asked participants to rate their reasons to participate in the co-design process on a scale of 1 (not a reason to participate) to 5 (very much a reason to participate), and the responses are reported in Figure 10. These motivations were further reflected in the questions and inputs of participants during the workshops. Overall, by accounting for the survey answers and the interaction in workshops, three main motivations appear: a) *technological curiosity* – being interested in learning more about energy technologies such as optimizing photovoltaic panels and electric vehicle charging; b) *energy saving* – wanting to learn how to save (more) energy and possibly also save money in doing so; and c) *climate relevance* – wanting to do more to address climate change. These motivations highlight the participants' intrinsic motivation for energy saving (i.e., preaction and action) and showed their interest and openness to change.

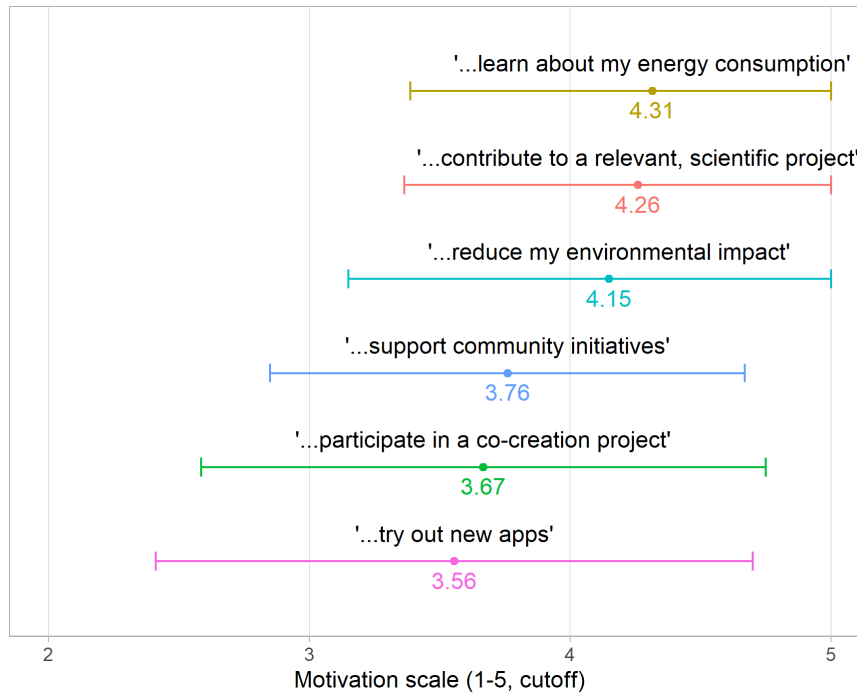


Figure 10: Rating of different motivations to participate in the co-design process showing mean and standard deviation bars.

Indeed, as the household participants were voluntarily recruited with only a small incentive (i.e., a gift certificate) for participation, their intrinsic motivation was necessarily high from the start. The participants' high intrinsic motivation could question whether they were the right participants for the design of an app for a less motivated audience. To address this, the personas, used to develop the user stories, were inspired from the characteristics of the real participants albeit setting the persona in the preaction phase. Participants were instructed to imagine the experience of a future app user from this perspective, therein we assumed that a future user would be similar to the co-design participants. The post-process survey response provided an additional perspective on future users: the participants did not expect their close network would be interested in the app and intervention. Thus co-design can provide insights for future design, but will not necessarily capture the entire potential intervention audience.

The participants' high motivation might have also affected the design results of the app. Namely, the app may have limited appeal for people with different attitudes and motivations, e.g., young people or uninterested and unexperienced homeowners. Nonetheless, the self-selected co-design participants are likely comparable with those that later on will self-select to join the app-based energy savings intervention, which will also be voluntary to join.

It is to be expected that voluntary behaviour change processes attract those already interested in addressing the impacts of their behaviour. This is unlikely to change without wider social movements or policy measures which change attitudes. For the energy transition, this will mean that people who will participate (in co-design or in an intervention) are already more aware of the relevance of energy consumption and are motivated to change. Amongst the case study participants, the three main motivators to participate (i.e., technological curiosity, energy saving, and climate relevance) provide some indication on how the energy savings app and project are perceived from the outside, despite this set of participants being particularly energy affine. For future intervention co-design, dedicated efforts should be made to attract more diverse participants who may have other non-energy related interests to participate in a co-design process (e.g., curiosity about an app development process).



4 Analysis of the effects of the SPP Community Energy Challenge

The project's second objective is to assess the effectiveness of the SPP app and Community Energy Challenge to produce tangible energy savings in the participating households. Namely, we aim to estimate the effect of the treatment with the SPP app on households' energy consumption, which is our primary outcome.

4.1 Methodology

We arrange a real-life intervention in the regions of EKS, SWW and TBW. To assess its effect, we develop and implement a quasi-experimental research design, with a self-selected treatment group of households identified through targeted communication and recruitment activities, and a control group identified in each region starting from the full set of household customers of the utility companies. The inclusion of a control group is needed to rule out possible effects on the outcome variable (the energy consumption) due to external factors, such as for instance the meteorological variability, the Russian war in Ukraine, or the Swiss Federal Office of Energy communication campaign launched at the end of August 2022 to spread the message that “energy is scarce” and invite the population to adopt energy saving measures within the household (<https://www.dont-waste.ch>, last accessed on November, 7 2023). Both the treatment and control households are exposed to and equally affected by the same exogenous events.

Household energy consumption is measured through data regularly collected by the utility company for invoicing and grid management purposes. On varying the utility, the related data comes from smart meters (TWB consumption for gas and electricity) or from manual readings of the electricity consumption meters (SWW and EKS). In the latter cases, manual readings are performed at specific intervals during the year, which differ between the utilities and within the sample of their household customers.

The effect of the intervention is calculated by comparing the average changes in consumption of the participating households (the treatment group) and other households from the utilities (the control group). The treated households self-selected themselves into the treatment, namely they were voluntary households that decided to answer a public call looking for project participants. Such a self-selection mechanism could not be avoided especially for the case of EKS customers, where they had to accept installation of the CLEMAP sensors in their home.

Self-selection of treated households can become a source of bias in the estimation of the true effect of the intervention if the *a priori* energy consumption trends of the two groups are not parallel before the intervention takes place. To empirically test the parallel trends assumption (i.e., in the absence of treatment, the difference between the treatment and control group is constant over time), we acquire baseline data on multiple pre-treatment time periods and visually inspect whether changes in the mean outcomes by group appear to be approximately parallel. For both treated and control household, we test also whether baseline values of the outcome are effectively uncorrelated with its changes over time. In the case of violation of the two assumptions, we implement statistical matching procedures to create a weighted combination of control households that mimics the trend of the treated households before the treatment (Ryan et al., 2015). Then, we estimate relative changes in energy consumption between the two groups (treatment and the matched control) over one year in a difference-in-differences (DID) set-up.

The resulting research design is shown in the left-hand side of Figure 11, which visually represents the DID panel approach we adopt. We collect energy consumption data before (i.e., baseline consumption or pretest data) and after the SPP Community Energy Challenge started (i.e., posttest data), compute their average differences over time for both the treatment and the control group, and verify how the changes differ between the two groups. This allows us to estimate the average treatment effect on the



treated households (ATT). To test whether impacts observed after the SPP Community Energy Challenge are maintained over time, or if rather there is a relapse to previous consumption levels, we measure the outcome of the SPP Community Energy Challenge and estimate the ATT both in the short- and in the long-term. Finally, we perform a heterogeneity analysis of the effects: we check whether estimated ATTs differ between regions by considering sub-samples of the households based on their utility company. We also consider sub-samples of households based on their heating system, considering the heat pump households (regions of EKS and SWW) compared with the gas heating households (region of TBW).

Besides the main effect on energy consumption due to the treatment with the SPP Community Energy Challenge, we also measure whether and how other socio-psychological and contextual variables are affected by participation in the SPP Community Energy Challenge (right-hand side Figure 11). For this purpose, we send three surveys to the treatment group before the SPP Community Energy Challenge (Jan 2022), directly after the high-intensity phase (April 2022), and eight months later (Jan 2023), respectively. The surveys investigate socio-psychological aspects (e.g. motivations for participation, environmental attitudes, self-efficacy, perception of social norms, etc.), self-reported changes in household routines, as well as the level of engagement by app users with the different features of the SPP app. As the control group is an anonymous and unknown group of household customers of each utility, these surveys cannot be sent to the control households. Thus, the related variables are only measured for the treatment group households, following a one-group pretest-posttest design.

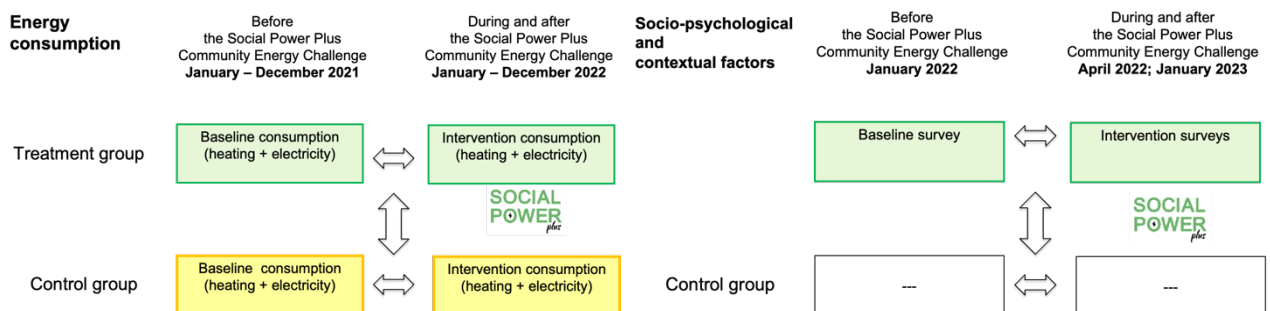


Figure 11: Summary of the quasi-experimental design to assess the effectiveness of the SPP Community Energy Challenge.

4.1.1 Timeline of the SPP Community Energy Challenge intervention

Figure 12 summarises the timeline of the whole set of activities we performed to implement and evaluate the SPP Community Energy Challenge intervention. As described in section 3.6, the SPP Community Energy Challenge intervention was designed to run for three “high-intensity” months, followed by eight additional “low-intensity” months for behavioural maintenance and follow-up. During the high-intensity months, all app features are available: novel challenges and tips are released every two weeks and a regional energy competition is held in the last week of every month. Furthermore, exchanges with the project team are frequent: through push notifications, messages posted on the pinboard, or group interaction via the online monthly meetings. During the low-intensity period, there are no new challenges or other gamified activity launched via the app. Rather the focus is on maintaining the achieved savings and behaviour changes achieved during the high intensity phase. Thus, feedback on household energy consumption remains active in the app and is updated on a daily basis (with hourly granularity). Further, reminders and energy-relevant information are provided to the participants in and outside the app (blog available on the SPP website) approximately once a month to support the maintenance of the behaviour (Fischer, 2008) and reduce the relapse to previous consumption behaviour (Ohnmacht et al., 2017).



	2022												2023								
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep
Pre-survey (survey 1)																					
SPP Community Energy Challenge - High-Intensity period																					
SPP Community Energy Challenge - Low-Intensity period																					
SPP Community Energy Challenge - Additional Low-Intensity period (not considered in the analyses of effect)																					
Intermediate-survey (survey 2)																					
Post-survey (survey 3)																					
Evaluation (interviews to utilities, research team workshop)																					

Figure 12: Timeline of the SPP Community Energy Challenge intervention.

The intervention started on February 1 2022 and was planned to run until December 31 2022. Due to the Russian war in Ukraine, the consequent energy crisis that affected Europe in 2022, and the explicit wish for long-term consumption comparison by the participants, we offered to extend the consumption feedback information in the app for interested participants until May 31 2023, again via “low intensity” interaction offered by the app. Participants interested in this extension were explicitly asked for their consent to continue to send their consumption data to the app and to analyse their data for research purposes. The analyses we performed to estimate the effects of the intervention does not take this extension into account and only considered consumption during the 2022 February- December period.

In summer 2023, at the end of all activities with household participants, we performed final interviews with the utility companies to collect their evaluation of the project. With the same goal, we also performed an evaluation of the project with the project research team (SUPSI, ZHAW, and CLEMAP data management company) in an online workshop supported by a Miro board interactive tool.

4.1.2 Availability of energy consumption data

Adopting the DID scheme requires collecting baseline data to represent energy consumption before the SPP Community Energy Challenge for the treatment and control groups. This strongly affects the choice of the timeline of the whole SPP quasi-experiment as to guarantee the availability of the required energy consumption data. As reported in section 3.6.4, smart metering infrastructure was available for a large number of households only in the case of TBW. For EKS and SWW, instead, electricity consumption data comes from manual meter readings for the control groups and baseline data collection of the treatment group households.

For EKS, a manual reading of consumption data is performed once a year, providing data for the calendar year, from January to December. For SWW, the manual reading is conducted four times a year between periods of generally equal length. However, the day of the reading may differ from household to household. Considering the constraints on consumption data and the timeline of the project, the assessment of the impacts of the SPP Community Energy Challenge is performed by comparing yearly energy consumption data over periods that are as close as possible to the intervention period (January – December 2022). To compare, baseline consumption data in years 2021 and 2020 is used for the three utilities.

Besides estimating the impact of the intervention over one full year of consumption, quarterly data are available to analyse its intermediate effects in the cases of TBW and SWW where consumption data is available within the year. More specifically, TBW provided four consumption data points representing consumption measured at regular quarterly intervals in 2020 and 2021 (January 1 – March 31; April 1 – June 30; July 1 – September 30; October 1 – December 31). SWW also provided consumption data on a quarterly basis for each household in 2020 and 2021. However, the meter readings are carried out manually –and therefore not simultaneously at the same time point– by the utility, thus generating misalignment in the quarterly starting dates. By carefully selecting the fitting reference quarters on an annual basis, SWW households were used together with TBW households to study the evolution of the treatment effect over time.

Table 5 summarises the periods for which energy consumption data is available for our estimates of the treatment effect. Based on such data availability, we investigate the effect of the intervention on both



EKS, TBW and SWW households looking at changes in average annual energy consumption between 2022 and the previous two years. The analysis is conducted using both the energy consumption for the year immediately preceding the treatment (2021) and two previous years (2021 and 2020) as a baseline. We also investigate how the effect evolves for the SWW and TBW households, as energy consumption data within the year is also available for them over four quarterly measurement periods.

Table 5: The energy consumption periods we consider to estimate the treatment effect, on varying the three utilities.

Utility	Household technical characteristics	Reading type	Available consumption data	Consumption measurement periods	Years
EKS	Household customers with electric heating	Manual	Yearly electricity consumption	January – December	2022 (intervention); 2021, 2020 (baseline)
SWW	Household customers with electric heating	Manual	Yearly electricity consumption	Different from household to household. In most of the cases, they correspond to January – December. In a few cases, the yearly period is November – October or December – November.	
			Quarterly electricity consumption (four readings per year)	Different from household to household (reflecting the above variability)	
TBW	Household customers with gas and electricity smart meter	Automatic (smart meters)	Yearly electricity consumption	January – December	
			Quarterly electricity consumption (four readings per year)	January 1 – March 31 April 1 – June 30 July 1 – September 30 October 1 – December 31	

4.2 Surveys and in-app automatic data collection

As indicated above, the assessment of the effectiveness of the SPP Community Energy Challenge in the reduction of energy consumption is performed via two complementary methods: besides the impact evaluation on energy consumption, against a control group of comparable households that were not treated with the SPP Community Energy Challenge, we also perform a series of surveys specifically targeting the treated households only, according to a one-group pretest-posttest research design.

The surveys are aimed at investigating socio-psychological and contextual factors to profile voluntary participants to the SPP Community Energy Challenge, evaluate progress towards the expected outcomes, and collect feedback on the SPP app and intervention. The surveys are also accompanied by the analysis of the data automatically collected by the SPP app, namely an analysis of the frequency and intensity of use of the different app features and of the qualitative materials posted in the pinboard.

The surveys are administered to all households in the treated groups at three time periods, namely before the start of the SPP Community Energy Challenge intervention (T1, January 2022), at the end of the “high-intensity” intervention period (T2, May 2022), and at the end of the “low-intensity” intervention period (T3, January 2023). For all such survey waves, the invitation to answer the survey, administered through Qualtrics, is sent to a reference person of each participating household, followed by up to three reminders to increase response rate. Specifically, the survey content is reported in Table 6, for a total duration of about 15 minutes.



The “baseline” survey at T1 aims to collect socio-economic information about the respondent, regarded as a proxy for the whole household, and on the house characteristics, followed by information about key variables that, according to previous research, are expected to influence household energy consumption behaviour and routines (beliefs, social norms, perceived behavioural control, thermal comfort perception). We also investigate self-reported data about the evolution of energy consuming routines at home (reported energy behaviour: from average values of thermostat settings into living rooms, to the number of cycles of washing machines or the number of hours per day during which IT appliances are running). Also, the motivation for joining the SPP Community Energy Challenge are explored. In the “short-term” and “long-term” surveys, respectively at T2 and T3, only some of these variables are investigated, namely only those that are expected to possibly change by participating in the SPP Community Energy Challenge. Additionally, in the “short-term” survey wave also feedback on the SPP app and Community Energy Challenge is explicitly explored, by means of both closed- and open-ended questions.

Finally, in the “long-term” survey wave we also explicitly explore the households’ perception of the ongoing energy crisis. The whole SPP Community Energy Challenge is occurring in a period during which household energy consumption has become a critical emergency at the federal level, due to shortages in energy supply caused by the Russian war in Ukraine. Presence of a control group allows us to account for the changes in consumption that are due to such external events, rather than to the SPP Community Energy Challenge. Due to the relevance of the conflict and its implications on energy consumption patterns and practices, we however added a new variable to be investigated in the last survey questionnaire, which is inspired by Ahorsu et al. (2022) and aims at providing insights on the perceptions by the households on possible energy shortages during the winter months.

Table 6: Contents investigated in the three survey waves.

Variables	Baseline (T1) January 2022	Short-term (T2) May 2022	Long-term (T3) January 2023
Energy Literacy	x	x	
Reported energy behaviour	x	x	x
Motivation	x		
Beliefs	x	x	x
Social norms	x	x	
Perceived behavioural control	x	x	
Personal details	x		x
Reported environmental behaviour	x		
Thermal comfort	x	x	x
App feedback		x	
Perceived utility of energy consumption feedback		x	
Energy crisis perception			x



The analysis of the in-app data is instead performed by considering all the data collected during the “high-intensity” period only (February 1 – May 1 2022), as that one is the only period during which all the intervention features are available and therefore a comparison on the level of engagement that they offer is possible. We consider the type and frequency of interactions with the app by participating households, according to the following indicators:

- number of app openings per household;
- number of goal setting per household;
- number of posted pinboard messages per household;
- number of started challenges per household;
- number of completed challenges per household.

For each indicator, we compute the average value of the indicator by only accounting for the households that performed the related action at least once. Finally, we also consider the analysis of the pinboard content, to verify if presence of the pinboard feature stimulates app users to interact with each other, thus possibly leading to social learning processes between the participants to the SPP Community Energy Challenge. All pinboard exchanges took place in German language, though our analyses are performed on their literal translation in English by the DeepL Pro software tool.

4.3 Recruitment of the participants (treatment group households)

Recruitment of the participants was performed from late Summer 2021, with the support of the SWW, EKS and TBW utilities. According to the project proposal, the goal was to recruit $n = 100$ participating households in each region. To achieve this target, we used a few complementary strategies that were customized by each utility. Such a customization was needed to better fit with the regional requirements for participation in the SPP Community Energy Challenge (see Table 4), which are in turn connected with the data collection tools available to us in each utility’s region (smart meters, CLEMAP sensors).

A customized post letter was sent by each utility to a large number of customer households meeting the requirements presented in Table 4 (3200 personal letters for EKS, 1100 letters for SWW, and 1500 letters for TBW), followed by communication activities on the utilities’ newsletters and social networks. TBW also offered a 50 CHF voucher for their online shop to the first 50 applicants. Participants applied online via the Social Power Plus website using an online registration form (<http://www.socialpower.ch/index.php/anmeldung/>). The project website reported all key information about participation in the SPP Community Energy Challenge, including the eligibility criteria and how sensitive individual data is managed, stored, and published. The latter elements were explicitly approved by the ZHAW Ethical Committee. Applications started in September 2021 and remained open until January 2022, to allow detailed applicant checks by both the SPP team. This rather large time window for application allowed the utilities to reinforce communication activities where applications were low, which in particular occurred in the TBW region. Once the eligibility checks confirmed the respect of the requirements listed in Table 4, no further selection among the applicants was performed. Therefore, the sample of households participating in the SPP Community Energy Challenge is a self-selected sample.

The recruitment process ended with 339 applicants, split among the three regions as reported in Table 7. In the regions of SWW and EKS we reached the number of target applicants (100) and even exceed it –which was appreciated, in order to account for possible drop-out during the quasi-experimental activities on the field. In EKS, where applicants were more than the maximum number of available sensors, we created a waiting list. In the region of TBW, instead, a slower application rate was registered than in EKS and SWW, and additional recruitment and communication activities were first implemented in October 2021 and then also in December 2021. Overall, despite these additional activities, that also included monetary rewards/vouchers, the number of applicants in the TBW region remained low, far from the 100 household target. According to the utility TBW, the reason for the low number of applications in their region can be due to a “participant fatigue”, namely to the fact that, exactly in the months before the launch of the recruitment for the SPP Community Energy Challenge, other initiatives around energy saving in households were taking place in the region.



Table 7: Number of applications received and accepted for the SPP Community Energy Challenge, for each region (“Treatment group”).

Utility	Number of received applications	Number of accepted applications	Comments
TBW	68	34	Houses with PV panels or shared heating system were not accepted
SWW	118	108	Houses with PV panels or shared heating system were not accepted
EKS	153	79	Certain households could not be accepted, due to technical problems occurring during the installation of the CLEMAP sensors (missing space, lack of compatibility between the sensor and the building technical set-up, including availability of Internet connections for data-transfer)
TOTAL	339	221	---

Once households were recruited, in January 2022 activities were performed in order to connect their energy consumption data with the SPP toolbox. In the case of EKS, this required that the CLEMAP sensors were installed in each participating household. This was performed directly by the SPP team, with a support of an external electrician, who was paid by dedicated project funds. In the case of SWW, this required that the households were equipped with the electricity smart meters and that the latter were properly activated, for data connection and sharing with the SPP toolbox. This was done directly by SWW and then verified by the SPP team. Finally, for TBW smart meters for gas and electricity were already available and connected for all applicant households. The SPP team therefore only set up the data connection procedures with the SPP toolbox and ensured that everything worked properly. Overall, Table 7 reports the numbers of accepted participants, namely the sub-sample of applicants where the eligibility criteria were fully met and no technical problem emerged during the installation of the CLEMAP sensors. In total, 221 households could join the SPP Community Energy Challenge as “treatment group” households. Namely, around 65% of the number of applicants.

