SCUOLA DI DOTTORATO
UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

Department of

Sociology and Social Research

PhD programme in Analysis of Socio-Economic Processes (ASEP) Cycle XXXV

HOME SWEET *ENERGY-SUFFICIENT* HOME: THE POTENTIAL OF PERSUASIVE APPS FOR HOUSEHOLDS' ENERGY AND CLIMATE TRANSITION

Cellina Francesca

Registration number 854291

Tutor: Professor Marco Gui

Co-tutors: Professor Giuseppe Vittucci Marzetti, Dr. Roman Rudel

Coordinator: Professor Maurizio Pisati

ACADEMIC YEAR 2021/2022

Francesca Cellina, registration number 854291
Home Sweet Energy-Sufficient Home:
The potential of persuasive apps
for households' energy and climate transition
Doctoral Dissertation, Academic Year 2021/2022
Supervisors: Professor Marco Gui (Tutor), Professor Giuseppe Vittucci Marzetti,
Dr. Roman Rudel (Co-tutors) and Professor Maurizio Pisati (Coordinator)

Doctoral School of the University of Milano-Bicocca
Department of Sociology and Social Research
Piazza dell'Ateneo Nuovo
20126 Milano (Italy)

Abstract

Climate protection goals require a transition in all energy consumption domains. In this work I focus on residential energy demand and assess the effects of persuasive smartphone apps promoting energy sufficiency (i.e. a reduction in the absolute amounts of energy demand, aimed at meeting people's basic needs within ecological limits) to support the transition to a low-carbon society. Thanks to the diffusion of information and communication technologies that provide novel real-time sensing and tracking possibilities, persuasive apps that trigger energy saving have increasingly spread worldwide, welcomed as promising tools to implement highly interactive behaviour change techniques. Rigorous analyses providing evidence on their effects are however still missing. Moreover, the optimal design of their features has still to be identified.

Previous research has found that app-based policy interventions are affected by critical limitations that had already emerged for behavioural interventions in general: a lack of scientific rigour in the empirical evaluations of their effects, poor grounding of app features on behavioural theories, and a tendency to rely on technocratic approaches. Furthermore, they are at risk of only producing short-term, transient effects. Scholars have therefore called for more research on persuasive apps: what are their energy and carbon saving impacts? Do they differ across heterogeneous user groups? Do they last in the long-term? Which features should the apps include, to favour greater engagement by users and therefore better support the energy and climate transition?

Tackling these research questions, I collect evidence on the effectiveness of three app-based interventions targeting energy saving in households, that were designed before my dissertation work and were run in Switzerland between 2016 and 2022. For the first two cases (enCompass and Social Power), under quasi-experimental research designs I perform fixed effects panel data regressions aimed at estimating the average treatment effect on samples of self-selected treated households, both in the short- and in the long-term (up to two full years after the end of the intervention). I also look for possible heterogeneous effects on varying the households' characteristics. For the enCompass case I additionally verify if the effects depend on the level of intensity of app use. For the third case (Social Power Plus), instead, I analyse two questionnaires that were administered to self-selected treatment group households, in order to collect both quantitative and qualitative insights on their evaluation of the app's features. For this case I also perform a qualitative analysis of app-mediated interactions between the involved households, to verify whether a social learning process was activated, thus contributing to shape the evolution of social norms and competences towards more sufficient energy consumption.

Due to the variety of their configurations, the three cases provide me with insights to better understand the actual potential and relevance of persuasive apps in the framework of low-carbon transitions. The enCompass and Social Power app-based interventions were significantly effective in reducing consumption and carbon emissions during the intervention, with average treatment effects respectively of 4.95% and 9.23% (statistical significance at the 0.05 level; effect size, measured through Cohen's d, respectively equal to 0.35 and 0.51). By considering households using electricity solely for non-heating purposes, enCompass even managed to reduce electricity consumption and CO_2 emissions by 14.46 % with respect to the baseline (large effect size, d= 0.91), with a 0.01 significance level. Analysis of the Social Power Plus case suggests that these results are mostly related with use of app features focusing on the individual level (energy consumption feedback and goal setting), which were more appreciated by the users than features acting at the social level (sharing of experiences on the in-app forum).

However, in the long-term (one or two years after the end of the intervention), the statistical significance of the treatment effects disappeared and practical significance estimates show that energy consumption reverted to pre-intervention (if not higher) levels. These results confirm the problem of long-term effectiveness already emerged in literature for other types of behavioural interventions and seem to challenge the body of literature that values social influence techniques as beneficial for a long-lasting change.

The evidence I found tends to dampen enthusiasm about behavioural policies based on persuasive app use: taken in isolation, persuasive apps seem not to be effective in driving long-lasting change for the needed energy and climate transitions —not even when leveraging social influence techniques. Also at the light of the review I performed on previous research and of the related theoretical backgrounds, I suggest that a critical reflection on persuasive apps is needed, and propose to rethink their role in sustainability transition processes.

From the research I performed in this dissertation, it emerges that a promising venue for future research, informed by Social Practice Theories, might be to keep using persuasive apps, but to include them in broader, trans-disciplinary and multi-stakeholder "living lab" processes. The aim of such living labs would be to collectively challenge and re-design current shared cultural and social meanings, material components, and competences around energy-demanding and carbon emitting practices. In the living labs, persuasive apps might serve as ancillary tools providing the community of involved stakeholders with monitoring, public commitment and experience sharing opportunities. Change would however mostly stem from the interactions by different stakeholders, including institutions, within the niche represented by living lab processes, rather than from the app themselves.

Acknowledgements

Countless people supported my effort on this dissertation. My supervisor, Professor Marco Gui, provided me with precious advice since the very beginning of this journey, together with members of my Dissertation Committee, Professor Giuseppe Vittucci Marzetti and Dr. Roman Rudel. I am indebted to all of them.

I am also very grateful to my employer, the University of Applied Sciences and Arts of Southern Switzerland (SUPSI), and particularly to the board of the ISAAC institute that I am serving since 2009, for giving me the possibility to enter an Executive PhD programme and fit it along my other duties and responsibilities.

I am indebted to colleagues of the Zurich University of Applied Sciences (ZHAW), with which, more than six years after the launch of the Social Power project, I am still fruitfully collaborating in Social Power Plus: Devon Wemyss, Evelyn Lobsiger-Kägi, Pascal Kienast and Vicente Carabias-Hutter. And a huge thank you to ISAAC's researcher Roberta Castri, initiator of Social Power, and Pasquale Granato, sparkling developer of the Social Power and Social Power Plus apps.

I am also indebted to Professor Piero Fraternali of Politecnico di Milano, leader of the enCompass project, and to all the enCompass research team.

A special thank you goes to the staff and board of the utility companies AEM, SWW, TBW, EKS and SES. My thanking also goes to the households who participated in the interventions, by interacting with the apps, answering the surveys, and sharing their experiences with us. Without them, this dissertation would not have existed.

I would also like to thank Dr. Tiziano Gerosa, who provided me with valuable suggestions on how to improve my analyses and supported my stubborn search for R packages providing the same features as Stata.

A final huge thank you goes to my family, who always provided me with the time to focus on this dissertation. Thank you Andrea for proofreading and for taking care of all the cooking and chores —though I would have saved more energy than you did, that was a fundamental help. Thank you Federica for allowing me in the same library as yours when we both had to prepare for examinations —I did not disturb you that much, did I? Thank you Davide for sacrificing yourself at playing video-games while I was writing the dissertation —alas, now it's game over! Sunday cycling rides will soon be back. And thank you to my parents Tino and Patrizia and all my relatives for coming to visit us in Orino during Summer week-ends. I was staring at the laptop all the time but did not miss a single word of what you all were saying.