4.4 Definition of the DID analytic sample

DID is one of the most popular approaches for evaluating the effect of a treatment in a quasi-experimental scenario. DID analyses can be conducted any time an outcome is measured at least twice (before and after the stimulus occurred) on a sample made up of units that have been exposed to a stimulus (the treatment group) and units that have not been exposed to it (the control group). Applying this method to our case requires average changes observed in energy consumption data of the SPP Community Energy Challenge participants to be compared to those observed in a control group of comparable households that are not treated with the SPP Community Energy Challenge. To do so, we requested the SWW, EKS and TBW utilities to provide longitudinal energy consumption data for all their household customers, that respect the requirements reported in Table 4. The utilities provided us with the yearly energy consumption data registered in 2020, 2021 and 2022 for all such households. TBW and SWW, in particular, provided us with quarterly energy consumption data over the same periods (measured at the time intervals reported in Table 5). For the SWW households, in some cases the same physical household was associated to two smart meters, one measuring the consumption of the heat pump and one measuring the consumption of lighting and electric devices. The utility company verified that we could access the full amount of consumption data for each household.

Longitudinal energy consumption data were linked to all treatment group households with accepted application that effectively used the SPP app at least once –namely, the households that at least accessed the app once to register their in-app account. The choice not to consider households that



never logged on to the SPP app is made to avoid households that have not received treatment with the SPP Community Energy Challenge (case of total non-compliance) being accounted for in the estimates of effect. Table 8 shows that, of the 221 households that were accepted for participation in the SPP Community Energy Challenge, a total of 172 households (78%) accessed the app at least once and were successfully linked to energy consumption data provided by utilities. On the other hand, a total of 4'851 non-treated households with complete data on their energy consumption over time were identified by the utilities, in line with the requirements on the presence of electric heating/heat pumps and gas/electricity smart meters (Table 4).

Table 8: Number of treated and untreated households that respect requirements for participation and for which complete energy consumption data is available, for each region.

Utility	Number of treated households	Number of untreated households	Total number of households
TBW	20	777	797
SWW	83	992	1'075
EKS	69	3'082	3'151
TOTAL	172	4'851	5'023

4.4.1 Computation of baseline energy consumption data

For the cases of EKS and TBW households, yearly consumption data covers the period January 1 – December 31. For the case of SWW households, instead, yearly consumption is not always aligned with the calendar year, and the start of the consumption reading period is shifted of up to two months. For some households, in fact, reading of the yearly consumption data starts in the month in November, for some others in December (rather than January). Moreover, in some cases readings are not always performed on a quarterly basis: in some cases the duration of the period from one reading to the next one is not always exactly three months. Figure 13 shows the reading periods (blue area) we identified among the 172 households of the treatment group, compared to the posttest yearly time window (grey area), by reporting the number of households characterised by each reading period.

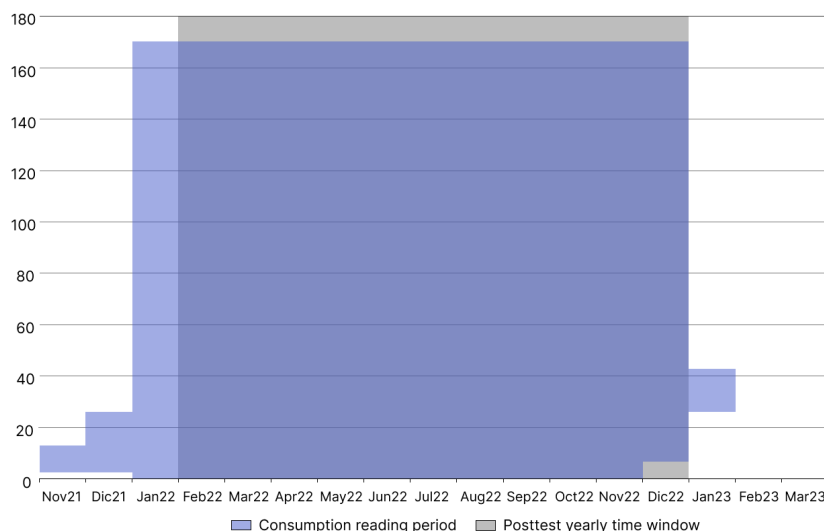


Figure 13: Number of treated households by reading period (blue area) compared to posttest yearly time window (grey area).

What all time windows have in common is that they start before the intervention period (grey area, January 1 – December 31 2022) by at least one month. None of the utilities has indeed provided us household energy consumption readings always conducted in February for all years under study (i.e.,



2020 to 2022). This means that we must consider as part of the posttest period at least one month of energy consumption by households prior to their exposure to treatment, namely the participation in the SPP Community Energy Challenge.

For some SWW households ($n = 23$), in particular, the reference reading for computing annual post-intervention consumption was even taken up to two or three months before the beginning of the SPP Community Energy Challenge (i.e., December or November 2021 respectively). Even if these households experienced an anticipation and, consequently, an extension of the monitoring period useful for estimating the annual impact of the intervention, the overall length of their monitoring window remained relatively similar to that of the other household, averaging 370 days compared to 361 for the rest of the treated group. Also, the variability found in the start dates of consumption reading periods is not critical for the ATT estimation under our DID design. In proportion, treated and untreated (control) households share similar patterns in the non-standard start dates (November 2021: $n_t = 10$, $n_c = 87$; December 2021: $n_t = 13$, $n_c = 282$) and end dates (November 2022: $n_t = 10$, $n_c = 107$; January 2023: $n_t = 17$, $n_c = 208$) of their energy consumption readings. This reinforces the likelihood of achieving a satisfactory degree of comparability in pre-intervention energy consumption trends between the two.

However, a relevant issue remains to be discussed with respect to the mismatch between the outcome reading dates and the transition from the pretest to the posttest period. Indeed, accounting for at least one full month of consumption that has not been influenced by the treatment implies that we are probably underestimating the overall effect of the intervention on pre-post changes in outcome across groups. Such an under-estimate could be especially relevant in the short run, for analyses of the effect over the first quarter of the year (months of January – March). That is, the influence of a pre-treatment month of energy consumption on the impact estimates decreases as the considered number of months of actual treatment increases. Such analyses, that we can only perform for the cases of SWW and TBW, will by definition result in an underestimation of the ATT, as they will account at most for two months of treatment (February and March), while the consumption in the preceding month(s) will not be affected by the intervention – and therefore we expect it is identical between treatment and control households. Furthermore, energy consumption for the winter months at the turn of the year is expected to be high, due to increased heating needs. This implies that we can expect an even smaller ATT due to the lack of exposure to intervention of the treated households in the month of the year when energy consumption is higher. The results of the impact assessment we present in the following sections need therefore to be regarded as a likely underestimation of the ATT, especially with respect to the short-term analyses conducted on a single quarterly basis.

4.4.1 Parallel trend assumption

The parallel trend assumption is critical to ensure internal validity of a DID model. To be fulfilled, it requires that in the absence of treatment, the average difference in outcome levels between the treatment and the control group is constant over time. Divergences in their outcome trends correspond to a violation of the assumption, leading to biased estimation of the causal effect (Wing et al., 2018).

In a setting with multiple pretreatment periods such as ours, it is possible to partially validate this common trend assumption by simply plotting average changes in the outcome by group and by searching for potential deviations across them in pretest periods. We thus perform a graphical inspection of the yearly energy consumption trends among the whole groups of treated and untreated (control) households, with a focus on annual average changes before the beginning of the SPP Community Energy Challenge.

The resulting trends in average energy consumption for 2020, 2021 and 2022 across treatment (T) and control (C) households are reported in Figure 14.a (total energy consumption, namely consumption for both heating and non-heating purposes). At first glance, it appears that there is no perfect alignment in group variations over time before the SPP Community Energy Challenge intervention (2020 and 2021). Treatment households started with a lower average level of energy consumption in 2020. In 2021, they increased their energy consumption more than the controls did: the percentage change for the



consumption for the control households is equal to +8.1%, while for the treated households it is equal to +13.4%. Such a misalignment appears even more clearly by considering consumption data over the different regions, and especially in the case of EKS (Figure 16.d). Here the growth in energy consumption for treated households in 2021 is more evident: their initial advantage over controls is more than halved (respectively, controls increased their consumption by 6.1%, while treated households increased by 12.8%). TBW-treated households, contrary to EKS, reported a slight increase in energy savings in 2021 compared to the control households (controls increased their energy consumption by 8.4%, while treated households increased by 7.2%, Figure 16.b), while there are no notable graphical changes in the annual trends for the SWW households (controls increased their energy consumption by 14.5% while treated households increased their consumption by 17.4%, Figure 16.c).

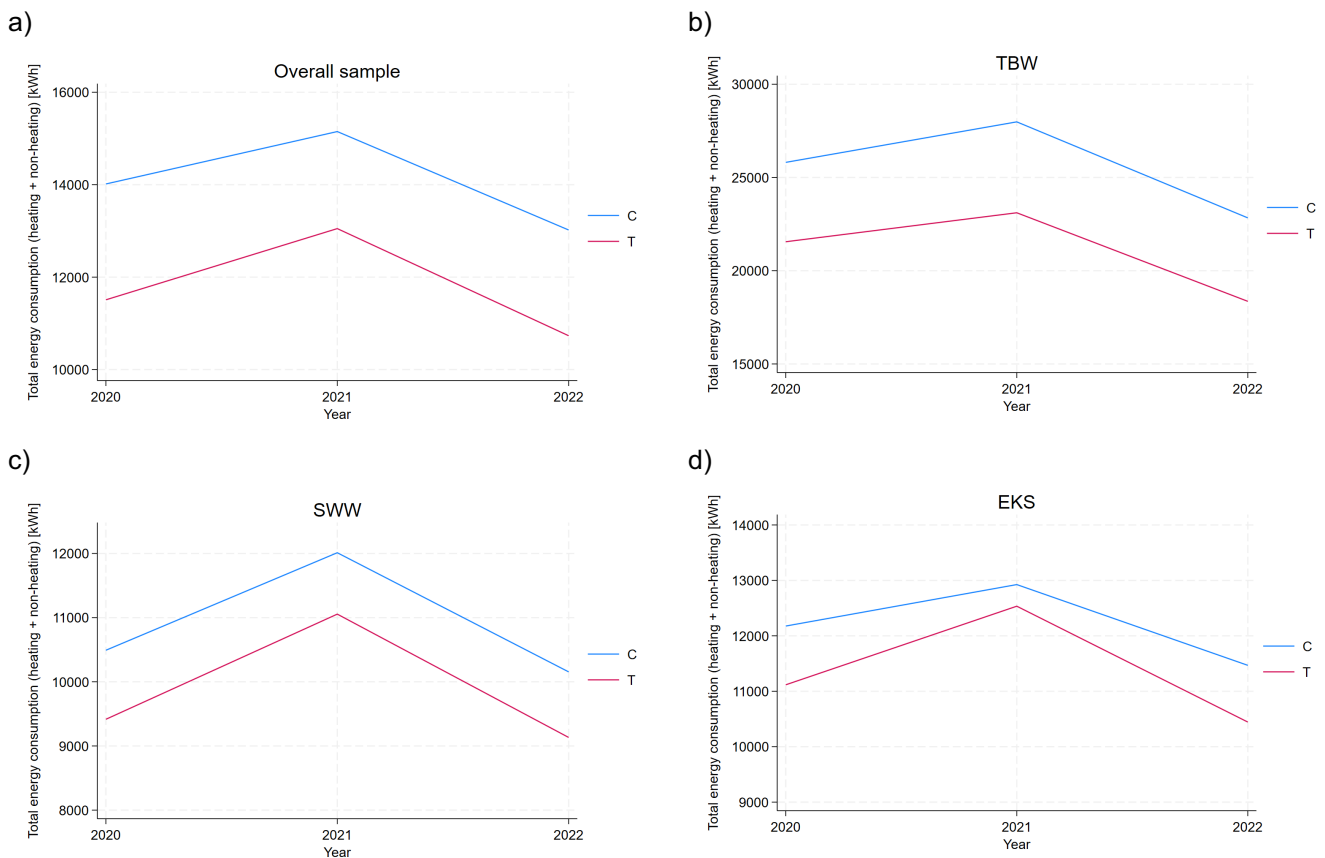


Figure 14: Evolution of yearly energy consumption (heating and non-heating purposes) in the treated (T) and untreated (C) households.

We also examined the relationship between the pretest average values and the pretest-posttest changes in household energy consumption, finding significant associations for both the treated and control groups. More in detail, model estimates resulted in significant negative coefficients for both the year 2020 (treated: $\beta = -.058$, $p < .001$; controls: $\beta = -0.69$, $p < .001$) and the year 2021 (treated: $\beta = -.087$; $p < .001$; controls: $\beta = -0.67$, $p < .001$). These findings highlight an additional issue towards the application of a standard DID estimation strategy to our data. To use such an approach it is indeed crucial that baseline outcome levels are unrelated to expectations for pre-post changes in outcome. Namely, if the treatment group is selected based on previous levels of the outcome, like in our case, there is potential bias from regression to the mean (Chay et al., 2005). If outcome pre-post changes are both influenced by pre-intervention household energy consumption, using standard DID estimation we can expect to obtain biased ATT point estimates. Balancing treatment and control group on pre-intervention levels in such a situation can greatly reduce bias (Ryan et al. 2015).



These last results, taken together with those that emerged from the visual trend analysis, hint at possible presence of multiple sources of bias, if a standard DID estimation approach is applied to our data. To properly account for this situation and reduce biases in the estimate of the treatment effect, we thus adopt a pretest longitudinal scheme. We consider baseline outcome values for the years 2020 and 2021 and use the propensity score matching (PSM) method before applying DID (difference-in-differences), to equate treated and controls on their baseline annual energy consumption data.

4.4.2 Propensity score matching algorithms and identification of the control group households

Due to the heterogeneity we found in pretest outcome levels and trends within and across the three EKS, SWW and TBW regions, we opt for matching treated and control households for each region separately. This strategy allows us to account for the clustered nature of the intervention, which encompasses regions of Switzerland characterized by differences in energy service provision and energy infrastructure (which results in differences in households' energy equipment) and, to a lesser extent, slight variations in exposure to meteorological conditions (see section 4.5). Besides optimizing the matching strategy by constraining the choice of control twins only within the same area of the corresponding treated, this choice allows to estimate the conditional treatment effects on the treated (CATT) by region, without incurring in –or, at least, minimizing– selection bias.

In line with recent simulation studies suggesting that balancing baseline outcome levels can significantly reduce estimates bias (Lindner and McConnel, 2018; Ryan et al. 2015), treated and control households are matched looking at all their available pretreatment annual energy consumption levels. All models are estimated by using the propensity score matching (PSM) method (Rosenbaum and Rubin, 1983). The propensity score is defined as the probability of being in the treatment group as a function of observed covariates (Austin 2011; Stuart 2010). Compared to controlling linearly for explanatory variables in a DID regression, PSM has the main advantages of condensing the observed covariates into a single measure used for matching, without imposing a predefined functional form on their relationship with the outcome (Rosenbaum and Rubin, 1983). Thus, it guarantees a more appropriate weighting of covariates by avoiding the risk of extrapolating twins beyond the region of common support (i.e., including treated households without suitable control twins in the analysis).

Given all pretest levels of annual energy consumption, we estimate propensity scores for each region with the following logistic regression model, that calculates the households' probability of being treated:

$$P(D_i = 1 \mid y_{it,t < t^*}) = \Lambda \left(\phi + \sum_{j=1}^{t^*-1} \pi_j y_{i,(t^*-j)} \right),$$

where D_i is an indicator equal to one if unit i is in the treatment group, Λ is the logistic cumulative distribution function, ϕ is the intercept and π is the coefficient for $y_{i,(t^*-1)}$, which is the last outcome before intervention begins for unit i .

Treatment and control households are matched by exploiting several algorithms, searching for the optimal trade-off between bias and efficiency (Caliendo and Kopeinig, 2008). For each of the three EKS, SWW and TBW regions, we separately test the following set of matching algorithms with a fixed caliper threshold of 0.02 for propensity score standard deviation:

- one Nearest Neighbor (1-NN) without replacement of controls;
- one Nearest Neighbor (1-NN) with replacement of controls;
- two Nearest Neighbors (2-NN) with replacement of controls;
- three Nearest Neighbors (3-NN) with replacement of controls;
- radius.

The Nearest Neighbour (NN) matching estimator simply select units from the control group as twins for a treated unit based on their closeness in terms of propensity score. We consider several variants of



NN matching, characterized by alternative approaches to replacement and the number of control twins to be extracted. We first use a NN matching without replacement, that allows untreated units to be considered only once. For the remaining NN models we instead allow control units to be freely selected more than once. This implies that weights can be used, and for a single treated household more than one twins are possible (respectively one, two, and three for each treated household, on varying the matching algorithms). Finally, we remove numeric constraints on the selection of controls and conduct a radius matching, that retains all units that lie within the predefined propensity range of 0.02 identified by caliper.

Table 9 below summarizes several measures of the balancing between treated and control households on pretest outcome levels before and after each PSM analysis. The balance is first checked for annual consumptions through t-test for equality of means in the treatment and control groups, both before and after matching. For good balancing, all t-tests should be non-significant after matching. We also report the standardized bias and percentage absolute bias (Rosenbaum and Rubin, 1985). The former indicates a successful matching when all variables obtain values lower than 5%, while the latter is a global diagnostic indicator, which is useful for comparing the overall bias detected between matched and unmatched samples or samples matched using different algorithms (Caliendo and Kopeinig, 2008; Zhang et al., 2019). Finally, we report sample sizes for treatment and control groups before and after matching. This allows us to assess which matching algorithm maximizes common support, i.e., to compare the degree of overlap in the propensity score distribution between the treatment and control groups.

Our results for all three regions suggest that matching procedures shrink the differences observed in average energy consumption across unmatched households and, consequently, reduce standardized biases for both pretest years (2020 and 2021). However, only some of the selected matching algorithms contribute effectively to reducing the bias well below 5%, which is the ideal condition suggested by the literature. This is especially the case of the 1-NN algorithm with replacement for SWW, which results in standardised biases under such threshold for both pretest measures of annual energy consumption, and in a reduction in absolute bias of more than 20% compared to the unmatched sample. EKS and TBW households are instead better matched by the 3-NN algorithm with replacement, which brings the standardized bias for EKS to zero and to values below 2% for TBW, leading to an overall bias reduction of 24% and 46% respectively.

Table 9: Performances of alternative PSM algorithms in balancing pretest outcome levels by region: average levels, percentage biases and common support before and after matching. Results of the best performing algorithm are reported in bold.

Region	Year	Performance indicators	Before matching	After matching					
				1-NN (nr)	1-NN	2-NN	3-NN	Radius	
EKS	2020	$M_{Treated}$	11.116	11.116	11.116	11.116	11.116	11.116	
		$M_{Controls}$	12.177	11.271	11.271	11.146	11.117	11.794	
		p-value	0.180	0.880	0.880	0.976	1.000	0.534	
		% bias	-16.6	-2.4	-2.4	-0.5	-0.0	-10.6	
	2021	$M_{Treated}$	12.535	12.535	12.535	12.535	12.535	12.535	
		$M_{Controls}$	12.925	12.746	12.746	12.575	12.531	13.037	
		p-value	0.636	0.838	0.838	0.969	0.997	0.653	
		% Bias	-6.0	-3.2	-3.2	-0.6	0.1	-7.7	
	% overall bias			24.3	3.5	3.5	0.7	0.1	10.8
	$n_{treated}/n_{controls}$			69/3.082	69/69	69/69	69/138	69/203	69/3.079
SWW	2020	$M_{Treated}$	9.416	9.507	9.507	9.507	9.507	9.507	
		$M_{Controls}$	10.492	9.428	9.469	9.446	9.428	9.751	
		p-value	0.044	0.889	0.947	0.913	0.889	0.657	
		% bias	-25.4	1.8	0.9	1.4	1.8	-5.8	



	2021	M _{Treated}	11.055	10.885	10.885	10.885	10.885	10.885	
		M _{Controls}	12.011	10.622	10.772	10.675	10.590	11.230	
		p-value	0.116	0.675	0.855	0.734	0.635	0.592	
		% Bias	-19.6	5.4	2.3	4.3	6.1	-7.1	
	% overall bias		25.4	10.2	4.3	8.6	12.2	8.4	
	n _{Treated} /n _{controls}		83/992	82/82	82/77	82/151	82/217	82/990	
TBW	2020	M _{Treated}	21.554	21.554	21.554	21.554	21.554	21.554	
		M _{Controls}	25.812	21.442	21.442	21.764	21.641	23.596	
		p-value	0.107	0.953	0.953	0.917	0.966	0.408	
		% bias	-46.5	1.2	1.2	-2.3	-0.9	-22.3	
	2021	M _{Treated}	23.109	23.109	23.109	23.109	23.109	23.109	
		M _{Controls}	27.984	23.067	23.067	23.214	23.165	25.475	
		p-value	0.086	0.985	0.985	0.962	0.980	0.387	
		% Bias	-48.5	0.4	0.4	-1.0	-0.5	-23.5	
		% overall bias		48.3	4.1	4.1	6.0	2.0	27.7
		n _{Treated} /n _{controls}		20/777	20/20	20/20	20/38	20/58	20/777

For each region, results of the best performing matching algorithms are reported in bold characters. We thus chose the control groups in each region according to the best performing algorithm. Only 1 treatment observation did not meet the common support condition in the region of SWW, with a negligible loss of information on treated households in the matched samples. Based on these results, we can confirm that the NN matching strategies succeeds in balancing the covariates, thus making the common trend assumption more plausible, without significantly altering estimates consistency outside the common support. We thus construct our final analytic sample by appending the three matched samples obtained through the 1-NN algorithm for SWW and the 3-NN algorithm for EKS and TBW, respectively. Our impact evaluation of the SPP Community Energy Challenge is thus performed by considering the combination of the subsamples of treated and control households identified in each region, adopting proper weights according to the best performing matching procedures that emerged in each region.