Contents

1 Introduction

> *To slow down, let alone reverse, increasing carbon emissions and temperatures requires the total reorganization of social life, nothing more and nothing less.*
>
> — **John Urry** (Sociologist)

How to favour the transition to a more sustainable society, that can satisfy human needs for everybody while remaining within planetary resource boundaries, has become a topical issue in the current policy debate and research agenda worldwide. Scholars have identified the critical planetary boundary conditions (collective maximum consumption or emission thresholds) that should not be passed to guarantee a safe operating space for humanity (Rockström et al., 2009; Steffen et al., 2015), as well as the requirements to ensure such a space is also just, fair, and equitable across countries and generations. A number of societal models have also been conceived, to ensure that well-being and good and thriving life conditions are offered to every individual, such as for instance, the "Environmental space" model by Spangenberg (2014), the "Prosperity without growth" model by T. Jackson (2009), the "Doughnut economics" model by Raworth (2012), or the "Consumption corridors" societal model by Fuchs et al. (2021).

Inspired by such models, programmes, policies and regulations have been envisioned, and in some cases even introduced, in order to tackle humanity's "grand sustainability challenges" (Markard et al., 2020) and promote a radical shift from current societal and economic organisation of Global North countries. So far, however, they have resulted in limited impact, as it is shown by the still highly ambitious targets for change in both the Global North and South set by the Agenda 2030 and the Sustainable Development Goals, as well as by their current level of achievement (J. Sachs et al., 2021).

In this dissertation, I focus on Sustainable Development goal 13 "Climate action" and on the globally challenging transition to a low-carbon society. This is increasingly demanded by the growing evidence of responsibility of human activities on climate change and of their irreversible impacts on natural and human systems, that are currently pushed beyond their ability to adapt to changed climate conditions. Specifically, by means of quasi experimental research and empirical data collection on the field that I had the opportunity to be involved into in the last few years, I analyse the effectiveness of demand-side policies aimed at fostering the energy transition and at mitigating the climate impact of energy consumption activities in households.

1.1 Demand-side mitigation of climate change

The interdependence of climate, ecosystems and biodiversity, and human societies is now widely acknowledged, as well the related global trends in biodiversity loss, unsustainable consumption of natural resources, land and ecosystem degradation, rapid urbanisation, and social and economic inequalities (Pörtner et al., 2022). Furthermore, climate change has now become an urgent problem not only in the scientific community: strong social movements have recently started to exert political pressure and call for "system changes" (Goodman, 2022).

After the UN Climate Change Conference of the Parties held in Glasgow in December 2021 (COP 26), in particular, climate change has started to be publicly acknowledged and framed as both an environmental and a social justice problem (Sovacool, Newell, et al., 2022), by the spread and affirmation of concepts that were first introduced more than fifteen years ago. For instance, J. T. Roberts and Parks (2006) talked about a "climate of injustice", since the consequences of the climate crisis are expected to be dramatically worse in the South than in the Global North, and to compel millions of people to leave their countries and turn into environmental refugees (Urry, 2009).

Complementary to technological progress, changes in energy consumption patterns by individuals are increasingly recognised as one of the pillars in climate mitigation strategies aimed at meeting the substantial reductions in greenhouse gas (from now on, GHG) emissions required to achieve a "net zero" or "zero emission" society (Sovacool, Newell, et al., 2022). Since the Fifth Assessment Report on the mitigation of climate change (Edenhofer, 2015), scientists of the Intergovernmental Panel on Climate Change (IPCC) have acknowledged the influence on energy use and related carbon emissions by behaviour, lifestyles, and culture, and remarked that behavioural change interventions have high mitigation potential. Scholars such as Creutzig et al. (2018) have in particular called for mitigation of climate change by means of "demand-side solutions", namely strategies encompassing technology choices, consumption, behaviours, lifestyles, products, and service provision, in a socio-technical transition framework[1]. They have suggested to design demand-side solutions according to the "avoid-shift-improve" approach, that was originally developed in the 1990s in the transport domain, and to apply them on energy related demand in the building, food, and manufacturing of products and services domains.