4.5 Heating and cooling degree-days

The evolution of energy consumption data is influenced by a number of external factors related with the meteorological conditions, such as the duration of the day/of the night, the presence of clouds, the external temperature. When comparisons over time need to be made, summary indicators that are usually considered to assess the level of energy demand that is needed to heat or cool buildings coincide with the “degree-days” indicators. Sometimes these indicators account for the energy demand for both heating and cooling purposes, sometimes they separately account for the energy demand for heating and cooling.

In Switzerland, these indicators are defined by the SIA (Swiss Society of Engineers and Architects) 381/3 standard. Basically, if the average daily temperature at a location is below 12 °C, it is assumed that there is a need for heating and therefore, that day has to be accounted as a heating day. The amount of heating energy that is needed during that day is supposed to be directly proportional to the difference between a reference indoor temperature (for Switzerland set to 20 °C) and the outdoor temperature. The “heating degree-days” (HDD) are thus defined as the difference between such a reference indoor temperature and the average daily outdoor temperature, during heating days. For instance, if the average temperature of the day is 8°, the number of heating degree-days for that day is 12. If instead the average temperature is above 12°, the number of heating degree-days for that day is 0. HDDs can be summed over months or years, and can be used to easily identify the different amount of energy that is needed for heating reasons for such periods. A similar procedure is used to estimate

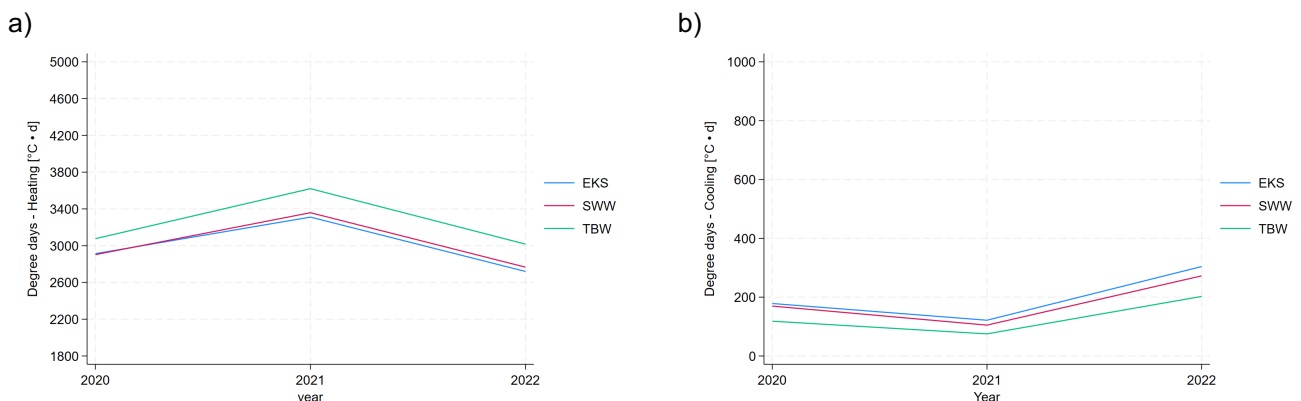


the cooling degree-days (CDD), which instead refer to the amount of energy needed for the purpose of cooling buildings.

Outdoor temperatures are regularly collected for a large number of locations throughout Switzerland by the Federal office of Meteorology and Climatology Meteoswiss. For the regions of the SPP intervention, long historical series of average daily temperatures can be directly downloaded from the OpenWeather online database (<https://home.openweathermap.org>, last accessed on June, 19 2023). In order to include heating and cooling degree-days into our analyses, we consider historical data on the cities of Winterthur (lat: 47.49882; lon: 8.723689), Wil (lat: 47.46356; lon: 9.04791), and Schaffhausen (lat: 47.69589; lon: 8.638049), and compute monthly heating degree-days (HDD) and cooling degree-days (CDD) according to the SIA 381/3 standard for years 2020, 2021, and 2022. The yearly values we obtain can qualitatively help to understand the evolution of yearly consumption data over the three regions EKS, TBW and SWW.

Figure 15 reports the evolution of the HDD (a), CDD (b) and TDD (c, total degree days, namely HDD + CDD) indicators over time. If we consider TDD, which we regard as a proxy for the overall evolution of the energy demand due to outdoor weather conditions, it appears that the energy demand was higher in year 2021 in all regions, while years 2020 and 2022 had lower energy demand (nearly comparable between the years). To get further insights on the evolution of the degree-day indicators over time, and identify (dis)similarities in temperature patterns among the years we are considering, we separately compute these indicators for each quarter. The result, reported in Figure 16, shows that the lower number of HDD registered in 2022 compared to 2021 is due to higher temperatures in Fall and Winter months. Also, year 2022 shows higher numbers of CDD during Summer months. Furthermore, year 2020 shows a much more similar trend in TDD to year 2022 than to year 2021.

Overall, the evolution of the TDD indicator over the years is consistent with the “inverted u-shaped” consumption trend that appears by plotting the 2020 – 2022 yearly consumption data by the households of the three regions, which is reported in Figure 14.d. The charts reported in the figure clearly indicate that, the evolution on consumption we observe among the three years is driven by the evolution in weather conditions, summarised by the degree-days indicator. Figure 16 reports the evolution of average yearly energy consumption per region, by considering the weather correction based on the TDD indicator (yearly consumption divided by TDD). The plot shows that in the region of TBW a decreasing trend in consumption is already taken place (with -6.3% and -12.90% temperature-corrected consumption respectively between 2021 and 2020 and 2022 and 2020). In the region of EKS, such a decreasing trend is also occurring, though it is less marked (respectively, -4.3% and -3.7%). In the region of SWW, instead, weather-corrected consumption is more stable (respectively, +1.7% and -2.1%).



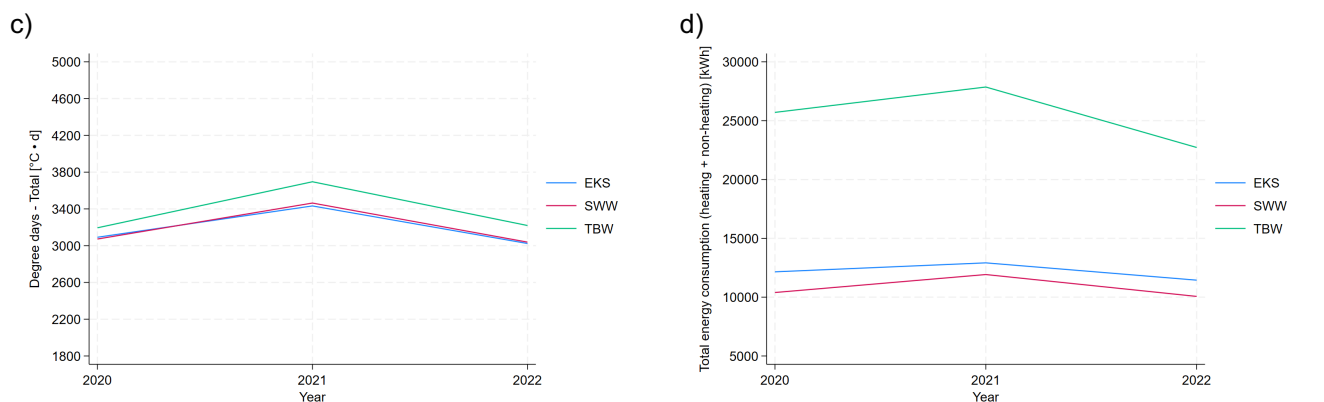


Figure 15: Evolution of the degree-days indicators (per region) over the three years we consider to estimate the effect of the SPP Community Energy Challenge, compared with the evolution of the energy consumption measured for all households (again, per region).

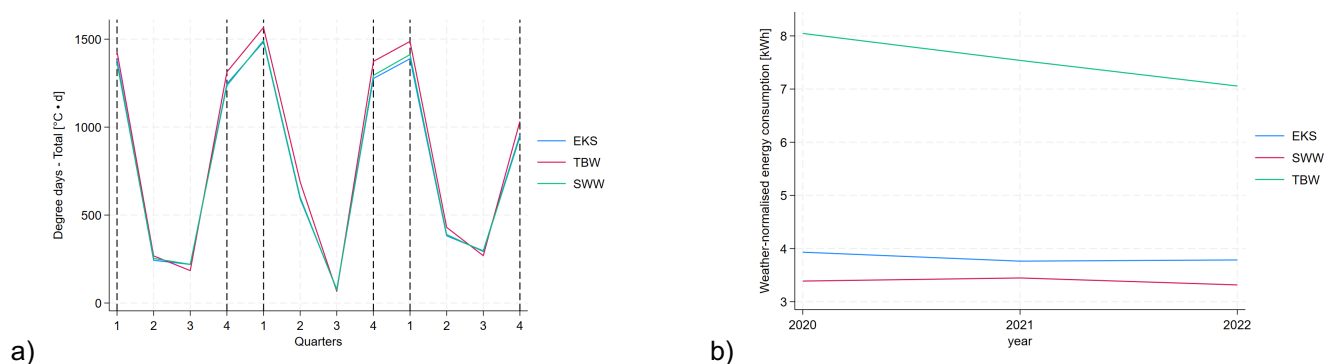


Figure 16: a) Quarterly evolution of the degree-days indicators per region. b) Weather-normalized energy consumption, obtained by dividing the yearly energy consumption by the yearly total degree-days, in each region.

4.6 Impact evaluation

Our main evaluative objective is to counterfactually assess the effectiveness of the SPP app and Community Energy Challenge in stimulating tangible energy savings in the participating households. Namely, we aim to estimate the effect of the SPP intervention on households' energy consumption –our primary outcome– net of changes in control group households. SPP is specifically designed to overcome the limitations of only theory-driven approaches previously used to develop energy-saving behaviour interventions, by incorporating user knowledge and preferences through a living lab approach. Furthermore, SPP aims to improve the scope and quality of the feedback offered to households regarding their energy consumption, through the provision of energy consumption feedback broken-down in consumption for heating and non-heating purposes. Based on such methodological and content-related innovations, we expect that:

H₁. The SPP Community Energy Challenge has a significant effect on reducing the annual energy consumption of the treated households compared to the control group.

Beyond this general research purpose, we are also interested in testing whether the impact of the SPP Community Energy Challenge differs between regions by considering sub-samples of the participating households based on their utility company or on their heating system (households with heat pumps in the regions of EKS and SWW compared with households with gas heating in the region of TBW). The idea behind this is to understand whether the SPP Community Energy Challenge affects participants' energy consumption in the same way regardless of energy supply services and facilities available to them. This allows to gather helpful information on the scalability potential of the SPP Community Energy Challenge across heterogeneous user market segments. We expect that:



H₂. There are no significant variations in the average effect of the SPP Community Energy Challenge between the three regions respectively served by the EKS, SWW, and TBW utility companies.

In line with this hypothesis, we also expect that:

H₃. There are no significant variations in the average effect of the SPP Community Energy Challenge between heat pump households compared with gas heating households.

Finally, we look at quarterly changes in household energy consumption to assess whether the effect of the SSP Community Energy Challenge is maintained over time. Previous research in this field suggests that app-based interventions aimed at promoting long-lasting pro-environmental behavioral changes often fail to achieve the desired result. While being successful in reducing electricity consumption in the short term, previous real-world tests of app-based energy savings challenges resulted in serious downward trends after their end, in the long-term (e.g., Wemyss et al., 2019; Cellina et al., 2024). To face these issues, besides the “high intensity period”, the SPP Community Energy Challenge also included an eight-month “low-intensity” period for behavioural maintenance and follow-up. Based on this treatment approach, we expect that:

H_{4a}. The effect of the SPP Community Energy Intervention on household energy consumption remains constant throughout the first year of implementation.

Conditional to an eventual rejection of *H_{4a}*, we also investigate alternative explanations for changes in the magnitude of the treatment effect over time, especially looking at seasonality. Seasonal meteorological variations have a strong influence on energy consumption for heating and cooling homes. Considering that some of the energy-saving behaviour suggested by the SPP Community Energy Challenge precisely focus on reducing heating and cooling consumption, we might expect to find an increase in the average treatment effect in the months of greater exposure to hot or cold temperatures. In light of that, we consider the following final hypothesis:

H_{4b}. Seasonality explains possible changes in the magnitude of the effect of the SPP Community Energy Intervention in the months after its implementation.

We thus test all these hypotheses in a two-group longitudinal quasi-experimental framework, whose analytical strategy is described in detail in the following section.

4.6.1 Analytic strategy

We estimate the average treatment effect on the treated (ATT) due to participation to the SPP Community Energy Challenge by means of panel regression models (Wooldridge, 2010), which allow to account for repeated measurement periods over time (Sergici and Faruqui, 2011). Specifically, following Wooldridge (2015) we use a Fixed Effects panel data regression estimator, which is among the most used panel data estimators. Differently than simple regression models, which produce unbiased estimates only provided that the “zero conditional mean assumption” is met¹, fixed effects allow correlation between the error term and the independent variable. They remove the unobserved effect by means of differencing between adjacent time periods, and then allow estimate of the model coefficients via an ordinary least square (OLS) estimation. Such a process allows to get rid of the unobserved heterogeneity between households that is constant over time. This characteristic is particularly relevant for our case, in which households participating to the SPP Community Energy Challenge are self-selected. Due to self-selection, it is likely that a correlation exists between the independent variable of the model (the dichotomous variable indicating whether a household was treated or not) and the unobserved variables that characterise the self-selected households of the treatment group. In particular, we cannot exclude that households that decided to join the SPP Community Energy Challenge have higher environmental attitudes than average households in the same region, and/or that they have higher than average education levels (and thus possibly also

¹ It assume that the expected value of the error term is zero for any value of the independent variable; namely, that the error term is not correlated with any of the independent variables in the model.



earnings), which makes them more inclined to actively join a research project. The Fixed Effects estimator is therefore well-suited to the SPP case —much more than the Random Effects, the other widely used estimator for panel data regressions, whose use is instead advised (and more efficient) when one thinks the unobserved effect is uncorrelated with any of the explanatory variables.

Among the family of Fixed Effects estimators, we specifically opt for the Two Ways Fixed Effect (TWFE) estimator (Wooldridge, 2021). This model is frequently used in evidence-based policy-making, since it allows to include both unit and time fixed effects in ordinary least squares estimation. Namely, it allows to remove the effects due to both unobserved specific characteristics of the households and secular changes in the external context that affect all units in the same way at the timing of the intervention, such as for instance meteorology and weather factors (the evolution of outdoor temperatures that drives the need for home heating of cooling, e.g. the degree-days). The general shape of TWFE regression models is as follows:

$$y_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + \mathbf{c}_i + f_t + u_{it}$$

for $t = 1, \dots, T$ and $i = 1, \dots, N$,

where:

- y_{it} is the observed dependent variable (here, the log transformation of energy consumption², which varies between the households and over time);
- \mathbf{x}_{it} is the vector of observed independent explanatory variables. Here, it only consists in a dichotomous variable indicating if the household i received the SPP Community Energy Challenge treatment at time t . This varies among the households but it is time-invariant;
- \mathbf{c}_i is the vector of unobserved household-specific effects, which are time-invariant (household fixed effects);
- f_t is the vector of observed time-specific effects, which are constant across the households though vary over time (here, for instance these may be the “Heating degree days” or “Cooling degree days” variables);
- u_{it} is the unobserved idiosyncratic error term, which varies between the households and over time;
- t is the subscript for the year, varying from 1 to T , which is equal to 3 when the model is used to estimate the impact over a year (years 2020, 201 and 2022) and to 12 when the model is used to estimate the impact over quarterly periods (four quarters per three years);
- i is the subscript for the household, varying from 1 to N , which is equal to the sum of the households of the treatment and control group.

Through the estimation of panel data regression models on our analytic sample using TWFE and PSM weights, we can therefore quantify the causal effect of SPP accounting for:

- violations of the parallel trend assumption;
- unobserved variables affecting the observed behaviour (energy consumption) of households of the treatment and control group;
- the evolution of external factors that drive energy demand for heating, cooling, use of appliances and lighting, affecting all households at the same conditions.

More specifically, we empirically test H_1 on the annual energy consumption data for all regions by fitting a TWFE panel regression model in which the treatment interacts with a dummy variable identifying the pretest and posttest periods, with PSM weights and robust standard errors at the household level to account for autocorrelation within each household over time.

² We use the log transformation of the energy consumption values (y), as the difference in the natural logs over time represents the percentage change in y from period to period (Wooldridge, 2015).



H_2 and H_3 , respectively focussing on the analysis of conditional average treatment effects by region and heating system, are separately assessed by introducing two additional interaction terms to the treated condition in H_1 model specification: respectively, the household region (0 “EKS”; 1 “SWW”; 2 “TBW”) and the heating system (0 “heat pump”; 1 “gas boiler”).

For the analyses on the evolution of treatment effects over time aimed at testing $H4a$ and $H4b$, we consider quarterly energy consumption data, that are only available for TBW and SWW households. The quarterly consumption for all TBW households always refers to the following fixed periods: January-March, April-June, July-September, and October-December. Instead, as anticipated in section 4.4.1, SWW households have different time windows for energy consumption, characterised by internally inconsistent starting dates and lengths. To account for these inconsistencies, we only consider SWW households characterised by the same quarter readings for 2021 and 2022, with the first quarter term for 2022 from March onward. SWW Households characterised by a first quarter reading ending in February (the same month of the treatment start), or missing one of the quarterly readings during the period of interest (2021-2022), are excluded from the analysis on the wrong or missing quartile, respectively. Adopting this selection strategy finally resulted in the exclusion of 6 out of 82 treated households and 6 out of 77 controls from the SWW subsample, for a total of 12 excluded cases.

Before estimating quarter-based conditional treatment effects and testing $H4a$ and $H4b$ hypotheses, we also re-evaluated the tightness of the balance obtained from previous PSM analysis on the new, restricted version of the SWW sample. We computed the same set of performance indicators used in section 4.4.2 on the restricted sample, to assess the balance between treated and control households before and after matching on the entire analytical sample. Despite the loss of 12 cases, we found no noteworthy deviations in pretest outcome levels across the treatment and control group: we still obtained standardised bias values under the 5% threshold for both pretest measures (% bias₂₀=2.0; % bias₂₁=1.9) and a relatively small absolute bias of 8.6.

We thus tested $H4a$ by fitting a regression model in which treatment interacts with both a dummy variable identifying pretest-posttest periods and a categorical variable distinguishing the same quarters over the years (0 “Jan-Mar”; 1 “Apr-Jun”; 2 “Jul-Sep”; 3 “Oct-Dec”), with PSM weights and robust standard errors at the household level to account for autocorrelation within each household over time.

Finally, we estimated $H4a$ by substituting the quarter interaction term with a categorical variable summarising tertiles of the distribution of quarterly average of total degree-days (TDI) for the SWW and TBW regions (0 “Lower tertile”; 1 “Central tertile”; 2 “Higher tertile”). Table 11 provides a summary description of the number of household observations by quarters and the total degree-days distribution.

Table 102: Number of household observations by quarters and total degree-days distribution.

Total degree-days (TDD)		Quarters				Total
		Jan-Mar	Apr-Jun	Jul-Sep	Oct-Dec	
Tertiles	Lower	0	509	1527	0	2036
	Central	0	1018	0	1212	2230
	Higher	1527	0	0	315	1842
Total		1527	1527	1527	1527	6108

4.7 Estimate of treatment effects

4.7.1 Effects on annual household energy consumption

We first performed the weighted panel regression model to estimate the effect of participation in the SPP Community Energy Challenge on the annual household energy consumption, by accounting for all three EKS, SWW, and TBW regions. As shown in the first row of Table 11, estimating the model on the entire analytic sample of households with a common support indicates that the treatment significantly results in a 4.9% ($p = .030$) reduction in annual energy consumption compared to the previous years.



Note, this estimate already accounts for the evolution in outside temperatures, namely for the Total degree-days (TDD) indicator that characterises every year. In fact, in each region both households of the treatment and control group are exposed to the same weather conditions. As use of the DID estimator means we are looking for the differences between the treatment and the control group of the same region in the differences between energy consumption of the same years, the difference in the energy consumption of these two groups across the years is already net of changes in the TDD indicator. This means, we can confirm H_1 about the expectation of an overall significant impact of the intervention on the reduction of households' annual energy consumption.

We then tested whether the households' reference region of residence significantly moderate the effect of being treated on their annual energy consumption. The idea behind that is to evaluate whether specific characteristics related to energy demand, provision and infrastructures that are common across households located in the same regions, can somehow moderate the magnitude of the estimated ATT. Our findings, also reported in Table 11, suggest that outcome trajectories over time are not differentially affected by reference regions. All the estimates of conditional average treatment effect (CATT) are indeed negligible in magnitude and statistically non-significant, confirming our expectations around H_2 and the null moderating effect of regional level characteristics, which can therefore be retained. Similar results are also obtained for the third panel regression model whose results are reported in Table 11, that includes the additional interaction term to compare treatment effects across households characterised by different heating systems. We used households with heat pump (namely households in regions EKS and SWW) as reference group and compared them to households with gas heater (namely households in TBW region). Again, we found no significant differential effects between the two groups. This result allows us to also retain also H_3 and to confirm the absence of significant moderators of the treatment effect depending on the type of heating system available to the household.

Table 11: Treatment effect on annual household energy consumption: overall effect and heterogeneity by region and heating system.

		N _t /N _c	ATT (SE)	p-value	CATT (SE)	p-value
Analytic sample		171/338	-0.049** (0.022)	0.030	-	-
Region	EKS	69/203	-0.049 (0.043)	0.249	-	-
	SWW	82/77	-0.043 (0.027)	0.107	0.006 (0.050)	0.906
	TBW	20/58	-0.067 (0.043)	0.120	-0.018 (0.060)	0.768
Heating system	Heat pump	151/280	-0.046* (0.024)	0.059	-	-
	Gas boiler	20/58	-0.067 (0.043)	0.120	-0.021 (0.049)	0.668

*pvalue≤0.10; **pvalue≤0.05; ***pvalue≤0.01

Finally, we performed a sensitivity analysis by replicating all the above analyses by using the annual energy consumption for year 2021 as the only baseline outcome level to obtain the DID estimates. The results of this sensitivity analysis confirm a high level of stability in the magnitude and direction of the impact estimates, which do not significantly vary by considering one or two reference years in our model specification.