The recent Sixth IPCC Assessment Report on Mitigation of climate change, which has been released in Spring 2022, extensively discusses the potential of demand-side mitigation

[1]Note that a well-developed branch of research, grounded in engineering, grid management and process automation concepts, uses the concept of "demand-side management" to refer to the use of "hardware or software adopted by electricity end users that could potentially provide grid services" (O'Shaughnessy et al., 2022, p. 973). Scholars in this domain adopt a purely technological perspective, which deals with demand-side solutions within building energy technologies (flexible loads, distributed generation, distributed storage), electric mobility, and industry, which is however far from my perspective. My research is in fact grounded in the above-mentioned conceptualisation of "demand-side solutions" provided by Creutzig et al. (2018).

measures, by accounting for changes in infrastructure use, end-use technology adoption and socio-cultural behavioural change (Pörtner et al., 2022). Overall, the report estimates that demand-side measures and new ways of end-use service provision could reduce global greenhouse gases emissions by 40-70% by 2050, compared with baseline scenarios. Particularly, policy measures targeting socio-cultural, behavioural and lifestyle factors could rapidly produce a reduction in global GHG emissions by 5% overall, by mostly acting on the food, transport, and buildings end-use sectors. Furthermore, according to the same report, such mitigation measures would not result in decreased quality of life: to the contrary, they would support the increase in well-being for everyone, through decreased air pollution, healthier and soberer lifestyles.

For instance, Millward-Hopkins et al. (2020) have recently estimated the minimum final energy requirements for decent living standards for the entire world population in 2050. By accounting for a list of basic material needs underpinning human well-being, and by considering a combination of the most efficient available technologies, as well as radically sufficient demand-side consumption patterns, they have found that global energy consumption in 2050 could be up to 60% lower than today (and up to 95% lower than current highest per-capita consumers), while still guaranteeing decent living standards for the world population. Similarly, focusing on the United Kingdom, Barrett et al. (2022) have estimated that a 52% reduction in energy demand would be possible by 2050, without compromising on citizen's quality of life, and rather resulting in more active lifestyles, lower air pollution, and improved work-life balance. Considering that currently 66.6% per cent of the world energy consumption is satisfied by fossil fuels such as gas, oil, and even coal (IEA, 2021b), such a potential in energy consumption corresponds to a huge potential in the decrease of GHG emissions. According to Barrett et al. (2022), rapidly activating such cuts in GHG emissions in the global North would allow to avoid highly ambitious interventions on carbon dioxide removal and especially carbon capture and storage technologies, which are still highly expensive and not sufficiently proven.

1.2 Households in transition to a low-carbon society

In order to support the transition to a low-carbon society and achieve international climate change mitigation goals, concerted action by a number of actors is needed. Governments are called to foster structural change, for instance by ending fossil fuel support, providing low-carbon infrastructures, redefining work policies, or supporting the evolution of social norms and habits, and to collaborate with supply-chain private actors and local communities (Schanes et al., 2019).

Interventions aimed at modifying the material and social context in which individuals and household practices are embedded, as well as at directly targeting such practices, are also needed (Allcott and Mullainathan, 2010; Stern, Janda, et al., 2016). The key role of individuals in the shaping of low-carbon and energy-efficient collective practices has in fact been widely acknowledged (Schot et al., 2016). Also recently, among the five challenges that humanity has to address in order to tackle the overarching "grand sustainability challenges", Markard et al. (2020) have explicitly listed the need for change in consumer practices and routines.

Without aiming to leave all climate climate change mitigation responsibilities to individuals, Ivanova et al. (2020) state that two thirds of global GHG emissions are directly or indirectly linked to household consumption, and argue that a change in consumption practices is *also* needed in order to reach current ambitious net-zero carbon emission goals. Similarly, even though it provides lower figures (households are accounted for being responsible for 25% of the European Union's final energy consumption in 2017), also the European Environment Agency (2020) calls for a reduction in energy consumption by households, in order to achieve the low-carbon transition goals.