4.7.2 Evolution of the average treatment effect over time (only for SWW and TBW regions)

This stage of analysis deals with the heterogeneity of treatment effects as a function of the time elapsed since the start of the SPP Community Energy Challenge. First, we test the H_{4a} hypothesis about the maintenance of the effect of the intervention on energy consumption over the first year since its start. Conditional to the rejection of H_{4a} , we also test the additional H_{4b} hypothesis to explain the possible heterogeneity of the treatment effect over time. H_{4b} focuses on the potential role played by seasonality as a contributor to differences in the estimated impact depending on the season (and particularly, as a contributor to differences between the hot summer months and the cold months).

The results are reported in Table 12. Also in this case, the estimate already accounts for the evolution in outside temperatures, namely for the Total degree-days (TDD) indicator that characterises every year. As treatment and control households in each region are exposed to the same weather conditions, the difference in the energy consumptions of the two groups across the years is already net of changes in the TDD indicator.

The first panel regression model allocates all quarters in a categorical interaction term, only differentiating between pretest and posttest periods. We use this preliminary model specification only to assess the consistency between the ATT estimates produced by using quarterly and annual consumption data. The result goes in this direction, showing a 4.8% more reduction in the quarterly energy consumption of treated households. This ATT estimate is strictly aligned with the yearly estimates made on the integral analytical sample, reported in Table 11.

Then, we consider another model specification aimed at investigating the evolution of the treatment effect over time, by substituting the period interaction term with a categorical variable that identifies the four quarters across the years. Before moving onto the interpretation of model results for single quarters, however, it is essential to point out that – due to a one-month mismatch between the energy consumption reading dates and the transition from the pretest to the posttest period – treatment effect estimates on the first quarter are most likely underreported (see section 4.4.1). For this reason, here we only report and discuss CATT estimates from the second quarter onwards.

The model results confirm that the impact of the SPP Community Energy Challenge varies in magnitude depending on the quarter of reference, with a gradual decline as the quarters pass. That is, we found a noticeable average treatment effect for the period Apr-Jun, corresponding to a 11% ($p = .012$) reduction in household energy consumption. The Jul-Sep quarter shows a substantial –albeit not significant– drop in the estimates compared to the former, while the last quarter of 2022 even brings to a null average treatment effect. Of particular interest for our analysis is the difference found in the CATT of the second and fourth quarters ($\beta = .112$; $p = .086$). Albeit slightly above the 0.05 significance level, this result further strengthens the interpretation of a downward trend in the impact of the SPP Community Energy Challenge on energy consumption over the first year after its launch, leading us to reject H_{4a} .

Table 12: Treatment effect on quarter household energy consumption for SWW and TBW regions: overall effect and heterogeneity by quarter and degree-days distribution between quarters.

		N _t /N _c	ATT (SE)	p-value	CATT (SE)	p-value
Restricted analytic sample		102/135	-0.048** (0.023)	0.041	-	-
Quarters	Quarter 1 (Jan-Mar)	102/135	-0.022 (0.039)	0.574	0.095** (0.047)	0.045
	Quarter 2 (Apr-Jun)	102/135	-0.110** (0.043)	0.012	-	-



	Quarter 3 (Jul-Sep)	96/129	-0.063 (0.051)	0.220	0.067 (0.058)	0.252
	Quarter 4 (Oct-Dec)	96/129	-0.009 (0.044)	0.842	0.112* (0.065)	0.086
TDD (quarter average)	Lower tertile	102/135	-0.077 (0.047)	0.104	-	-
	Central tertile	102/135	-0.036 (0.035)	0.308	0.043 (0.045)	0.340
	Higher tertile	96/129	-0.039 (0.036)	0.277	0.035 (0.054)	0.522

* pvalue≤0.1; **pvalues≤0.05; ***pvalues≤0.01

We finally estimated a final model to test H_{4b} , namely whether variations in total degree-days (TDD) depending on the season significantly moderate the effect of the treatment on household energy consumption over one year. Even if regression coefficients reported in Table 12 suggest that the impact of the SPP Community Energy Challenge is higher in the quarters when households had to be heated/cooled less (the lower tertiles TDD quarter), the estimated differences across ATTs conditional to the distribution of degree-days are not significant. These results suggest to also reject H_{4b} . We thus conclude that the downward trend in the magnitude of the treatment effect we found over the first year after the launch of the SPP Community Energy Challenge is not due to seasonal effects/is not affected by variations in energy consumption due to seasonality.

Again, results of the sensitivity analysis conducted by replicating all the above models by using the annual energy consumption for year 2021 as the only baseline outcome level, confirm a high level of stability in the magnitude and direction of the impact estimates.

4.8 Discussion

The above results show that the SPP Community Energy Challenge produces statistically significant effects in terms of a reduction in energy consumption. The elements collected through the three surveys allow us to verify if the households' self-reported routines are consistent with the ATT we found, and also to discuss which specific changes in the households' routines were involved in the observed decrease in consumption. The survey materials, combined with the data automatically collected by the in-app analytic system, as well as with elements that emerged during the online meetings with participants to the SPP Community Energy Challenge, allow us to discuss which specific features played a role in driving the effect we found.

4.8.1 Changes in household routines

To discuss whether and how changes in households' routines took place, we consider the set respondents to the three surveys performed before (T1), during (T2), and after (T3) the SPP Community Energy Challenge. Before analysis, all survey responses were cleaned to remove those by households that never logged into the SPP app. This is coherent with the choice to remove the set of households that never logged on the SPP app from the sample of treated households (see Section 4.4). Furthermore, to analyse the evolution of self-reported routines over time, we only consider the analytical sample of respondents to all the three survey waves. Response rates on the three survey waves largely differed: the baseline survey at T1 had 181 complete responses by registered app users, while surveys at T2 and T3 respectively had 135 and 118 complete responses by registered app users. By considering



the sub-set of those that responded to all three survey waves, we obtain an analytical sample of 96 respondents.

Survey respondents are mostly males (73%), middle-aged (average age is 53 years old), and highly educated (more than half have a university degree). The composition of the households is quite balanced, with 41% consisting of families with kids and the remaining share consisting of single adults, of which 27% are older than 65 years and 32% are younger than 65 years.

In a general question of the T2 survey investigating whether respondents had tried new energy saving activities at home, 54% of them answered affirmatively. Indeed, One-Way repeated measures ANOVA tests comparing self-reported routines data collected at T1, T2, and T3 (Table 13) show statistically significant decreases in most energy consuming routines performed in the household. Particularly, a significant decrease appears in self-reported living room temperature settings, characterized by medium to large effect size (measured through the Cohen's d parameter), respectively in the comparison between T1 and T2 and T1 and T3. Also, statistically significant decreases are found in the self-reported data about use of the oven in cooking practices, use of the tumble dryer in washing practices, and use of computers, most likely for recreational practices. Furthermore, statistically significant effects appear in the self-reported use of the dish-washer but they only appear in the short-term, and in the self-reported use of the washing machine, but they only appear in the long-term. This result is probably due to the combined presence of SPP and of the communication campaign promoting energy saving in households, launched by the Swiss Federal Office of Energy in August 2022, namely exactly between T2 and T3. The campaign recommends five ways to avoid wasting energy (turn the heating down, cook with lids on pans, turn lights off, switch off equipment the right way, take a shower instead of a bath), which are strictly connected with the challenges offered by the SPP app.

Table 13: Household energy consumption self-reported routines before, during and after participation in the SPP Community Energy Challenge, at T1, T2, and T3.

Survey question	T1			T2			T3			Effect size (Cohen's d)		ANOVA	Post-hoc tests (Bonferroni correction)	
	n	Mean	SD	n	Mean	SD	n	Mean	SD	T1 vs T2	T1 vs T3	p-value	Significant pairwise comparisons	Adjusted p-value
At what average temperature (°C) do you heat your living room during the day?	93	21.17	0.90	92	20.51	1.03	95	20.22	1.00	0.77 medium	1.07 large	3.2E-21 ***	T1 vs T2 T1 vs T3 T2 vs T3	2.39 E-10 *** 2.09E-16 *** 1 E- 3 ***
On average, how many showers do you take per week?	95	5.66	2.47	96	5.50	2.70	96	5.25	2.63	0.09 very small	0.17 very small	0.167	---	---
On average, how many baths do you take per week?	94	0.50	1.18	95	0.34	0.69	95	0.28	0.63	0.17 very small	0.25 small	0.045	T1 vs T3	0.053 *
How many times per week does your household use the oven?	96	3.90	1.91	96	3.42	1.63	96	3.25	1.78	0.27 small	0.37 small	8.17E-04 ***	T1 vs T2 T1 vs T3	0.027 *** 0.001 ***
How many times per week does your household use the dishwasher?	95	4.91	2.62	95	4.49	2.68	96	4.73	3.22	0.26 small	0.006 very small	0.133	T1 vs T2	0.036 **
How many times per week does your household use the washing machine?	96	3.93	2.20	96	3.69	2.05	96	3.68	2.15	0.16 very small	0.18 very small	0.168	T1 vs T3	0.264 *
How many times per week does your	86	2.37	1.92	86	1.97	1.68	84	2.02	1.65	0.29 small	0.29 small	0.006 ***	T1 vs T2 T1 vs T3	0.029 ** 0.027 **



household use the tumble dryer?															
On average, how many hours per day are TVs running in your home?	90	3.13	2.21	90	2.81	2.22	87	2.76	1.71	0.14 very small	0.22 very small	0.177	---	---	
On average, how many hours per day are computers running in your home?	96	6.92	6.42	93	5.88	6.02	95	4.19	4.14	0.23 small	0.52 medium	1.19E-05 ***	T1 vs T2 T1 vs T3 T2 vs T3	0.094 * 6.60 E-06 *** 0.032 **	
On average, how many hours per day are tablets running in your home?	79	3.54	4.82	79	2.96	4.87	80	2.84	4.65	0.11 very small	0.16 very small	0.399	---	---	

* pvalue≤0.1; **pvalue≤0.05; ***pvalue≤0.01

In all the cases for which statistically significant differences appear, however, the effect size is always small or very small – with the only exceptions consisting in the use of computers, whose change at T3 compared with T1 has a medium effect size, and in the change in the thermostat temperature setting in the living room. Indeed, the latter appears to be the most relevant change in the households' self-reported routines. Namely, it seems that energy saving for heating purposes played a major role in driving the overall reduction in energy consumption represented by the ATT.

Table 13 also shows that a decrease is reported in the frequency of weekly showers and baths. However, such decreases are not statistically significant. Even though a decrease in frequency is not the only way to sufficiency (e.g., shorter showers or lower temperature baths could also decrease consumption), the lack of statistical significance may hint at the difficulty of shifting cultural conventions on cleanliness-related topics. We cannot even exclude that, as survey responses are self-reported by household representatives that self-selected themselves, these results do not fully represent the diversity of ways in which energy consuming routines are performed within the household, by different household members. This might for instance be the case for typically gendered activities, such as a possible tendency of men to mostly focus on heating-related activities (including setting of the thermostat), and of women to mostly focus on washing routines. As most of the survey respondents were male, the women's point of view might be missing from our results. Future research might aim at exploring in more depth what happens within the household, by explicitly investigating the dynamics and relationships that take place between householders, when energy consuming routines are at stake.

Indeed, attempts to shift cultural conventions on such topics are at the core of many applied works inspired by social practice theories (Shove, 2003; Hand et al., 2005) and have already been tested in living lab settings with promising results (Godin et al., 2020). In the SPP case, the analysis of the pinboard posts, reported in the next Section (Figure 18), shows very limited discussion on social conventions around cleanliness and showering, which might have contributed to hinder such a shift. These results suggest that different intervention strategies should be enacted to favour more engagement by users in questioning current cleanliness-related social norms: attempts to shift them by simply launching challenges via an app like SPP turned out to be insufficient to ensure the householders' engagement. Particularly, future research might wish to explicitly engage more diverse household representatives, regarding their socio-economic characteristics. Managing to include both women and men, characterized by different education levels and age, would allow to clarify whether the apparent lack of questioning current social norms around cleanliness is actually due to app-based approaches or rather is a consequence of the specific composition of the household representatives actively engaged in the SPP Community Energy Challenge. The decrease in the average thermostat temperature reported in Table 3 is instead significant and characterised by the highest effect size ($M_{\text{difference}} = 0.52$, Cohen's $d = 0.58$, intermediate effect size) – which strengthens the idea that changes in thermostat settings played a major role in driving the energy saving effect we found. This is confirmed by research on the building sector: considering current building structures and technologies, the largest energy saving that households can obtain is related to heating (Kemmler and Spillmann, 2021).



Changes in energy consumption for heating are intertwined with the temperature level that is considered to offer sufficient comfort, and with a trade-off with costs (Sovacool et al., 2020). From a sufficiency perspective, lowering indoor temperatures requires adapting one's routines to maintain thermal comfort. The intermediate survey at T2 investigated which routines were changed by users when they set lower temperatures (Figure 18, items based on Matschoss et al., 2021). Respondents declared they increased their use of blankets during both day and night, decreased their use of extra electric heaters, and gave in less frequently to the temptation to turn the temperature up again. However, they declared a decrease in use of warmer clothes –which is the most common way to keep warm in lower indoor temperature settings (Matschoss et al., 2021).

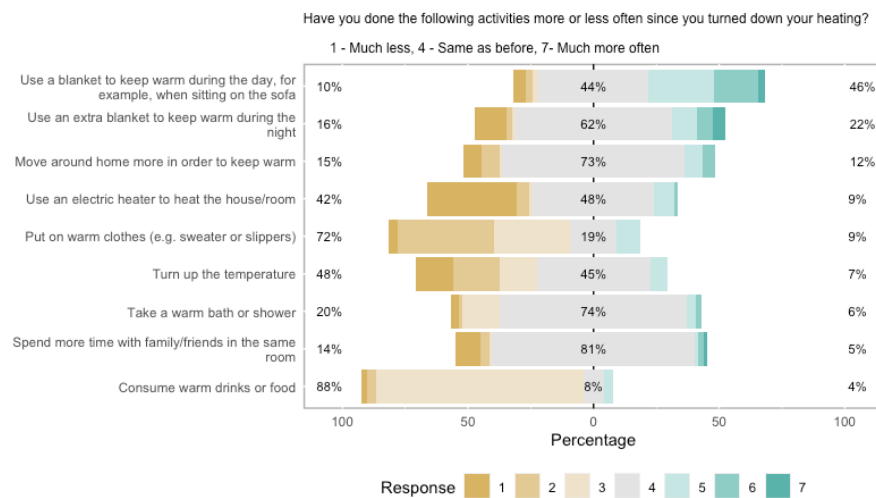


Figure 17 Household practices when heating temperature setting is turned down (intermediate survey at T2, n=132).

The significance and effect size of the average change in temperature setting we found is partially in contrast with findings by Sovacool et al. (2021), who argue that policies aimed at encouraging households to reduce the thermostat setting temperature may have limited effectiveness because individuals want comfortable, warm homes. Heating decarbonisation –they argue– can more easily be obtained by providing individuals with the needed comfort level (via energy efficiency solutions), rather than by relying on them to voluntarily reduce their heating demand (via energy sufficiency solutions).

Such a perspective is actually very well represented by a few open comments by app users that have already implemented energy efficiency interventions in their houses and optimised their energy consumption, such as: “We will continue to not sleep in winter coats, knitted socks, night caps. [...] I will continue to take warm showers, eat warm and relax in the evenings with pleasure. [...] As far as energy saving is concerned, I'm very well informed and trained in terms of energy consumption, so my wife and I have been living very environmentally conscious for a long time. My consciousness is oriented towards the future and not back to the Middle Ages, as some people understand when it comes to saving energy.” Or, more in general about changes in showering practices: “I have already implemented many things. Cold showers were out of the question. I installed a new door in my house with a better insulation value. [...] I found some of the tips from the other participants to be absurd and self-aggrandizing (e.g. take a cold shower every day). Instead, I bought a certified water saving shower head”. A household participating to the SPP Community Energy Challenge also reported that, thanks to a data analysis with the app and a better understanding of their consumption patterns, they opted for switching to “a boiler with less capacity and with air-water heat exchanger”. This therefore suggests that in some cases participation to the SPP Community Energy Challenge resulted in energy efficiency intervention. However, the average results on self-reported household thermostat setting routine changes reported in Table 13 suggest that some degree of sufficient heating behaviour is also possible and was voluntarily adopted in real-life conditions.



4.8.2 Topics of interest by the participants

Analysis of pinboard posts can help us shed light on the self-reported evolution of practices that we observe through the survey responses. In total, N = 257 messages were posted on the pinboard: n = 193 were by members of the SPP community and n = 64 by three app administrators from our research team. Many posts reported issues around data transmission or data reliability and were therefore responded to by the app administrators (our research team). Overall, 78 posts out of 193 (about 40%) met the goals the pinboard was developed for, namely sharing positive or negative experiences, such as someone having better skin because they showered less often or that one could increase the life of garments by not using a tumble dryer. An exemplary selection of such posts is reported in Table 14.

Table 14: A selection of pinboard posts dealing with experiences by members of the SPP community on challenge topics.

Routine	Technical suggestion	Behavioural suggestions
Heating	We have programmable thermostatic valves in three rooms and two bathrooms. The bathrooms are at 23 degrees in the morning and evening and 18 otherwise. In the rooms they are programmed according to absences. This can be done to within 10 minutes and to within 0.5 degrees. Of course, this only makes sense with radiator heating systems and if the temperature can be lowered in the rooms for a sufficiently long time. The valves need about 1 set of AA batteries per year. From a cost point of view, the purchase is therefore questionable, but from a CO ₂ -saving point of view it is certainly worthwhile. We have to save about one thousandth of our gas heating costs to offset the CO ₂ footprint of the batteries.	Study: often froze when sitting in front of the PC in winter when the room temperature was low. Remedy: standing desk combined with small movements!
Showering	@DH: I know that too! I have the PROSECCO shower head from AquaClic. This was advertised by the city of Winterthur in 2016. I can definitely still enjoy myself with 8 l/min. With newer models (e.g. in the kitchen), the flow can be changed at the touch.	When I was a child, baths or showers were taken on Saturday. That went on and was normal back then. On days when I don't have to leave the house and don't do any sweaty activities, I wash myself - as I used to - with a flannel. My skin and hair are grateful and energy is saved.
Cooking and food conservation	---	Bring the pasta water to the boil with the lid closed, add salt, put in the pasta, stir once, put the lid back on. Turn off the cooker. After about 10 minutes, the pasta is "al dente" and nothing spills over. The glass ceramic hob stays spotlessly clean!
Laundry	---	Except for underwear, I like to hang my clothes up to air after wearing them. If they don't smell the next day, they can be worn again, the difference is considerable depending on the fabric and the wear! Of course, a T-shirt sweated through in high summer doesn't have a second chance, but especially in winter certain pullovers or trousers can be reused for several days. Of course we alternate, never wear the same clothes in a row.
Dishwashing	We have connected our dishwasher to the hot water system because we produce the hot water with the heat pump. So the treatment is a bit more efficient than from the machine itself. The washing machine is next. As soon as I have the solar thermal in operation, the energy consumption drops to almost zero.	The biggest difficulty is the uneven use of dishes, in our house lots of glasses, espresso and coffee cups, but few plates/pots, so the bottom half of the dishwasher is often not quite full.
Gardening	Since we only have a very small lawn area, about 25 sqm, it can be kept in good shape very well with a manual reel mower instead of the electric mower, and even with a better cutting pattern! I always shook my head when our ex-neighbour used a petrol mower for his 50 square metres!	I also do the scarifying by hand, which is sweaty, but I put it down as a fitness programme.



Considering only user posts and excluding posts related to technical issues, the most frequent post category deals with heating-related routines (Figure 18). Three factors might have driven the interest in heating-related discussions: heating challenges were the first to be launched at the beginning of the intervention, when interest and engagement was highest; SPP displays heat energy consumption data separately, thus highlighting the impact of heating more explicitly; and members of the SPP community consisted solely of households living in independent houses and are thus directly responsible for the consumption and costs of their heating system. Indeed, the prominence of the heating-related posts in the pinboard might be among the reasons why the changes in self-reported heating-related routines (thermostat temperature setting) were statistically significant and characterized by the largest effect size: pinboard posts may have started to activate a change in shared cultural conventions and social norms about the indoor temperature setting that is expected to be “normal”.

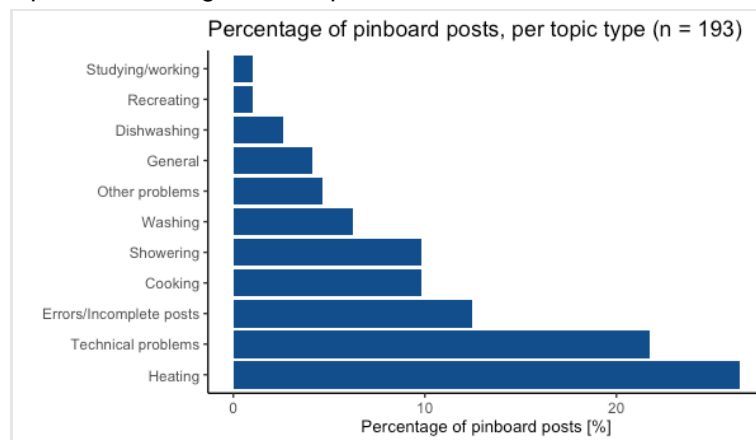


Figure 18: Classification of pinboard posts by members of the SPP community.

Most peer suggestions referred to actions that could immediately be implemented without investment in technical infrastructure or equipment. Some posts reflected on energy consumption in a broader sense. For instance, an almost philosophical discussion arose on the trade-off between comfort and health, and how energy saving raised the question about the difference between comfort and habit. Further, users reflected on grey energy embedded in smart energy tools, or how the impact of a single full bathtub frustratingly outweighed an otherwise highly efficient lifestyle. A few other posts, instead, had highly technical, expert-based perspectives, and mentioned investment in energy efficiency of the building or heating system. Quite surprisingly, all the posts that included a direct question to other participants to the SPP Community Energy Challenge and that received an answer (fully reported in Table 15), dealt with highly technical and expert topics, such as how to set the parameters of the heating plant in order to ensure low consumptions. This suggests that the participants to the SPP Community Energy Challenge were characterized by higher experience than average on energy-related technical topics – a consideration on which we will come back in Section 4.8.4.

Table 15: The full list of pinboard posts reporting questions by members of the SPP community, together with the answers provided by other peer users. Nicknames of app users, modified for this document, are reported in brackets.

Question	Answer
Is 55 degrees, once a week, enough to keep the domestic water free of legionella? Normal temperature = 50 degrees. (H6)	I think I would have heard if anyone in our estate had had a problem with legionella. (35 detached houses, >20 years old). Our heat pumps don't go above 50 degrees. But I guess it's not really safe that way. (CHB)
	My new heater heats the hot water to 60 degrees every Sunday. that's what the professionals recommended. (H1)



<p>@ AS: Good idea to use a booster in low tariff. I'll be happy to try that out. How many degrees higher than the daytime temperature did you set the booster temperature? If it was too high, I think the heater would often kick in, which wouldn't be desirable. (MS)</p>	<p>@MS: Interesting, I have evening sun.... I set the increase before the end of the nightly low tariff to 3°. This increase does not translate into a correspondingly higher room temperature because of system inertia, and since it tends to be warmer (AAS)</p>
<p>@BA: May I ask what the smart home retrofit cost and who installed it for you? I would be very interested. Merci in advance! (KH)</p>	<p>@KH: The installation also included the automation of a blind and other small extras that can be integrated into the same SmartHome system (e.g. smoke detectors) and cost +/- CHF 5000. To be honest, we have to say that the installation of the room temperature (BA)</p>
<p>@BH: What kind of models <i>[of programmable thermostatic valves, that user BH described in a post before, editor's note]</i> do you have? and is there anything special that needs to be taken care of during installation or can a loaner do it themselves? (HN)</p>	<p>We have Danfoss Eco. They can be programmed with an app on your mobile phone via Bluetooth. It's quite intuitive. Installing the valves is pretty good (I'm not a handyman). It comes with adapters for a lot of different connections. (BH)</p>
<p>My wooden house is heated with the HP with controlled ventilation. How much can the room temperature be reduced during an absence (approx. 1 month)? What do the heating experts say about this? Has anyone experience with this? (EA)</p>	<p><i>[Not answered by the community]</i></p>
<p>Here, too, a question for the experts: From how many days of absence is it worthwhile to switch off the purely electric boiler by fuse? Or just reduce temperature significantly? (AH)</p>	<p>From an energy-saving point of view, this is worthwhile from the first day onwards, depending on many factors, of course, but basically you save from the first day onwards, provided that the boiler (if left on) continues to heat every day even when you are away, which should be the case most of the time. (DH)</p> <p>Addendum: it is better to reduce the temperature instead of switching it off completely because of frost protection: my WP heating is programmed in such a way that a minimum temperature of 10 degrees is still ensured when the boiler is "switched off". (DH)</p>
<p>Consumption No heating We are absent, boiler off via fuse and RCCB <i>[Residual Current Circuit Breaker, a safety measure for electrical circuits, editor's note]</i> I am now completely amazed to observe that there is regularly a consumption peak of 2 to 4 kWh in the morning at 6 am and at 11 pm. The rest of the day is about 0.75kw/h. These are the fridges, for example. But I can't explain these two peaks for the life of me. Do you have any clues? (AH)</p>	<p>@AH what are your large consumers that could draw over 2kW? Boiler? HP <i>[heat pump, editor's note]</i>? Air conditioner? Electric car? Cooker/oven? Is the heat pump/boiler really separate, or just the control system? It could be due to the frost protection programme, if you are absent. (DH)</p>

4.8.3 Participants' feedback on the SPP Community Energy Challenge features

The quantitative results about the average treatment effect on the treated show that participation to the SPP Community Energy Challenge produces statistically significant results in terms of energy saving at home. We now consider automatically collected in-app data and elements from the intermediate survey at T2 (at the end of the "high intensity period"), to discuss which features of the SPP Community Energy Challenge have possibly played a major role in producing the effect we found.

A summary of key in-app actions performed by participants to the SPP Community Energy Challenge is reported in Figure 19, which represents the evolution over time of the total number of actions by all app

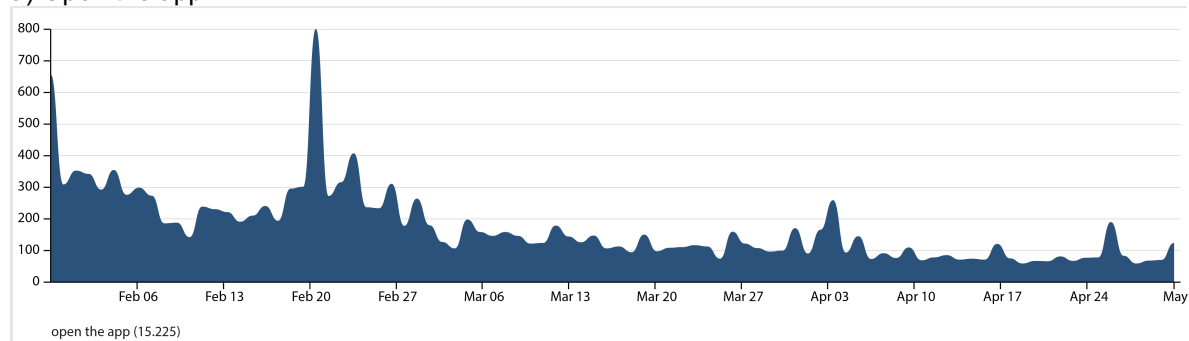


users during the “high intensity” period of the SPP Community Energy Challenge, when all app features were available. Though the number of app openings remained quite high over time, with about 80 openings per day even in the last month (Figure 19.a), the plots suggest that the level of engagement with the SPP app’s key features was limited. This is confirmed by the descriptive statistics reported in Table 16, which shows the number of users that performed each action and the average number of actions per user. Individual goal setting was the most frequent action performed. Data shows that goals were set on average 9.5 times (Table 16), namely on average they were changed 7.5 times after the initial setting of the two heating and non-heating goal targets, which was compulsory at the start of app use, and that changes in the heating target were most frequent (Figure 19.b). In contrast, the number of users engaged in challenges or in posting pinboard messages was smaller (Figure 19.c and Figure 19.d): respectively, 24% and 21% of the users engaged in a challenge or posted a pinboard message (Table 16). Furthermore, only 6% of users formally completed at least one challenge by sharing their experience in the pinboard. Among those that performed such in-app actions at least once, the frequency of in-app actions was however satisfactory: Table 16 shows that users that posted at least one pinboard message, on average posted 3.8 messages; users that at least started one challenge, on average started 2.4 challenges; and users that at least completed one challenge, on average completed 2.5 challenges.

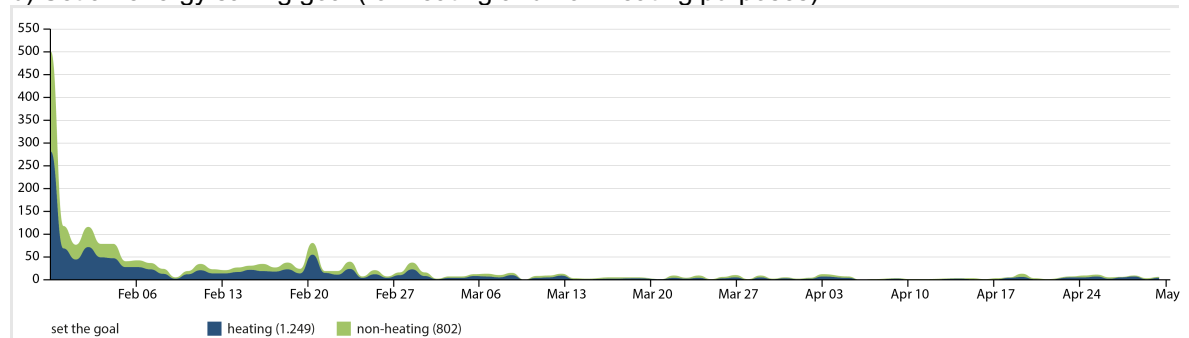
Table 16: Descriptive statistics on the number of in-app actions per user.

		Open the app	Set a goal	Post a pinboard message	Start a challenge	Complete a challenge
Users that performed the action at least once	n	203	203	42	48	12
	%	100.0	100.0	20.7	23.7	5.9
Number of actions per user (among those that performed the action)	Mean	70.8	9.5	3.8	2.4	2.5
	SD	106.0	9.1	3.6	1.7	2.2

a) Open the app

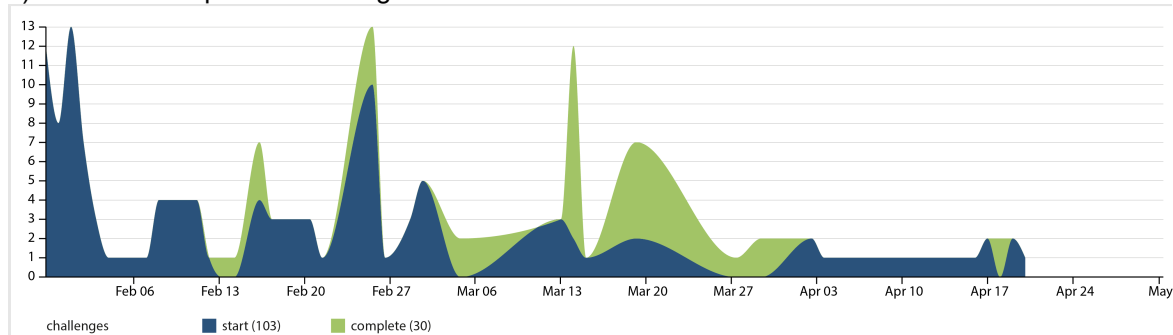


b) Set an energy saving goal (for heating and non-heating purposes)





c) Start and complete a challenge



d) Post a pinboard message

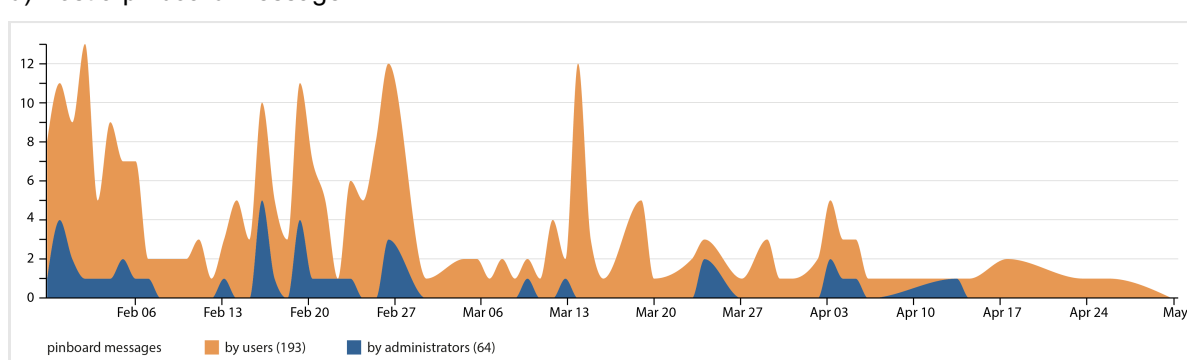


Figure 19 Evolution over time of key in-app actions by SPP users. Note differing counts on y-axis scales.

Elements about the evaluation of in-app features from the perspective of app users were also collected via the user's direct feedback in the intermediate survey at T2, which is graphically represented in Figure 20. To collect as much information as possible from the users' perspective, for this analysis we consider the full sample of the T2 survey respondents (however by filtering out those that never logged in the SPP app). The resulting sample is made of $n = 132$ respondents.

In line with the above figures, survey respondents indicated appreciation for individual-level features (energy consumption feedback, goal setting, and tips). The pinboard feature, on the other hand, is rated as the least useful to motivate energy savings, maintain user engagement, or support learning something new about saving energy. Specifically regarding pinboard posts and active social interaction processes, 75% of respondents declared they did not like to share their experience via pinboard posts and 65% of them declared pinboard-mediated interaction with other users did not motivate them to save energy. Thus, the majority of users did not engage in active social interaction. Responses also indicate that passive social interaction processes have only occurred in a limited number of cases: 43% of respondents indicated they did not like to read about other participants' experiences. Overall, combined with 19% stating a neutral position, these figures show participants were marginally engaged with the pinboard through reading the posts of their peers.

The respondents' perceptions about user engagement and social interaction processes emerge also from the open comments on the whole SPP experience: though a few respondents explicitly indicated their appreciation for being part of a community (*"It was fun! It was nice to see that other people are also dealing intensively with the topic"*), others very explicitly stated their lack of interest for reading about suggestions by their peers via the pinboard: *"Above all, I want as much data as possible about my house and my use. I don't have time for chats in an app at the moment"*.

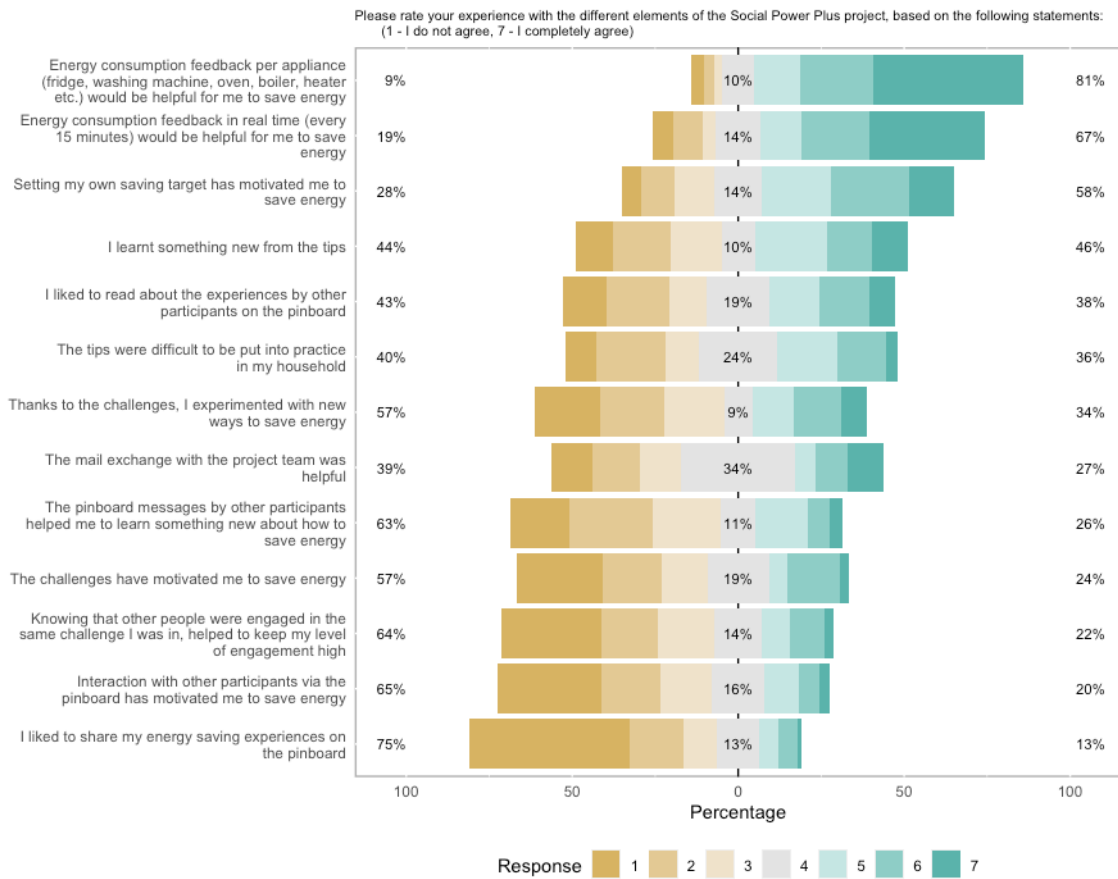


Figure 20 Feedback on app features according to the intermediate survey at T2 (n = 132).

A slight majority of respondents also indicated that energy saving challenges neither inspired them to experiment with new ways to save energy nor motivated them to save energy. Overall, respondents found individual level features to be the most engaging and useful features of the app. Overall, 81% of respondents would like to receive the breakdown of consumption at the appliance level (*“The only useful information the app provides is the distribution of energy between heating and the rest. However, this is too high-level and you have no idea how the use of individual devices affects the energy consumption”*), and 67% of them would like to monitor consumption data in real time (*“An hourly consumption reading only makes sense to me if it can be tracked simultaneously. The day after, I usually don’t remember what I was doing at what hour the day before”*). Some respondents also indicated they would be eager to freely browse through historical series of recorded consumption data or even download consumption data, for further analyses by themselves. *“I don’t want a “fancy” app, but simply to be able to download the current performance as a csv file with the best possible temporal resolution”* is the straightforward open feedback by a respondent.

Similar comments also emerged during the live meetings that were performed once a month during the high intensity period of the SPP Community Energy Challenge. During such meetings, the research team offered explanations about the way the different app features worked and the participants were invited to freely comment on them, presenting their opinion and perspective, including suggestions on how to modify them in order to make them more appealing and better fitting to their needs, interests, and expectations. The full list of such comments is reported in Table 17. Also from this list, the interest by the participants in highly detailed, quantitative, and in many cases technical information clearly emerges. Participants to the live meetings in fact ask for high resolution data, for historical series



showing evolution of the consumption in the long-term, for download and export possibilities, that allow them to perform further analyses via external spreadsheets, and for real-time information that can allow them to clearly identify the source of consumption. When real-time data is not possible, they ask for activity diary features, that allow them to identify high energy consumption activities by comparing them with the energy consumption charts. Finally, push notifications notifying users of unusually high consumption levels, as well as customized tips, recommendations and warnings based on their consumption patterns would be highly appreciated.

Table 17: The list of comments and suggestions by the participants, collected during monthly live meetings.

Feature	Comments and suggestions by the SPP Community Energy Challenge participants
My Home	<ul style="list-style-type: none"> • In the comparison with similar households, also consider the characteristics and age of the house building • Benchmarks should be established for certain types of houses and households, besides the household composition as currently done
My consumption	<ul style="list-style-type: none"> • An average line based on the previous week/day would be interesting • Weekly consumption charts are too small on smartphone display (enable side scrolling through different days) • Add more graphical representations and other analytical functions for «My consumption» • Allow to compare certain weeks against each other • Allow to always access highly detailed data from the past (e.g. from month to week to day to hour) e.g. through a zoom function in the long-term graph • Show higher resolution data in terms of timespan • Combine "heating" and "non-heating" energy consumption visualisation into one single chart, to make the amount of different consumption types clearly visible. In particular, for the chart representing the "weekly consumption since the beginning of the year" • Allow richer consumption visualisation over different timespans (not only yesterday, last week) • Show baseline values also in total consumption • Show absolute values in kWh for a day/week • Show statistics of energy consumption data • Add dates in the tables • Show savings in CHF • Include «degree-day» indicators for a long-term comparison (e.g. year to year) • Access to smart meters real-time consumption data (rather than access to them the day after) would allow to more correctly break consumption down between the heating/non-heating categories
Tips	<ul style="list-style-type: none"> • Offer tips for efficiency measures and new technologies • Offer personalized tips based on the household's consumption patterns
Pinboard	<ul style="list-style-type: none"> • The structure is not so clear, maybe offer just one pinboard per topic (heating, showering, etc.), instead of splitting it for single challenges • Share best practice examples from other households, including kWh and monetary savings • Peer-to-peer consulting is very promising, especially if one can show, what the results in energy savings were in his/her case
Challenges	<ul style="list-style-type: none"> • There are great saving potentials for geothermal probe heat pump (for heating and hot water) with better settings: include this as a new, separate challenge
New features	<ul style="list-style-type: none"> • Develop appliance use diaries to compare consumption with activities • Send push notifications if consumption increases significantly • Offer financial incentives to increase participation/motivation/interest for the general population, e.g. the more you save, the less you pay for your energy
General suggestions	<ul style="list-style-type: none"> • Allow app use on iPad • Enable download of whole dataset as Excel files

4.8.4 Characteristics of the participants to the SPP Community Energy Challenge



Previous literature on energy-saving smartphone applications (apps) has shown that their effects were not maintained in the long-term (Wemyss et al., 2019; Cellina et al., 2024). When co-designing the features of the SPP Community Energy Challenge and app, we assumed two key reasons impair such apps' capability to ensure the long-term change needed by the energy and climate transitions. First, as conceptualised by Kaufman et al. (2021), most of such apps solely address the individual level, by adopting a “strategic behaviour” perspective grounded in individual behaviour theories. Thus, they fail to engage with the socio-technical context in which energy consumption is embedded, namely organisational, geographical and “practice-as-context” aspects (Heiskanen et al., 2019) or, more broadly, with infrastructures, other practices, and social actors (Raven et al., 2021), that can enable or impede changes in households' energy demand. Second, as argued by Strengers (2014), these apps tend to focus too much on engineering and quantitative consumption feedback data, being mostly designed for “resource men” archetypes: individuals that are interested in their own energy data, understand it, and are open to change their energy consumption patterns by responding to feedback signals. As remarked by Strengers (2014), most individuals do not understand energy consumption data, do not know the meaning of “kWh”, and cannot guess the implications of “tons of saved or emitted CO₂”. Furthermore, household energy consumption routines are constrained by work, school, and leisure schedules from household members and, more broadly, by the socially shared conventions of “normality” that affect them – shared conventions that transition processes should explicitly address, if long-term change is sought.

Acknowledging the above as critical limitations, in SPP we coupled the “strategic behaviour” perspective with the “everyday perspective”, as it was conceptualized by Kaufman et al. (2021). Rather than focusing on energy consumption *per se*, in fact, the SPP app and Community Energy Challenge explicitly deal with real-world challenges: they tackle daily household routines (e.g., heating, showering or cooking) with the aim of re-crafting them towards energy sufficiency (Moser et al., 2015; Seidl et al., 2017; Jungell-Michelsson and Heikkurinen, 2022) and reducing consumptions in absolute terms. SPP also accounts for the social context in which energy consumption is embedded, by creating opportunities for peer-to-peer household interaction via the pinboard in-app forum. Challenges and pinboard features also make the experimentation coherent with living lab processes grounded in Social Practice approaches, which try to reduce energy demand by acting on both the individual routines and on the social norms and conventions that support them (Heiskanen et al., 2018; Sahakian et al., 2021; Matschoss et al., 2021).

The above insights from the users' evaluation of the app features, together with the analysis of the open-ended feedback about their overall SPP experience, suggest that many SPP households actually differ from the target general population we expected to engage, and for which the SPP's app features were designed. Some households clearly indicated the changes they have introduced in their energy consumption routines and especially in their use of appliances. However, it seems that many households are already highly experienced with saving, having in some cases already invested in energy efficiency retrofits and being familiar with the amount of energy consumed by their house, thanks to self-monitoring activities they had already started before the SPP Community Energy Challenge. This tendency is suggested by the high interest in real-time, highly broken-down feedback about energy consumption reported in the previous section and is confirmed by some survey open comments, such as *“If someone has no idea how to save, the app is ok. But if you have optimized your household practically everywhere, the app is rather frustrating. Either I don't wash my clothes anymore, only eat cold, have a barbecue every day...”* or *“Since I've always been very careful about saving energy, my energy saving potential was very low. Only reducing the heating from 19 to 18 degrees made a difference”*.