1.2.1 Households' carbon footprint and mitigation potential

Households, their dynamics and their material settings are thus an important target for government policies. Indeed, the energy and climate transition potential of household actions have been explored since at least two decades. Early attempts by Druckman and T. Jackson (2009), for instance, aimed at quantifying the carbon footprint[2] of average households in the United Kingdom (UK). Their estimates, that account for energy consumption for space heating and lighting, consumption of goods and services, as well as personal transportation (vehicle and aviation), indicate that in 2004 about 55% of carbon footprint of UK households was due to GHG emissions embedded in consumed goods and services, and that only 30% was due to direct energy consumption. Druckman and T. Jackson (2010) also estimated that overall GHG emissions could decrease up to 37% in the UK, if households were to abandon current over-consumption scenario and adopt a less materialistic "Reduced Consumption Scenario", which guarantees subsistence (food, heating, shelter) and also offers everybody the means to participate effectively in society. Such a scenario would also increase well-being, by reducing stress and anxiety, and produce a more egalitarian society, thus ensuring "a good life" living standard for everybody.

Specific estimates of the carbon footprint of households have first been performed for the UK by Gardner and Stern (2008): at the time of the analysis, households were responsible for 38% of the country's carbon emissions. The authors of the same study estimated that households could reduce their energy consumption, and hence their carbon emissions, by almost 30%, without major economic sacrifice or decrease in well-being.

More recently, by considering the whole supply chain of product and services and their life-cycle emissions, and computing the overall climate change mitigation impact

[2] Following Wiedmann and Minx (2008), I define the carbon footprint as "a measure of the exclusive total amount of carbon dioxide emissions that is directly and indirectly caused by an activity or is accumulated over the life stages of a product". This includes activities of individuals, populations, governments, companies, organisations, etc., as well as production processes of goods and services. The carbon footprint accounts for both direct (on-site, internal) and indirect emissions (off-site, external, embodied, upstream, downstream).

associated with ninety types of changes in consumer behaviour, Moran et al. (2020) have estimated a 25% potential for reduction of the overall European carbon footprint. They accounted for changes in consumption patterns, substitution with goods and services with lower carbon footprint, and reduction in the overall amount of consumption. Their analysis has shown that the highest mitigation potential is associated with measures dealing with food, transport, and building efficiency.

Also Wynes and Nicholas (2017) provided a valuable insight on the climate impacts of changing individual patterns and practices, by identifying the "top" list of high impact changes in practices to be pursued for climate mitigation purposes: having one fewer child, living car-free, avoiding airplane travel, and eating a plant-based diet. Similarly, Ivanova et al. (2020) identified the "top 10" changes in household practices from a climate mitigation perspective, via a meta analysis of previous literature estimating the CO_2 equivalent reduction potential. According to them, the "top 10" changes are related with shifting diet (becoming vegetarian or vegan), transport mode (from use of active mobility, public transport or electric vehicles to car-free living in urban areas), and energy source (moving to renewable-based heating and electricity), reducing overall travel demand (less air travel and exploitation of digitalisation alternatives), and performing home energy retrofits.

Similar mitigation options are identified by the latest IPPC Assessment Report on Mitigation of climate change (Pörtner et al., 2022), that stresses the need for urgent implementation of car-free living, plant-based diets, use of low-carbon sources of electricity and heating at home, and local holiday-making. The report however also remarks that simply adjusting the set-point temperature in Winter and Summer could result in CO_2 emission savings between 5 and 25% compared with average households' emissions in the Western countries.