Many SPP app users seem thus to coincide with the highly energy efficient “resource man” conceptualized by Strengers (2014), despite the app design also aiming to attract users who have less energy saving experience. On the one hand, this might be a consequence of the eligibility requirements imposed by current metering infrastructure: to join the SPP experiment, households were required to live in independent houses and be equipped with smart meters accounting for their heating (gas or electric heat pump) and non-heating electricity consumption. The presence of heat pumps, in particular in older buildings, is not common in the current building stock in Switzerland (SFSO, 2017). Thus, it is



the result of a specific choice by the household, which in turn might also have affected the intention to implement other energy saving measures already before joining SPP. This would also explain the wish from users for detailed energy consumption feedback, and customised recommendations to further optimise their energy consumption.

On the other hand, presence of many “resource men” in the SPP households might also be the direct outcome of the self-selection process behind app use. The data we collected suggest that apps like SPP might tend not to raise the interest of their target user group but rather to reach already technically experienced individuals. The very fact that the SPP experiment was built around use of a technological device (an app) providing energy data feedback through direct connection with another technological device (a smart meter), might have restricted the potential audience of interested users to those that were already personally engaged with energy consumption data monitoring and technical optimisation processes. Namely, we cannot rule out that app-based processes may tend to attract “resource men” rather than users that have no prior engagement with their energy consumption data and with its optimisation through technology-based, energy efficiency approaches. If so, radically different communication approaches would be needed to raise the interest of more diverse user groups. Despite this, however, use of the app and participation to the SPP Community Energy Challenge on average produced nearly 5% electricity savings over an entire year. This suggests that managing to engage non-resource men, that are characterized by less prior interest in energy saving activities, and therefore have more room for changing their routines and use of appliances, would likely result in higher energy saving impacts.

4.8.5 Preaching to the converted?

If the “resource men” that tended to be attracted by the SPP Community Energy Challenge were already experienced with energy efficiency measures and had already implemented them in their households, a risk of “preaching to the converted” clearly emerges. Particularly, if those that could benefit the most from SPP participation are not sufficiently interested to do so, the overall value of these apps would be limited. This indicates a clear obstacle to possible large-scale deployment to the broad population, once the ongoing smart meter roll-out process will be completed and smart meters will be available to the whole population. Under such conditions, apps and challenges like SPP could potentially be available to any household; however, most of them would not be interested and would ignore them. Even thinking of a possible future in which apps and challenges like SPP are universally offered under “opt-out” frameworks, the outcomes of the SPP case suggest that they would simply be ignored and not accessed by a large share of the population —probably those who would benefit the most from them.

Such a scenario is for instance coherent with findings by Nilsson et al. (2014) in a randomised intervention providing electricity consumption feedback through In-Home-Displays to Swedish households. They found that a lack of interest in energy saving, together with insufficient understanding of the information provided by the display itself, were barriers to energy savings reported by the households (that were randomly allocated to the feedback treatment instead of self-selecting themselves into it). Therefore, even though “opt-out” strategies can potentially increase the audience for use of SPP-like apps, we expect that their practical impact would not be higher than the one we found for SPP. Future research is therefore needed to scrutinise whether differences actually exist between the target population and the participants to the SPP Community Energy Challenge by identifying how many similar participants exist and what household characteristics they have. For this purpose, one could rely on the measurements on pro-environmental values and beliefs used by the Swiss household energy demand survey (SHEDS), that was administered in Switzerland from 2016 to 2020 to a probabilistic panel sample of about 5'000 households, which is representative of the Swiss population (Weber et al., 2017).

In any case, the problem of “preaching to the converted” is not a new one. For instance, it had already emerged two decades ago in the EcoTeam Programme analysed by Staats et al. (2004): due to the effort required by programme participation, participants were “already ahead of the population with respect to their pro-environmental behaviour” (page 363). More recently, Bird and Legault (2018)



explicitly discuss the effect of energy consumption feedback, and more broadly energy saving behavioural interventions, on “high achiever” households, namely those that have already achieved low consumption levels. While the authors argue that prompts or feedback messages will not provide additional benefits to them, they however do not provide recommendations on how “average-” or “low-achievers” should be handled. Dealing with In-Home-Displays (IHDs) providing electricity consumption feedback, Buchanan et al. (2015) remark that IHDs may only appeal to a niche subset of the population, which may limit the overall aggregate effects of feedback on energy consumption. They explicitly remark that, as intervention participants are made up of volunteers who actively decided to join, the findings in current research may have over-estimated the overall IHDs benefits and cannot be generalised to the whole population. They conclude that clear opportunities exist to target “energy non-engaged” or “energy stagnant” households, as they were profiled by Murtagh et al. (2014). “In targeting such consumers rather than ‘preaching to the converted’, government policy may have more scope for achieving much wider scale results” (Buchanan et al., 2015, p. 92). However, also in their case the problem of how to reach non-engaged households remains open.

A similar result is found by Puntiroli and Bezençon (2020), who analysed the long-term effect (ranging from one to seventeen years, depending on the household) of owning an IHD device that provides electricity consumption feedback. They considered 276 Swiss households (138 owning the device and 138 acting as a control group, identified via matching techniques) and found that only households with high bio-spheric values (environmentally concerned households) who had owned the IHD device for at least three years, reduced their energy consumption. Households that were less concerned with environmental topics, instead, did not decrease their consumption, even though they owned the IHD device. Even if the authors do not explicitly discuss it, the problem also remains for other approaches such as the one attempted by Sahakian et al. (2021) who aimed at challenging social practices through living labs. The authors indicate that they engaged 306 households across eight European countries (on average, nearly 40 households per country), that were selected through open calls performed via advertising on local (social) media and on-street campaigns. Households were requested to take detailed notes of their activities through diaries, to answer three surveys (before, during and after living lab activities) and to participate in in-person meetings and group interviews. Under such a demanding plan of activities, we assume that people’s intrinsic motivation to comply with the project—and thus self-selection bias—played a key role and highly influenced the results of living lab activities. Though the authors explicitly claimed that they were not looking for statistical representativeness (and rather, they aimed at getting a heterogeneous sample and at including traditionally less represented groups, such as unemployed or single-parent households), we suppose that their results would have been largely different if they had interacted with less intrinsically engaged households. And in any case the problem remains about how to broaden the change in social practices experienced in the living lab to the wider, less intrinsically motivated and non-engaged population. The challenge of how to go beyond the converted—no matter what the behavioural intervention is—remains a key issue that future research has to experiment and deal with.

4.8.6 Limitations of this work

We acknowledge the results we obtained are not exempt from limitations. First of all, we adopted a quasi-experimental research design, which is weaker than truly experimental approaches in regards to the capability to produce robust causal impact estimates. Moreover, despite the care we devoted to matching comparable control groups to the treatment group households, to perform such a matching we could only rely on the households’ yearly energy consumption in the two years before the SPP Community Energy Challenge. Therefore, the matched groups might actually largely differ from the treatment groups in many unobservable variables that we could not access. This may affect the internal validity (Vine et al., 2014; Frederiks et al., 2016) of our quasi-experimental process, and thus the related estimates of average treatment effect. The lack of availability of information on house(hold) characteristics also precludes the possibility to control for household socio-demographic variables or for



house building and technical equipment variables in panel regression models, thus further affecting the quality of the results we obtained.

Another relevant limitation lies in the problem of self-selection of intervention participants, who were volunteers answering a public call to join project activities. This may affect both the internal and the external validity of our quasi-experimental results (Vine et al., 2014; Frederiks et al., 2016). If households that decided to participate in project activities (the treatment group) were systematically different from those who decided not to participate (e.g. due to higher education, income, age, or pro-environmental attitudes), then the quasi-experimental results and the estimates of the treatment effect would be affected by a problem of internal validity. On the one hand, in fact, the treatment effect on the broader population would have been over-estimated. Voluntary participants would in fact have had higher intrinsic motivation to interact with the app's features. On the other hand, the treatment effect might have been under-estimated, as voluntary participants might have had less available room for change. Due to their intrinsic motivation, before joining the intervention they might also have already implemented at least part of those energy sufficient behaviours that the apps aimed at supporting. And as those behaviours would have been unobserved, they would not have been properly controlled for in the estimate of the effects of the intervention (Sergici and Faruqui, 2011). The same self-selection issue might also have raised problems of external validity: if the households involved in the SPP Community Energy Challenge were not representative of the interventions' target populations, as we suspect, reliable conclusions could not be drawn about the overall effectiveness of offering the same apps to such populations. Namely, the lack of external validity would affect opportunities for scaling up app use (with the same impacts) beyond the intervention participants.

Future research might first of all verify the existence and amount of self-selection biases, by administering new questionnaires to the participating households (treatment groups) and comparing key variables (environmental values, beliefs and attitudes, as well as income and education) with the responses by the Swiss probabilistic sample monitored by the Swiss household energy demand survey SHEDS by Weber et al. (2017). If self-selection biases actually emerge by comparing the three case samples with the SHEDS one, future research should identify novel ways to tackle and reduce them.

One possibility to overcome limitations stemming from self-selection of participants would be to still recruit the sample of possible intervention participants under a voluntary, opt-in scheme, and then to randomly allocate treatment and control within such a sample. For the SPP case this was not possible due to the small sample sizes of households that applied for the SPP Community Energy Challenge. Randomisation within that sample would in fact have required to halve (or in any case largely reduce) the size of the treatment group, that was already very small. This would have very likely precluded the possibility to find any statistically significant results, therefore from the very beginning we discarded this intervention design option. Nevertheless, future works should strive as much as possible to adopt such a randomised design, which would also solve the problems about poor matching that we have discussed above. The small size of the samples is, indeed, a critical aspect in itself, which weakens the results we have presented. In particular, the lack of statistical significance we found for the heterogeneity analyses we performed in order to assess treatment effects for different sub-groups of households might be due to the small sample sizes and a related insufficient statistical power.

Finally, for the analysis on routines we relied on questionnaire data. The responses we obtained might have been affected by social desirability bias, as the goal of the SPP Community Energy Challenge (reducing energy consumption) had been clearly communicated in project-related materials and by the SPP app itself. Similarly, we used questionnaire data to collect feedback on the SPP app and Community Energy Challenge features. While this approach allowed us to collect evaluations on the different features directly from the perspective of the app users themselves, this does not allow us to strictly identify which features were more effective in producing the energy savings. The research design we adopted does not allow to disentangle effects by a single intervention technique or feature: we could only obtain aggregated estimates of the overall effectiveness of the SPP Community Energy Challenge and could not identify the amount of saved energy due to the presence of each feature. As the insights from the surveys suggest, in fact, some of them were more effective —and should therefore be



recommended to policymakers— while other ones are more likely to have been ineffective and could be discarded. However, the current research design does not allow us to strictly perform this kind of statements. Use of experimental designs allowing to estimate the causal effect of each single intervention component would have been more advisable. However, such designs would again have required availability of a large sample of participants, to be split in sufficiently powered sub-samples for random allocation to different intervention categories. The limited size of registered participants would not have allowed to do so. Quantitative, strict analyses on the effectiveness of the different SPP app and Community Energy Challenge features are therefore left to future applied research.

4.9 Recommendations for future app-based interventions

The strict analyses we performed to estimate the average treatment effect of the SPP Community Energy Challenge suggest that tangible energy saving results can be obtained –and that they can probably be even higher, if households with no prior engagement with energy-related topics could be involved. The problem of how to engage them, and thus of how to go “beyond the converted”, remains however open. Nevertheless, households that voluntarily opted for joining the SPP Community Energy Challenge were highly interested in quantitative information about their energy consumption data. The practical experience we gained through SPP thus suggests to try to meet this expectation, because this is what potential app users and challenge participants are looking for. Particularly, the SPP experience suggests that specific attention needs to be devoted to two other key elements, namely the capability to provide highly disaggregated consumption feedback by interacting with the smart meter network, and the identification of viable business models that can bear the costs for the development and management of app-based interventions.

The SPP Community Energy Challenge and related app were in fact designed with the aim of raising the attention of non-engaged households, that are characterised by limited prior interaction with both the concepts of energy efficiency and sufficiency. This only happened for some of the households that voluntarily decided to join the SPP Community Energy Challenge; most of the users, instead, called for much more detailed energy consumption feedback, than the feedback that the SPP could offer, asking for real-time and highly broken-down (appliance-level) consumption information. Managing to meet their expectations is therefore crucial in order to enable the success of app-based or, more broadly, digital technology enabled processes of change. Introducing such novel features, which go beyond current smart meter technology, would add additional efforts to the utility companies who are interested in launching app-based challenges or alike. How to financially cover the needed IT infrastructure, development and management of the app and challenge, in order to ensure they are economically viable and can self-sustain over time, therefore requires in-depth analysis. To investigate these aspects, we performed three final interviews with representatives of the utilities (SWW, TBW, and EKS), followed by a workshop with the research team members (CLEMAP, ZHAW, and SUPSI researchers, including the app’s software developer). The most relevant insights we collected, summarised in the next sections, allow us to draw recommendations for future activities that aim at supporting the energy transition by leveraging smart meters, energy consumption feedback, and app-based interventions.

4.9.1 Smart meters and disaggregated consumption feedback

Exploiting consumption feedback from smart meter data is a challenge and opportunity experienced from many sides: the utilities follow regulation on large-scale smart meter roll-outs, the research team test potential added value of the technology, and the data management company (CLEMAP) looks for viable business models.

Utilities have no possibilities yet to use smart meters for demand side management processes and exploit the meters’ bi-directional communication features; however, they still recognize the potential of smart meters, both for the novel grid management services they enable (e.g. remote updates on firmware and troubleshooting) and for the data to improve internal operations and feedback to their customers. On the one hand, smart meter data can be used for grid forecasting and optimisation, as



well as for stress tests and simulations. On the other hand, smart meter data can be used to provide the utilities' customers with highly customised energy consulting services and build awareness around energy consumption topics among their customers.

Outcomes of the interviews and workshop discussion suggest that currently, possible benefits of smart meter data are not fully exploited, due to a number of factors. First of all, extremely strict regulations in terms of data ownership and privacy protection make accessing and using the data collected by the smart meters quite difficult. This is particularly critical for Swiss-based utilities, which are quite small and have limited resources to tackle legislation-related challenges. Furthermore, smart meter data-sets are usually stored in complex data structures, that are difficult to manage for a single utility without the support of a technology/data management company. Using smart meter data thus necessarily implies that the utility company bears additional management costs. Also, the way smart meter data-sets are arranged and organised may sometimes be in contrast with the data storing and management systems that the utilities are equipped with – and that usually consist in platforms that were neither designed nor optimised to store smart meter data. For these reasons, therefore, the utilities acknowledge the opportunity to leverage already existing webportals, platforms, or apps providing smart meter data feedback – rather than internally developing their own webportal, platform, or app. Availability of white-label app/platform solutions, that each single utility can customise to meet their corporate image standards, would also allow smaller utilities to offer novel high-quality services to their customers. Such a possibility will be particularly relevant in the near future, when customer webportals or platforms, which provide consumption data to customers, will become mandatory for all utilities.

Furthermore, thanks to future technological development in the field of gateways, there are expectations that in the future smart meter data will become more easily accessible to any customers, and that non-intrusive load monitoring algorithms will be able to produce reliable and accurate disaggregated estimates of the consumption data of the different appliances. This addresses one of the main requests by participants to the SPP Community Energy Challenge, namely the provision of high granularity feedback data, broken down per appliance. The advantages of being able to provide such a disaggregation are very clear, and related with possibilities to optimise the households' consumption patterns. However, in the case of SPP, when an attempt to break consumption data down into heating and non-heating consumption was performed, the challenges affecting the implementation of this type of feedback clearly emerged in all their complexity. On the one hand, consumption data with higher frequency than 15 minutes would be needed, but this type of data is hardly accessible due to privacy protection regulations. Furthermore, considerable effort and investment in the development of high-accuracy disaggregation algorithms is still needed. Currently however there is a lack of direct economic added value as a market demand and business model have not emerged. Customers would like to get the benefits of disaggregated consumption feedback, but they are unwilling to pay to cover development cost. Or, better: each of them might be willing to pay small amounts of money, which would require many users to ensure the algorithm development costs are covered.

Utilities, or other grid system operators, instead, currently see very limited marketing opportunities in disaggregation algorithms and therefore are not willing to invest in them. Unless specific regulations force energy markets to invest in the development of disaggregation technologies, or totally new business models are identified, it is likely that the large-scale use of high granularity, smart meter data to provide detailed appliance-level consumption feedback will not take place in the near future.

4.9.2 Possible business models for future self-sustaining app-based processes

Potential business models for SPP, or similar processes, have to consider the high costs of the development and maintenance of the technological infrastructure for activities such as the SPP Community Energy Challenge. To support participants, a user helpdesk is particularly necessary, especially during the “high intensity” period. Finally, promotion and communication costs for user recruitment have to be considered, including potential incentives such as vouchers for discounts on the energy bill or in the utility online shop, that aim at recruiting users of the app, webportal or platform,



and/or participation in community challenge processes. It is therefore important to understand if, for processes like the SPP Community Energy Challenge, self-sustaining business models can exist.

From the perspective of the involved utilities, the SPP Community Energy challenge was too complex, concerning the data management, recruitment, and participant helpdesk aspects. Furthermore, in the current non-liberalised market for household customers, utilities have no incentive for improving customer relations and providing them with additional services and thus it is not likely that, under the given framework conditions, utilities will invest in funding SPP-like activities. On the other hand, however, it is foreseen that webportals for utilities' customers will become compulsory. Also, utilities might start to prioritise sustainability issues and to start to value engagement in sustainable activities – at least for image reasons towards their customers. They might also start to value webportals and/or apps as a way to activate direct communication channels with their customers. Currently in fact they only have access to a subset of their customers' email; through a webportal or app, instead, they could easily send push notifications or other communications to all their customers that logged in. Finally, if scenarios of energy shortages should get more likely, for instance due to climate change or geo-political crises, utilities might start to value energy saving apps and platforms as less expensive alternative strategies for managing black-outs.

Furthermore, customers may be willing to pay small monthly fees if they find value (e.g., cost savings, environmental benefits, etc.) in participating in a process such as the SPP Community Energy Challenge. Their willingness to pay for app-based services might even increase in the future if, coherently with the emerging trends due to the Russian war in Ukraine, energy prices will continue to increase. Additionally, the SPP process has clearly shown that other customers may be willing to pay for high quality/granularity consumption data provided on a good platform/app – and of course this requires that they are offered highly reliable and accurate data.

In order to ensure reliable high granularity data, changes in the legislation would be beneficial. The definition of standard interfaces and data management processes would facilitate smart meter data acquisition and analysis for easier processing for such a platform/app. Furthermore, other forms of customer privacy protection should be considered, in order to ensure that smart meter data can be used in useful and meaningful ways.

A possibly viable and self-sustained business model might therefore combine financial support by a utility company, coupled with direct support by the customers via small monthly fees. Utilities could exploit the same “white label” product, made of modular components, that each utility company could customise to meet their needs. Such a product might consist of elements like those offered by the SPP app and Community Energy Challenge (i.e., feedback, tips, and gamification elements), with other services offered by the utility, such as energy efficiency consulting or “smart home systems”, that include heating system monitoring and/or photovoltaics production management. According to the interviews, utility companies would particularly appreciate integrated products, such as a competence centre for energy consulting, that includes an app or web-based portal that offers free services for all customers. This can include additional services (e.g. the disaggregation of energy consumption data) for a fee. Both the consulting services and the app or webplatform would be managed by external providers and not directly by the utilities, as white label products. Doing so, the utilities would have a limited amount of effort, which would be compatible with their limited resources. Furthermore, as such a competence centre and the related app/webplatform would be offered to all the customers, it might provide an ideal framework to support the creation of broad multi-stakeholder processes aimed at going beyond the “converted” and the “resource men” and at collectively re-examining what is considered normal in households' routines and energy consumption practices and thus focus on energy sufficiency principles. Table 18 summarises the recommendations for self-sustaining business models for future app or webplatform based processes, by considering the key stakeholders represented by the utilities, the app or platform developers, and the policy-makers broadly dealing with the energy system regulation.

Table 18: Recommendations for future self-sustaining app-based processes.

Stakeholders	Recommendations
--------------	-----------------



Utilities	<ul style="list-style-type: none"> • Consider co-development of a customer webportal or even broader energy consulting centre with other utilities • Explore willingness to pay from the customer-side • Consider offering multi-stakeholder processes aimed at engaging the broader population in energy sufficiency processes • Consider offering energy efficiency measures, particularly coupled with smart home management systems and renewable distributed power plants • Deploy a future proof smart meter infrastructure which allows demand side management and fast data transfer
App, web-platform, web-portal developers	<ul style="list-style-type: none"> • Develop one white label platform/app for consumption data feedback to provide a high-quality solution to the large number of (small/medium) utilities in Switzerland • Include the following app features: <ul style="list-style-type: none"> • drill down function for consumption data • insightful long term comparisons • personalised tips and recommendations based on individual households' energy consumption data • financial savings from changed behaviour • download function (Excel, other common formats)
Policy makers	<ul style="list-style-type: none"> • Consider other forms of customer privacy regulations that favours the use of smart meter data in meaningful way • Consider defining standard interfaces and data management processes

4.9.3 Going beyond the converted

The insights we collected from the SPP Community Energy Challenge suggest that the energy saving impact we found (the average treatment effect) might be under-estimated, as computed by using a treatment group of households that were already sensitive to energy efficiency measures. On the other hand, they also suggest that “average households” were not spontaneously interested in joining the SPP Community Energy Challenge, and that therefore similar activities are at risk of “preaching to the converted”. The above recommendations focus on how to best develop app-based interventions and tools in order to meet the expectations of the involved users. However, a crucial question also deals with how to go beyond current user types.