Indeed, even though the households' carbon footprint has been shown to be higher in other domains than residential buildings, recent research has confirmed that energy consumption at home is a relevant domain to be investigated. A recent meta-analysis by Khanna et al. (2021) of 133 behavioural interventions targeting a reduction in energy consumption and carbon emissions in residential buildings has in fact estimated a worldwide carbon emission reduction potential of 0.35 Giga ton CO_2 per year, which is about 6.25% of carbon emissions of the whole residential building sector estimated in 2018 by the International Energy Agency (IEA). While the authors acknowledge the limited effect of such behavioural interventions with respect to national and international net-zero emission policy targets, they also suggest to combine behavioural interventions with structural interventions, such as the replacement or upgrade of heating and non-heating equipment and appliances. Therefore, they urge the scientific community to keep providing rigorous assessments of the evidence of the effects by behavioural interventions in residential buildings.

1.2.2 Households' willingness to change

All the above estimates need to be handled with caution, since the actual impacts are highly dependent on the context, which is made of infrastructural, technological, political, and cultural factors, including the dominant social norms in a specific time and place. Also, the estimated GHG saving potentials are strictly dependent on the initial starting practices and patterns of energy use of the households. And, finally, all those estimates refer to the *theoretical technical carbon saving potentials* of household action. In real life, practical constraints often prevent achievement of such potentials. Furthermore, indirect rebound effects can happen. This is why, by adopting a broader system level perspective, the overall emission reduction potentials may turn out to be lower than above anticipated (Druckman, Chitnis, et al., 2011). However, despite these uncertainties in the actual climate change mitigation potential by changes in household behaviours, there is a wide agreement among scholars that both high and low impact changes in behaviour need to be urgently put into practice.

For instance, Wynes and Nicholas (2017) call for forward-looking policy goals and measures aimed at implementing both low and high climate change mitigation actions. They remark that the political unpopularity of high mitigation impact actions does not justify a focus on moderate or low-impact actions at the expense of high-impact ones: forward-looking policy goals and measures are needed in order to guarantee the implementation of both low and high climate change mitigation impact. Further, even though they are not sufficient to achieve a deep decarbonisation of our society, low impact changes in behaviour are still necessary as well.

However, in spite of the relevant mitigation potentials, research has shown that households might not be spontaneously willing to put such potential changes into practice. Early work in the energy domain, for instance performed by Steg (2008) or Stern (2000), has remarked that many people attribute low priority to saving energy, giving instead higher priority to status, comfort, or effort. Therefore, they are likely to reduce or change their consumption patterns only when this entails limited costs in terms of money, effort or convenience. This is for instance the case for recycling, while a reduction in car use is much less likely to occur. A recent empirical study by Dubois et al. (2019) on German and Swedish households has provided further evidence for this phenomenon: without external intervention or conditioning, households are mostly willing to only implement lifestyle changes that are characterised by lower climate change mitigation impact, such as upgrading home appliances to highly energy efficient ones, or switching to renewable electricity. High impact changes, such as giving up the car or substantially reducing the number of long-distance air travels, are instead less likely to be voluntarily adopted by households, since they would require way too relevant changes in dominant behaviours.

Also the latest Assessment Report by the IPCC on mitigation of climate change (Pörtner et al., 2022) remarks that willingness to adopt has been observed for certain measures (full load to laundry appliances, use lids while cooking, turn lights off, defer electricity usage and heating, ventilation and air conditioning systems, reduce set-point temperature by 1 °C) but not for others (turn off appliances on standby, use clothes for longer period, avoid leaving the TV on while doing other things, defer use of oven, ironing or heating systems, adjust set-point temperature by 3°C, move to a low energy house or smaller apartment). To guarantee such changes in behaviour are soon put into practice, therefore, external triggers are needed, through dedicated policy interventions.

1.3 App-based persuasive policy interventions

With the rise of the information age and diffusion of the network society (Castells, 1996), digital innovation has enabled novel synergies in the potential of people and technology to tackle societal challenges in a wide range of fields (Sareen and Haarstad, 2021), such as provision of health-care services, improvement of wellbeing of the elderly population, support to migrants, delivery of teaching and learning environments, and also tackling the climate crisis (Stokes et al., 2017). Digital technologies in fact are believed to help mobilise and engage large communities, facilitate bidirectional communication, and support information sharing and feedback at the individual and collective level. In particular, they help capture, analyse and communicate novel data, which can be leveraged to support real-time decision-making. For these reasons, digital technologies have been increasingly used within behavioural policy interventions aimed at supporting change.