One possibility to explore in future research is how the use of SPP-like apps fit within broader participatory processes, that explicitly challenge and question the material and cultural contexts that shape current households' routines. This is for instance aligned with recommendations by Whitmarsh et al. (2021), who critique standard psychological approaches used to tackle climate change wherein “people act alone and in isolation from others. Even social norms are conceived as individual perceptions of expectations and obligations held by the individual” (p. 78). The authors thus call for the “profound and participatory social transformation required to respond to the climate crisis” (p. 78). A similar suggestion for future research and policy-making is also aligned with proposals by Della Valle and Bertoldi (2022) to find “a point of intersection between sociological and individualising approaches” (p. 7). Operatively, principles by Social Practice Theories may support and inform interventions aimed at challenging and experimenting with social norms, cultural and material aspects of everyday life, helping to treat them as “constituents of behaviour”, instead of static “contexts for behaviour to take place” (Jensen et al., 2019). In particular, interventions that create possibilities to discuss, challenge, and re-negotiate collective conventions and norms about daily life and the related practices might be a strategy to ensure that change is maintained in the long-term. As suggested by Hargreaves et al. (2013), instead of “leaving the complex dynamics of energy consumption unquestioned, and thus tacitly supporting and sustaining normal patterns of consumption” (p. 133), intervention features should try to explicitly and collectively question current social practices and their constituents. Such interventions would require active engagement by different social actors: not only households, but also public



institutions and private companies that sustain performance of given practices over time—including those that occur in non-energy related domains—and that can influence and shape their evolution.

A promising approach to perform this type of interventions is offered by living labs: sites for experimentation and open innovation that, through co-creation processes, operate as intermediaries among citizens, research organisations, companies, and government agencies. A review of initiatives activated in the last decade within living labs is offered for instance by Heiskanen et al. (2018). Among the most promising ones are those experimented within the ENERGISE European research project, that launched challenges to re-position heating-related and laundry-related practices in households, with the aim of reducing the related energy demand (Sahakian et al., 2021). To our knowledge, however, those activities only involved household representatives, in a fully bottom-up approach, overlooking key representatives of institutional actors that characterise top-down approaches, as well as intermediaries. Our suggestion would be to also integrate the various stakeholders, in a mixed approach.

Such living labs would entail a number of in-person meetings for a given period of time, engaging voluntary households and also representatives of key public and private institutions. In-person relations would at least initially be needed to create bonds between the participants, to support the creation of feelings of empathy between them, and to consolidate their intrinsic motivation to maintain their engagement. Apps might have a role in such a process, as they might favour the making of public commitments to change practice(s), offer opportunities to manually keep track of practices within the household, monitor (nearly) real-time evolution of household consumption, and offer opportunities for social interaction via digital forums. Apps like Social Power Plus might fruitfully be exploited within living labs; however, they would not be the only tool, or the core of the intervention. Apps would become devices to support groups of households, together with the relevant stakeholders, in the collective re-definition of social conventions around certain energy practices that consume energy. They would not be promoted as a way to improve one's own consumption patterns, but as a tool supporting the collective re-definition of everyday practices.

Besides the discussion on collective conventions of “normality” (the “meanings” of our actions, according to the conceptualisation of social practices by Shove et al., 2012), and opportunities to collectively improve “competences” around novel practices, interaction in living labs would also offer opportunities to re-think, and possibly also tangibly operate on, material structures that shape current and future practices (“materials”). This would be feasible as, besides households, other relevant actors would also be engaged into living lab activities. As those would be offered opportunities to directly engage themselves in the transition process, it is more likely that they would share the sense of urgency and need for collective solutions, thus being more open to implement and try them out.

As participation in living lab activities is highly time and resource intensive, the voluntary, self-selected participants would undoubtedly be already highly motivated individuals—more motivated than the average population, close to the “converted” individuals. However, apps for consumption feedback and monitoring, and technological innovation would not be at the core of living lab activities—to the opposite, they would be used as ancillary tools to enrich and support the transition in social practices by motivated individuals. The higher motivation of the participants in this case would thus be a strength, as, together with the broad living lab network, it would facilitate the tangible evolution in practices, also beyond the living lab.

Fuenfschilling et al. (2019) refer to living labs as ways to “facilitate a process where emerging and fluid ideas, practices, expectations, technologies, and new social relations can develop and align into a new, potentially more sustainable socio-technical configuration, that, if diffused more broadly, will radically alter the existing system” (p. 220). Indeed, living labs might be conceptualised as the “niche” spaces that are identified in the Multi Level Perspective (MLP) by Geels (2004). As noted by Raven et al. (2010), niches can act as “strategic locations for learning, building new social networks and improving the innovation so that it gains momentum for diffusion to other niches or even replace dominant regime practices” (p. 63). The expectation is therefore that, provided that living labs involve a broad group of actors and achieve sufficient size and critical mass for being able to bring about institutional change (Kemp et al., 1998), the novel practices resulting from the living lab, and particularly the novel shared



meanings, supported by novel competences and material structures, would have opportunities to leave the living lab niche in which they were originated and to scale beyond its boundaries, thus spreading to the broader population and to the whole societal system (Von Wirth et al., 2019). The MLP in fact suggests that, through interaction with the dominant regime, and thanks to the effect of factors operating at the landscape level, windows of opportunity might open-up for innovation in practices emerging from the living lab to replace the dominant regime. By prompting a shift in priorities and motivations not only at the individual level but also at the level of social actors, institutions, and their governance strategies, ultimately the changed context created by the Russian war in Ukraine and the related energy crisis, which act at the landscape level, might offer an opportunity for the niche innovation to scale and diffuse beyond the living lab itself, even though the niche has originally been populated by self-selected, highly motivated individuals.

5 Conclusions

In this document we introduced goals and outcomes of an approach aimed at first co-designing an app-based intervention aimed at supporting energy saving in households, and then testing it in real-life conditions with voluntary users. The intervention aimed at promoting both energy sufficient and energy efficient household routines by considering consumption for both heating and non-heating purposes (electric appliances, lighting). It was performed in three regions of German-speaking Switzerland, with the direct involvement by the three utilities SWW (Winterthur), EKS (Schaffhausen) and TBW (Wil).

The features of the Social Power Plus (SPP) app and of the related intervention (the SPP Community Energy Challenge) were developed through a co-design process that involved voluntary customers of the three utilities. Their involvement since the start of the design process allowed to complement theory-based insights for effective behaviour change interventions with the needs and expectations of potential users. We relied on the “Model of Action Phases” (MAP) theoretical framework, according to which behaviour change occurs by means of a sequence of stages, though which individual can progress if they are properly stimulated. By considering both insights from the MAP and previous literature as well as suggestions from the co-design process, through open and creative discussion we identified a set of specific features to support users to progress wherever they are in the behaviour change phases. The co-design process helped our research team to capture relevant user insights to design a more personal intervention, capable to foster new ideas and address contextual motivations, while at the same time firmly leveraging learnings from previous theoretical and applied research.

The resulting SPP app and Community Energy Challenge app were designed to highlight daily energy-consuming household routines and promote social interaction between its users via challenges directly addressing eight energy consuming household routines and pinboard features to facilitate the direct exchange of experiences between participants. The assumption was that social interactions, mediated by the app and supported by goal-setting, challenges, tips, and consumption feedback features, could lead to rethinking the way the routines are performed in the household, for a long-lasting energy saving effect.

The SPP app was used in a field intervention (the SPP Community Energy Challenge), that engaged about 200 self-selected voluntary households living in the regions of the three utilities. The intervention took place from February 1 2022 to December 31 2022. It was characterised by an initial “high intensity” period (three months), during which all the app features were available, and new in-app challenges and tips about specific household routines were launched every two weeks. Then, a “low intensity” period (eight months) took place, during which consumption feedback features were always maintained, but no new challenges were launched. During such a period users could regularly access their consumption data, read the tips and experiences of other participants via the pinboard, and thus stay motivated to keep implementing new ways of performing their routines with less energy.



To estimate the average effect on energy consumption produced by participation in the SPP Community Energy Challenge, we adopted a quasi-experimental approach, by identifying a comparable group of households (control group) among the household customers of the same utilities. Specifically, the control group was identified via propensity score matching statistical techniques, by *ex-post* selecting households with yearly energy consumptions in years 2020 and 2021 (baseline years) that were as similar as possible to those of the treated households. To estimate the effect, we considered overall energy consumption (heating and non-heating purposes) in the two baseline years, as well as consumption in year 2022, namely during the SPP Community Energy Challenge. We opted for a “difference-in-differences” estimator obtained by a fixed effects panel data regression model.

Our findings indicate that on average participation in the SPP Community Energy Challenge produces nearly 5% energy savings, computed over a period of 11 months from the launch of the challenge and considering overall consumption for both heating and non-heating purposes, net of changes in outdoor temperatures and therefore of the evolution of the “total degree-days” indicator.

In the first months after launch of the challenge, the effect was even higher, on average equal to about 11% savings; then, however it gradually decreased over time, probably due to the decrease of the novelty of the app and the interruption of stimuli due to the “low intensity” period. We tested whether such a decrease was due to seasonal effects (the evolution of the energy demand for heating and cooling during Winter and Summer months), but test results led us to reject this hypothesis. We also looked for differences between the regions, which could mostly reflect differences in the energy provision and metering infrastructures between the heating system (gas vs heat pump), and could not find statistically significant differences between the sub-groups. This suggests that the effect of participation to the SPP Community Energy Challenge is consistent across different contexts, particularly between regions and infrastructural characteristics. Unfortunately, no additional analyses on the heterogeneity of effects could be performed as only energy consumption data was available for the control group households.

The estimates of the average treatment effect are consistent with the before and after survey results on the changes to self-reported household routines. We found that, after use of the app, some household routines became more energy sufficient, in particular the change in thermostat settings for heating. The statistical significance and effect size of this change in heating-related routines suggest that it definitely played a role in driving the overall energy saving effect due to participation in the SPP Community Energy Challenge. Other routines, instead, appear not to have changed due to participation. The lack of change in routines around showering and laundry washing, in particular, hints at the difficulty of shifting cultural conventions on cleanliness-related topics, and might also be due to the bias in gender composition of the survey respondents and household representatives, that was largely biased towards middle-aged males. We suggest complementary approaches (vs. purely app-based) to promote a shift in socially shared norms and conventions about perceptions of normality around cleanliness and the related energy consumption. Complementary approaches might also manage to attract a more balanced composition of the sample of household representatives regarding gender, age, and social status, which might increase the chances that such a shift actually takes place. Regarding the features of the SPP app and Community Energy Challenge, our findings suggest that the average treatment effect is marginally related with the social interaction processes focusing on household daily routines that were enabled and mediated by the SPP pinboard. Survey responses suggest that individual feedback on energy consumption and the goal setting feature have played a major role in driving the effect. Most users also declared their eagerness for even more detailed information on energy consumption feedback with the expectation that it allows them to further optimise their consumption and well-being with limited routine change. The attempt to foster change through sharing via the pinboard did not prove successful within the set-up of the SPP Community Energy Challenge.

This result may have been influenced by two key factors. First, the social interaction features were mediated by the SPP app and were expected to occur between strangers that had no previous connection. Future interventions might therefore preferably address households that already have personal connections. However, if the app aims at addressing energy consumption for heating purposes



for a broader population, households living in building blocks with centralised heating systems would need different measurement technology for individual consumption feedback. This is coherent with sustainability transition theory showing the need for whole (socio-technical) system change to support changing practices and behaviour (Markard et al., 2012, Geels, 2023). Second, the specific characteristics of the app users may have influenced their expectations regarding the SPP app and the way they interacted with it. This is likely to have played a relevant role considering that participants to the SPP Community Energy Challenge self-selected themselves based on the recruitment communication material, where getting feedback on individual energy consumption was presented as a prominent incentive, and the social interaction was less in focus.

Finally, our analyses suggest that most users differ from the target households we expected to engage and for which the features of the SPP app and Community Energy Challenge were designed: they already were implementing some degree of energy efficiency measures in their house. Overall, the collected data suggests high affinity between many of the participating households and the concept of “resource man” by Strengers (2014) (individuals that are interested in and understand their energy data, mostly males), despite the project trying to go beyond these types of users.

Insights from the SPP experience thus suggest that no matter the specific features they offer, app-based change processes tend to attract “resource men” and do not manage to raise the interest by an audience who would probably benefit the most from the app. There is, therefore, a risk that the SPP Community Energy Challenge “preached to the converted”.

The “resource men” that participated in the SPP Community Energy Challenge asked for highly disaggregated real-time feedback in both the co-design process and survey feedback. Considering current technology and regulations, as well as the motivations and constraints of utility companies and the energy system operators, it is quite challenging to fulfil this request. In Section 4.9 therefore we recommended ways to combine these expectations while ensuring that the app development and challenge management can financially self-sustain itself. However, not all participants were “resource men”: a small group of participants to the SPP Community Energy Challenge in fact actively interacted with the pinboard and challenges, and shared experiences and knowledge about saving energy at home. This implies there is a potential for technical knowledge and competence sharing that might be explicitly leveraged for sustainability transition processes, for instance for personal peer-to-peer consultancy processes –either mediated by an app or by in-person interactions in living lab processes. The community-based app features implemented through the SPP challenges and the pinboard might therefore have a potential for the specific target group of such users. Namely, households (and houses) are very diverse in how energy is used and what they are interested in. Further applied research might aim at profiling the households that interacted the most with challenges and pinboard features to understand the effect on this sub-group. If the effect is significant and relevant, research should be performed to understand their characteristics and reasons for engagement in the SPP experiment and thus explicitly target (and reach) them through improved communication activities.

A promising approach might consist in giving less prominence to the presence and role of the app and rather to explicitly address communities of households, the social interaction aspects, and the experimentation with new sustainable practices, as typically performed in living lab approaches (Laakso et al., 2021; Sahakian et al., 2021). We expect that these approaches could also engage “non-converted” and “non resource men” households, as the apps would not be the core and only intervention measure. Rather, these living lab processes would explicitly focus on collectively questioning and challenging current practices and routines. Apps can thus be a supportive element to provide feedback and show the effects of changed routines. Future applied research will tell if, and to what extent, these approaches can tangibly support large-scale change in households’ energy consumption.

6 Outlook and next steps



In the previous chapters we have summarised the goals and results of the SPP project, also discussing their limitations. We also briefly hinted at additional analyses that would be worth performing to further understand the effect that SPP participation produces on energy consumption. These analyses would be particularly relevant to identify possible user groups to be directly targeted in future similar app-based interventions.

We have also already indicated possible strategies to stimulate a broader engagement by “non resource men” and support the participation by previously non-engaged households, namely the living lab approach and the explicit focus on questioning current household practices and the dominant concepts of “what is normal” around them. We therefore suggest that future research explicitly explores these approaches and their effect, adopting strict evaluation procedures. Ideally, analysis of the effect of living lab approaches should be performed under experimental settings, via randomisation of the allocation to treatment and control groups.

Finally, a key element for future research is the maintenance of the treatment effect in the long-term. Our analyses in fact found that participation in the SPP Community Energy Challenge resulted in a nearly 5% decrease of energy consumption over one year. As the SPP intervention concluded in May 2022, we could not verify yet if such an effect is maintained over a longer period of time or not. Knowing how persistent the effect is, is in fact essential to support policy-makers in the current context of limited financial and human resources. If evidence of impact is found during or immediately after app use, but it is not maintained one or two years after its use has been discontinued, investing in their development and deployment would be questionable, especially if funded by public money. If so, alternative intervention strategies might be preferable. Research that some of our team members have just published (Cellina et al., 2024) has shown that, for another app-based intervention, the relevant and significant saving effect found during a one-year long intervention, had disappeared the following year. This may not be the case of the SPP Community Energy Challenge, however research is still needed to ensure long-lasting effects are produced.

7 National and international cooperation

The whole SPP project is based on trans-disciplinary research activities based on collaboration between partners from SUPSI, ZHAW, and CLEMAP, and cooperation partners from the three utilities EKS, SWW, and TBW. Apart for the workshop presentations reported in the next section, no additional international collaborations have been activated during the project.

8 Communication

The project website, available at <http://www.socialpower.ch>, has been constantly updated in both German and English, with the aim of allowing any interested individuals, either already participating in project activities (living lab, SPP Community Energy Challenge), or simply interested in its goals and content, to find the relevant information. Short posts targeting both participants to the SPP Community Energy Challenge as well as a broader audience made of practitioners, policy makers, and entrepreneurs in smart city topics, have in fact been regularly published on the project website. Similarly, also CLEMAP published a blogpost explicitly addressing the disaggregation of heat pumps and the knowledge acquired under the SPP project (<https://www.clemap.com/blog/lastaufschlusselung-wp>).

Additional dissemination of project activities was performed at two levels:



- within the regions of the three utilities, mostly with the aim of recruiting participants to living lab activities and to the SPP Community Energy Challenge;
- and within the scientific community, in order to present the goals and approach of the SPP project, present intermediate results obtained, and collect suggestions for future activities.

Besides the project website, for recruitment the utilities made their communication tools and online social networks available to present the project. In particular, a short article introducing SPP was published on the EKS magazine “E on!”, with an interview to Evelyn Lobsiger-Kägi of the ZHAW team.

For dissemination within the scientific community, SPP-related activities were instead presented at three conferences. In 2021, Francesca Cellina of the SUPSI team offered an oral presentation at the 6th European Conference on Behaviour Change for Energy Efficiency (Behave 2021) organized by the Copenhagen Centre on Energy Efficiency; Devon Wemyss of the ZHAW team offered an oral presentation at the International Sustainability Transition (IST) 2021 conference, hosted by the Fraunhofer Institute for Systems and Innovation Research of Karlsruhe). Both contributions are included in the conference proceedings. The manuscript submitted at the IST conference, already developed as a full paper, was further elaborated and submitted to the Energy Research and Social Science journal, in which it was then published in Summer 2023 (see details in the next section).

In 2022, Francesca Cellina of the SUPSI team presented SPP preliminary results in a three-day workshop organized by the Monash University and hosted in their European offices in Prato (Italy, October 24-26, 2022) about “Household Innovation and Agency in Sustainability Transitions”. The venue offered the opportunity for a 45 minute intense presentation and discussion of the full paper manuscript titled “Social learning in household communities: Can energy-sufficient routines be fostered by means of app-based peer-to-peer interaction?”. A revised version of the manuscript has been submitted in October 2023 to a special issue about “Household innovation and agency in sustainability transitions” to be published in the Environmental Innovation and Societal Transitions journal, and it is currently under review.

Communication targeting a broad public of innovators, entrepreneurs, start-ups and researchers was also performed during the National Open Innovation Camp Lugano “NOI 2022”, which took place in Lugano in September 2022 (<https://luganolivinglab.ch/en/projects/national-open-innovation-camp-2022/#programma>). Particularly, during a two-hour workshop session the SPP was presented and discussed with NOI participants, who were engaged to find novel and effective ways to connect with “hard to reach households”.

Finally, within their energy saving campaign, in 2022 ZHAW produced and published a short video (<https://gemeinsam.zhaw.ch>), also shared in different Social Media channels) starring team member Evelyn Lobsiger-Kägi, where she talks about the largest energy saving potentials in the household and how we can unlock them, which was nourished by the experience we made in the SPP project. The SRF also held an interview in the context of “Kassensturz” with Pascal Kienast of the ZHAW team, where the experiences of Social Power Plus have been shared to a broader public (<https://www.srf.ch/sendungen/kassensturz-espresso/kassensturz/schleppende-digitalisierung-smartmeter-nutzung-noch-nicht-intelligent-genug>). As a consequence of this interview two utilities got in touch with the project to have further insights about project activities.

Now that the final project results have been obtained, the research team will work on at least an additional research manuscript, to be submitted to scientific journals such Energy Research and Social Science, Energy Policy, Sustainable Cities and Society, Applied Energy, or Computers in Human Behavior – all journals that the team members are experienced with, as they have previously published their work on them. Additional dissemination of the project results will be performed via presentation to scientific conferences and venues about smart cities and open innovation. Finally, we would like to summarise project results in a publication on the Energeia online magazine edited by SFOE.



9 Publications

Conference presentations and proceedings:

- Cellina, F., Carabias-Hütter, V., Castri, R., De Luca, V., Granato, P., Kienast, P., Lobsiger-Kägi, E., Wemyss, D., Social Power Plus: Empowering Households to Energy Sufficiency through Co-designed App-based Community Energy Challenges. In: Zhu, X., Prata Dias, G. (Eds.), Conference proceedings BEHAVE 2020-2021 - The 6th European Conference on Behaviour Change for Energy Efficiency, April, 21-23 2021, Online (hosted by Copenhagen Centre on Energy Efficiency). ISBN: 978-87-94094-01-6. <https://c2e2.unepdtu.org/wp-content/uploads/sites/3/2021/04/behave-2020-2021-conference-proceedings.pdf>
- Wemyss, D., Lobsiger-Kägi, E., Jud, S., Cellina, F. (2021). My energy transition at home: Results of co-creating an energy savings app with household members and researchers. International Sustainability Transition (IST) 2021, October 2021.
- Cellina, F., Lobsiger-Kägi, E., Wemyss, D., Profeta, G., Granato, P. (2022). Social learning in household communities: Can energy-sufficient routines be fostered by means of app-based peer-to-peer interaction? In: International Workshop on Household Innovation and Agency in Sustainability Transitions, October, 24-26 2022, Prato - Monash University Europe.
- Wemyss, D., Cellina, F., Lobsiger-Kägi, E., Kienast, P., Granato, P., Gerosa, T. When data doesn't motivate: potentials and limitations of app-based feedback for sustainable energy consumption. Contribution submitted to the eceee 2024 Summer Study, Accepted as oral presentation.

Scientific articles:

- Wemyss, D., Lobsiger-Kägi, E., Jud, S., Cellina, F. (2023). Leveraging realities of saving energy at home: Contributions of co-design to behavioural interventions, *Energy Research & Social Science*, Volume 104, 2023, 103258, ISSN 2214-6296, <https://doi.org/10.1016/j.erss.2023.103258>
- Cellina, F., Lobsiger-Kägi, E., Wemyss, D., Profeta, G., Granato, P. Households in energy transition: Promoting household energy-sufficient routines via app-based peer-to-peer interaction. Submitted to *Environmental Innovation and Societal Transitions*, Special Issue on "Household innovation and agency in sustainability transitions". Under review.

Additional scientific articles dealing with the estimates of the quantitative impact of the SPP Community Energy Challenge will be submitted to scientific journals in the next months.

10 References

- Abras, C., Maloney-Krichmar, D., Preece, J. (2004). User-centered design. In *Encyclopedia of Human-Computer Interaction* (pp. 445–456). Thousand Oaks: SAGE Publications.
- Ahorsu, D.K., Lin, C.Y., Imani, V. et al. (2022). The Fear of COVID-19 Scale: Development and Initial Validation. *Int J Ment Health Addiction* 20, 1537–1545.
- Allcott, H., Mullainathan, S. (2010). Behavior and Energy Policy. *Science*. Mar 5;327(5970):1204–5.
- Allcott, H. (2011). Social norms and energy conservation. *Journal of public Economics*, 95(9-10), 1082-1095.
- Allcott, H., Rogers, T. (2014). The Short-Run and Long-Run Effects of Behavioral Interventions: Experiment Evidence From Energy Conservation. *American Economic Review*. 104(10):3003–37.
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*. 50:179–211.



- Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate behavioral research*, 46(3), 399-424.
- Bamberg, S. (2013). Changing environmentally harmful behaviors: A stage model of self-regulated behavioral change. *Journal of Environmental Psychology*, 34(2013):151–9.
- Baptista, G., Oliveira, T. (2019). Gamification and serious games: A literature meta-analysis and integrative model. *Computers in Human Behavior*. 92:306–15.
- Baxter, J. (2020). Content Analysis. In: Kobayashi A, editor. *International Encyclopedia of Human Geography (Second Edition)*. Oxford: Elsevier; p. 391–6.
- Bear, J.B., Woolley, A.W. (2011). The role of gender in team collaboration and performance. *Interdisciplinary Science Reviews*. 36(2):146–53.
- Bergquist, M., Nilsson, A., Schultz, W.P. (2019). A meta-analysis of field-experiments using social norms to promote pro- environmental behaviors. *Global Environmental Change*. 58:101941.
- Bergvall-Kåreborn, B., Howcroft, D., Ståhlbröst, A., Wikman, A. M. (2010). Participation in Living Lab: Designing Systems with Users. In J. Pries-Heje, J. Venable, D. Bunker, N. L. Russo, J. I. DeGross (Eds.), *Human Benefit through the Diffusion of Information Systems Design Science Research: IFIP WG 8.2/8.6 International Working Conference, Perth, Australia, March 30-April 1, 2010*. Proceedings (pp. 317–326). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Bird, S., Legault, L. (2018). Feedback and behavioral intervention in residential energy and resource use: a review. In: *Current Sustainable/Renewable Energy Reports* 5.1, pp. 116–126.
- Böhm, S., Szewc, L. (2013). Smart metering with smartphones: user-centered design of a mobile application in the context of energy efficiency. In *International Conference of Design, User Experience, and Usability* (pp. 631–640). Springer.
- Breukers, S. C., Mourik, R. M. (2013). *The end-users as starting point for designing dynamic pricing approaches to change household energy consumption behaviours*. Arnhem.
- Brown, M.A., Sovacool, B.K. (2018). Theorizing the Behavioral Dimension of Energy Consumption: Energy Efficiency and the Value-Action Gap. *The Oxford Handbook of Energy and Society*. 201.
- Buchanan, K., Russo, R., Anderson, B. (2015). The question of energy reduction : The problem(s) with feedback. *Energy Policy*. 77:89–96.
- Bugden, D., Stedman, R. (2017). A synthetic view of acceptance and engagement with smart meters in the United States. *Energy Research and Social Science*. 47:137–45.
- Bundesamt für Energie (2017). *Analyse des schweizerischen Energieverbrauchs 2000 - 2016 nach Verwendungszwecken*.
- Butler, C., Parkhill, K.A., Lizecka, P. (2018). Rethinking energy demand governance: Exploring impact beyond 'energy'policy. *Energy Research & Social Science* 36, pp. 70–78.
- Caliendo, M., Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of economic surveys*, 22(1), 31-72.
- Chatzigeorgiou, I.M., Andreou, G.T. (2021). A systematic review on feedback research for residential energy behavior change through mobile and web interfaces. *Renewable and Sustainable Energy Reviews*. 135:110187.
- Cellina, F., Fraternali, P., Herrera Gonzalez, S.L., Novak, J., Gui, M., Rizzoli, A.E. (2024). Significant but transient: The impact of an energy saving app targeting Swiss households, *Applied Energy*, Volume 355, 122280.
- Chan, R.C.H. (2022). A social cognitive perspective on gender disparities in self-efficacy, interest, and aspirations in science, technology, engineering, and mathematics (STEM): the influence of cultural and gender norms. *International Journal of STEM Education*. 25;9(1):37.
- Chay, K. Y., McEwan, P. J., Urquiola, M. (2005). The central role of noise in evaluating interventions that use test scores to rank schools. *American Economic Review*, 95(4), 1237-1258.



- Cialdini, R. B., Reno, R. R., Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of personality and social psychology*, 58(6), 1015.
- Cockbill, S.A., Mitchell, V., May, A.J. (2020). Householders as designers? Generating future energy services with United Kingdom home occupiers. *Energy Research and Social Science*. 69:101615.
- Cockbill, S.A., May, A., Mitchell, V. (2019). The Assessment of Meaningful Outcomes from Co-design: A Case Study from the Energy Sector. *She Ji: The Journal of Design, Economics, and Innovation*. 5(3):188–208.
- Dalén, A., Krämer, J. (2017). Towards a User-Centered Feedback Design for Smart Meter Interfaces to Support Efficient Energy-Use Choices. *Business & Information Systems Engineering*, 59(5), 361–373.
- Darby, S. (2010). Smart metering: what potential for householder engagement? *Building Research & Information*. 38(5):442–57.
- Darby, S. (2020). Demand response and smart technology in theory and practice: Customer experiences and system actors. *Energy Policy*. 43:111573.
- Della Valle, N., Bertoldi, P. (2021). Mobilizing citizens to invest in energy efficiency. EUR–Scientific and Technical Research Reports. Ispra: Joint Research Centre.
- Dell’Era, C., Landoni, P. (2014). Living Lab: A methodology between user-centred design and participatory design. *Creativity and Innovation Management*, 23(2), 137–154.
- Delmas, M. A., Fischlein, M., Asensio, O. I. (2013). Information strategies and energy conservation behavior: A meta-analysis of experimental studies from 1975 to 2012. *Energy Policy*, 61, 729–739.
- Devine-Wright, P. (2012). Energy citizenship: psychological aspects of evolution in sustainable energy technologies. In *Governing technology for sustainability* (pp. 74-97). Routledge.
- Fischer, C. (2008). Feedback on household electricity consumption: A tool for saving energy? *Energy Efficiency*, 1(1), 79–104.
- Fraternali, P., Cellina, F., Herrera Gonzales, S.L., Melenhorst, M., Novak, J., Pasini, C., et al. (2019). Visualizing and gamifying consumption data for resource saving: challenges, lessons learnt and a research agenda for the future. *Energy Informatics*. 2(S1):1–13.
- Frederiks, E.R., Stenner, K., Hobman, E.V. (2015). Household energy use: Applying behavioural economics to understand consumer decision-making and behaviour. *Renewable and Sustainable Energy Reviews*. 41:1385–94.
- Frey, E., Rogers, T. (2014). Persistence: How Treatment Effects Persist After Interventions Stop. *Policy Insights from the Behavioral and Brain Sciences*. 1(1):172–9.
- Fuenfschilling, L., Frantzeskaki, N., Coenen, L. (2019). Urban experimentation & sustainability transitions. *European Planning Studies* 27.2, pp. 219–228.
- Geels, F.W. (2004). From sectoral systems of innovation to socio-technical systems: Insights about dynamics and change from sociology and institutional theory. *Research policy* 33.6-7, pp. 897–920.
- Geels, F.W. (2023). Demand-side emission reduction through behaviour change or technology adoption? Empirical evidence from UK heating, mobility, and electricity use. *One Earth*, 6(4), 337-340)
- Georges, A., Schuurman, D., Vervoort, K. (2016). Factors Affecting the Attrition of Test Users During Living Lab Field Trials. *Technology Innovation Management Review*. 6(1):35–44.
- 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):6111–9.
- Hargreaves, T., (2018). Beyond energy feedback. *Building Research & Information* 46.3, pp. 332–342.
- Hand, M., Shove, E., Southerton, D. (2005). Explaining showering: A discussion of the material, conventional, and temporal dimensions of practice. *Sociological Research Online*, 10(2), 101-113.
- Heckhausen, H., Gollwitzer, P.M. (1987). Thought contents and cognitive functioning in motivational versus volitional states of mind. *Motiv Emot*. 11(2):101–20.



- Heiskanen, E., Laakso, S., Matschoss, K., Backhaus, J., Goggins, G., Vadovics, E. (2018). Designing real-world laboratories for the reduction of residential energy use: Articulating theories of change. *GAIA-Ecological Perspectives for Science and Society*, 27(1), 60-67.
- Hekler, E.B., Michie, S., Pavel, M., Rivera, D.E., Collins, L.M., Jimison, H.B., et al. (2016). Advancing Models and Theories for Digital Behavior Change Interventions. *American Journal of Preventive Medicine*. 51(5):825–32.
- Herrmann, M. R., Brumby, D. P., Oreszczyn, T., Gilbert, X. M. (2018). Does data visualization affect users' understanding of electricity consumption?. *Building Research & Information*, 46(3), 238-250.
- Honebein, P.C., Cammarano, R.F., Donnelly, K.A. (2009). Will Smart Meters Ripen or Rot? Five First Principles for Embracing Customers as Co-Creators of Value. *Electricity Journal*. 22(5):39–44.
- Hurley, E., Dietrich, T., Rundle-Thiele, S. (2021). Integrating Theory in Co-design: An Abductive Approach. *Australasian Marketing Journal*. 29(1):66–77.
- Itten, A., Sherry-Brennan, F., Sundaram, A., Hoppe, T., Devine-Wright, P. (2020). State-of-the-art report for co-creation approaches and practices with a special focus on the sustainable heating transition: Shiftt work package 2 deliverable 2.1.1. Provincie Zuid-Holland.
- Itten, A., Sherry-Brennan, F., Hoppe, T., Sundaram, A., Devine-Wright, P. (2021). Co-creation as a social process for unlocking sustainable heating transitions in Europe. *Energy Research and Social Science*. 74(March):101956.
- Jensen, C. L., Goggins, G., Røpke, I., Fahy, F. (2019). Achieving sustainability transitions in residential energy use across Europe: The importance of problem framings. In: *Energy Policy* 133, p. 110927.
- Jungell-Michelsson, J., Heikkurinen, P. (2022). Sufficiency: A systematic literature review. *Ecological Economics* 195, p. 107380.
- Karlin, B., Zinger, J. F., Ford, R. (2015). The effects of feedback on energy conservation: A meta-analysis. *Psychological Bulletin*, 141(6), 1205–1227.
- Kaufman, S., Saeri, A., Raven, R., Malekpour, S., Smith, L. (2021). Behaviour in sustainability transitions: A mixed methods literature review. *Environmental Innovation and Societal Transitions*, 40, 586-608.
- Kemmler, A., Spillmann, T. (2017). *Energieverbrauch der Privaten Haushalte 2000–2020*, SFOE, Swiss Federal Office for Energy.
- Kemp, R., Schot, J., Hoogma, R. (1998). Regime shifts to sustainability through processes of niche formation: the approach of strategic niche management 10.2, pp. 175–198.
- Laakso, S., Jensen, C. L., Vadovics, E., Apajalahti, E. L., Friis, F., Szöllőssy, A. (2021). Towards sustainable energy consumption: Challenging heating-related practices in Denmark, Finland, and Hungary. *Journal of Cleaner Production*, 308, 127220.
- Lee, E., Kang, M., Song, J., Kang, M. (2020). From intention to action: Habits, feedback and optimizing energy consumption in South Korea. *Energy Research and Social Science*. 64:101430.
- Lindner, S., & McConnell, K. J. (2019). Difference-in-differences and matching on outcomes: a tale of two unobservables. *Health Services and Outcomes Research Methodology*, 19, 127-144.
- Marek, R., Bossart, R., Marek, K., Derungs, C., Winterberger, S. (2020). *EVISU: Energievisualisierung Bottom-Up*, Final Report for the Swiss Federal Office of Energy.
- Magnusson, P.R. (2003). Benefits of involving users in service innovation. *European Journal of Innovation Management*. 6(4):228–38.
- Markard, J., Raven, R., Truffer, B. (2012). Sustainability transitions: An emerging field of research and its prospects. *Research policy*, 41(6), 955-967.



- Mastelic, J., Genoud, S., Lasvaux, S., Padey, P., Périsset, B., Farsi, M., Weber, S. (2018). Projet UserGap: Influence des utilisateurs sur l'écart de performance dans les bâtiments collectifs à haute performance énergétique. Annual Report for the Swiss Federal Office of Energy.
- Matschoss, K., Fahy, F., Rau, H., Backhaus, J., Goggins, G., Grealis, E., ... Vasseur, V. (2021). Challenging practices: experiences from community and individual living lab approaches. *Sustainability: Science, Practice and Policy*, 17(1), 135-151.
- Mengolini, A., Gangale, F., Vasiljevska, J. (2016). Exploring community-oriented approaches in demand side management projects in Europe. *Sustainability*, 8(12), 1266.
- Michie, S., van Stralen, M.M., West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*. 6(1).
- Mitchell, V., Ross, T., May, A., Sims, R., Parker, C. (2016). Empirical investigation of the impact of using co-design methods when generating proposals for sustainable travel solutions. *CoDesign*. 12(4):205–20.
- Moser, C., Rösch, A., Stauffacher, M. (2015). Exploring societal preferences for energy sufficiency measures in Switzerland. In: *Frontiers in Energy Research* 3, p. 40.
- Mourik, R., Breukers, S. (2015). Did you behave as we designed you to? Monitoring and evaluating behavioural change in demand side management: from what to why. In *ECEEE Summer Study* (pp. 1881–1892). ECEEE.
- Murtagh, N., Gatersleben, B., Uzzell, D. (2014). 20 60 20-Differences in energy behaviour and conservation between and within households with electricity monitors. In: *PLoS one* 9.3, e92019.
- Nilsson, A., Jakobsson Bergstad, C., Thuvander, L. et al. (2014). Effects of continuous feedback on households' electricity consumption: Potentials and barriers. In: *Applied Energy* 122, pp. 17–23.
- Nikas, A., Lieu, J., Sorman, A., Gambhir, A., Turhan, E., Baptista, B.V., et al. (2020). The desirability of transitions in demand: Incorporating behavioural and societal transformations into energy modelling. *Energy Research and Social Science*. 70(October).
- Ohnmacht, T., Schaffner, D., Weibel, C., Schad, H. (2017). Rethinking social psychology and intervention design: A model of energy savings and human behavior. *Energy Research & Social Science*, 26, 40–53.
- Pallot, M., Trousse, B., Senach, B., Scapin, D. (2010). Living lab research landscape: From user centred design and user experience towards user cocreation. First European Summer School Living Labs.
- Preston, S., Mazhar, M.U., Bull, R. (2020). Citizen Engagement for Co-Creating Low Carbon Smart Cities: Practical Lessons from Nottingham City Council in the UK. *Energies*. 13(24):6615.
- Puntiroli, M., Bezençon, V. (2020). Feedback devices help only environmentally concerned people act pro-environmentally over time. *Journal of Environmental Psychology* 70, p. 101459.
- Raven, R., Reynolds, D., Lane, R., Lindsay, J., Kronsell, A., Arunachalam, D., (2021). Households in sustainability transitions: A systematic review and new research avenues. *Environmental Innovation and Societal Transitions*, 40, 87-107.
- Raven, R., Van den Bosch, S., Weterings, R. (2010). Transitions and strategic niche management: Towards a competence kit for practitioners. *International Journal of Technology Management* 51.1, pp. 57–74.
- Reed, M.S., Vella, S., Challies, E., de Vente, J., Frewer, L., Hohenwallner-Ries, D., et al. (2018). A theory of participation: what makes stakeholder and public engagement in environmental management work? *Restoration Ecology*. 26:S7–17.
- Rosenbaum, P. R., Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55.
- Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, 39(1), 33-38.



- Royston, S., Selby, J., Shove, E. (2018). Invisible energy policies: A new agenda for energy demand reduction. *Energy Policy* 123, pp. 127–135.
- Ryan, A. M., Burgess Jr, J. F., Dimick, J. B. (2015). Why we should not be indifferent to specification choices for difference-in-differences. *Health services research*, 50(4), 1211-1235.
- Sahakian, M., Rau, H., Grealis, E., Godin, L., Wallenborn, G., Backhaus, J., ..., Fahy, F. (2021). Challenging social norms to recraft practices: A Living Lab approach to reducing household energy use in eight European countries. *Energy Research and Social Science*, 72, 101881.
- Sanders, E. B. N. (2003). From user-centered to participatory design approaches. *Design and the social sciences*. CRC Press 18–25.
- Sanders, E.B.N., Stappers, P.J. (2014). Probes, toolkits and prototypes: three approaches to making in codesigning. *CoDesign*. 10(1):5–14.
- Sandmeier, E. et al. (2020). 2000-Watt-Gesellschaft leben: Reduktion des End-Energieverbrauchs durch Verhaltensänderungen, SFOE P+D project.
- Schuler, D., Namioka, A. (Eds.). (1993). *Participatory design: Principles and practices*. CRC Press.
- Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., Griskevicius, V. (2007). The Constructive, Destructive, and Reconstructive Power of Social Norms. *Psychological Science*, 18(5), 429.
- Schwartz, S.H., Howard, J.A. (1981). A normative decision-making model of altruism. *Altruism and helping behavior*. 189–211.
- Seidl, R., Moser, C., Blumer, Y. (2017). Navigating behavioral energy sufficiency. Results from a survey in Swiss cities on potential behavior change. In: *PloS one* 12.10, e0185963
- Sergici, S., Faruqui, A. (2011). *Measurement and verification principles for behavior based efficiency programs*. Brattle Group: Cambridge, MA, USA.
- SFSO, Swiss Federal Statistical Office (2017). *Erhebung der Energieträger von Wohngebäuden*.
- SFSO, Swiss Federal Statistical Office (2022a). *Current situation and change*.
- SFSO, Swiss Federal Statistical Office (2022b). *Federal Register of Buildings and Dwellings*.
- Shove, E. (2003). Converging conventions of comfort, cleanliness and convenience. *Journal of Consumer policy*, 26(4), 395-418.
- Shove, E., Pantzar, M., Watson, M. (2012). *The dynamics of social practice: Everyday life and how it changes*. Sage.
- Sovacool, B. K. (2014). What are we doing here? Analyzing fifteen years of energy scholarship and proposing a social science research agenda. *Energy Research and Social Science*, 1, 1–29.
- Sovacool, B., Osborn, J., Martiskainen, M., Lipson, M. (2020). Testing smarter control and feedback with users: Time, temperature and space in household heating preferences and practices in a Living Laboratory. In: *Global Environmental Change* 65,p. 102185.
- Sovacool, B., Demski, C., Noel, L. (2021). Beyond climate, culture and comfort in European preferences for low-carbon heat. In: *Global Environmental Change* 66, p. 102200.
- Staats, H., Harland, P., Wilke, H.A.M. (2004). Effecting durable change: A team approach to improve environmental behavior in the household. *Environment and Behavior* 36.3, pp. 341–367.
- Statista (2021). *Home ownership rate in selected European countries in 2019 and 2020, by country*.
- Steen, M., Manschot, M., de Koning, N. (2011). Benefits of Co-design in Service Design Projects. *International Journal of Design*. 5(2):53–60.
- Steen, M. (2013). Co-Design as a Process of Joint Inquiry and Imagination. *Design Issues*.29(2):16–28.
- Strengers, Y. (2014). Smart energy in everyday life: Are you designing for resource man? *Interactions*. 21(4):24–31.
- Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical science: a review journal of the Institute of Mathematical Statistics*, 25(1), 1.



- Vine, E., Sullivan, M., Lutzenhiser, L., Blumstein, C., Miller, B. (2014). Experimentation and the evaluation of energy efficiency programs. *Energy Efficiency* 7.4, pp. 627–640.
- Visser, F.S., Stappers, P.J., Lugt, R.V.D., Sanders, E.B. (2005). Contextmapping: experiences from practice. *CoDesign*. 1(2):119–49.
- Von Wirth, T., Fuenfschilling, L., Frantzeskaki, N., Coenen, L. (2019). Impacts of urban living labs on sustainability transitions: Mechanisms and strategies for systemic change through experimentation. *European Planning Studies* 27.2, pp. 229–257.
- Webb, L., Morgan, E., Carter, K., Webb, J., Goddard, N. (2016). A Living Lab Co-Creational Approach To Energy Demand Reduction in Non-Domestic Buildings: Understanding the Organisation. In: *BEHAVE 2016 - 4th European Conference on Behaviour and Energy Efficiency*. p. 1–13.
- Weber, S., Burger, P., Farsi, M. et al. (2017). Swiss household energy demand survey (SHEDS): Objectives, design, and implementation. Tech. rep. IRENE Working Paper.
- Wemyss, D., Castri, R., Cellina, F., Luca, V. De, Lobsiger-kägi, E., Carabias, V. (2018). Examining community-level collaborative vs competitive approaches to enhance household electricity-saving behavior. *Energy Efficiency*.
- Wemyss, D., Cellina, F., Lobsiger-Kägi, E., de Luca, V., Castri, R. (2019). Does it last? Long-term impacts of an app-based behavior change intervention on household electricity savings in Switzerland. *Energy Research and Social Science*, 47, 16-27.
- Wever, R., Kuijk, J. Van, Boks, C. (2008). User-centred design for sustainable behaviour. *International Journal for Sustainable Engineering*, 1(1), 9–20.
- Whitmarsh, L., Poortinga, W., Capstick, S. (2021). Behaviour change to address climate change. In: *Current Opinion in Psychology* 42, pp. 76–81.
- Wilhite H, Shove E, Lutzenhiser L, Kempton W. (2006). The Legacy of Twenty Years of Energy Demand Management: we know more about Individual Behaviour but next to Nothing about Demand. In: *Society, Behaviour, and Climate Change Mitigation*. p. 109–26.
- Wing, C., Simon, K., Bello-Gomez, R. A. (2018). Designing difference in difference studies: best practices for public health policy research. *Annual review of public health*, 39, 453-469.
- Wood, G., van der Horst, D., Day, R., Bakaoukas, A.G., Petridis, P., Liu, S. (2014). Serious games for energy social science research. *Technology Analysis and Strategic Management*. 2014;26(10):1212–27.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.
- Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*. Cengage learning.
- Wooldridge, J. M. (2021). Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators.
- Zeldin, A.L., Britner, S.L., Pajares, F. (2008). A comparative study of the self-efficacy beliefs of successful men and women in mathematics, science, and technology careers. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*. 45(9):1036–58.
- Zhang, Z., Kim, H. J., Lonjon, G., Zhu, Y. (2019). Balance diagnostics after propensity score matching. *Annals of translational medicine*, 7(1).
- Zhao, Q. Y., Luo, J. C., Su, Y., Zhang, Y. J., Tu, G. W., Luo, Z. (2021). Propensity score matching with R: conventional methods and new features. *Annals of translational medicine*, 9(9).