In the energy sector, the roll out of the smart meter technology a decade ago was accompanied by large expectations for this digital technology to impact demand side management approaches within the household domain (Akenji et al., 2021). In the early two-thousands, smart meters were in fact expected to support the transition to more efficient, sufficient, and flexible energy consumption practices, well-suited to the integration of renewable energies into the energy system and the transition to a low-carbon society. In particular, they were welcomed as promising devices to be integrated in behaviour change policy measures, due to their capability to provide data for consumption feedback through in-home displays that can inform, motivate, and ultimately persuade change (Darby et al., 2006; Fischer, 2008).

More recently, data feedback was incorporated into another digital innovation: the smartphone. An individual's smartphone can be the data collection technology itself through a smartphone app, either by automatically sensing an activity from the phone or another device (e.g., recording the length and temperature of showering) or by allowing for manual data tracking by the user (e.g., tracking consumption activities within and outside the house). Subsequently, the related app can provide immediate feedback and bridge the gap between data and its user, through various forms of engaging feedback: informative, personalised, real-time, entertaining, gameful, among other aspects. Furthermore, the app can incorporate a rich set of features, that act at both the individual and the group level in order to nudge (Thaler and Sunstein, 2009) and persuade (Fogg, 2003) change.

Thanks to their potential to reach a wide audience, smartphone apps aimed at persuading behaviour change have increasingly been adopted in real-world activities, to support health (Chatterjee and A. Price, 2009; Kientz et al., 2010; Chow et al., 2017; Orji and Moffatt, 2018) and sustainability transitions (Froehlich et al., 2010; Kimura and Nakajima, 2011; Agnisarman et al., 2018; Spaiser et al., 2019; Wee et al., 2021; Adaji and Adisa, 2022). As documented for instance in the review by Fraternali et al., 2019, they flourished in both the private sector and research, within processes aimed at addressing consumption practices by a variety of target groups (households, school communities, office employees), and in a variety of domains (energy consumption in buildings, transportation, water or food consumption, etc.). Most apps rely on gameful approaches, by either being shaped as serious games (Wood et al., 2014) or by including gamified features (Morganti et al., 2017; Beck et al., 2019).

Overall, these apps are expected to support a reduction in energy consumptions —and hence in CO_2 emissions— while at the same time maintaining individual wellbeing —if not improving it. For these reasons, many energy providing companies are increasingly launching their app and webportals for their customers, with the aim of both increasing their awareness on their energy consumption and supporting them in the reduction of their consumptions.

Previous studies have however found key limitations that may affect the behaviour change potential of such persuasive apps, as well as the analyses that were performed in order to assess their effectiveness in field interventions. Similarly to many previous behaviour change interventions, the apps' persuasive features can in fact be questioned for being poorly grounded in behaviour change theories (Michie and Prestwich, 2010; Michie, Carey, et al., 2018; Beck et al., 2019; Nielsen, Cologna, et al., 2021) or for relying too heavily on quantitative and technical consumption feedback (Strengers, 2014). The related field interventions have also been questioned for the observed critical decrease of the engagement of app users over time (Perski et al., 2017; Löschel et al., 2020), as well as for the opt-in framework unavoidably characterising their use. The latter in fact may not manage to raise the interest by non-intrinsically motivated individuals, which instead might be the key target group to mobilize (Hartman, 1988; Tiefenbeck et al., 2019; Cellina, Vittucci-Marzetti, et al., 2021). Furthermore, and more critically, the related field interventions can be questioned for the lack of scientific rigour in the evaluation of impacts and for the lack of assessment of the effects in the long-term (Delmas et al., 2013; Vine et al., 2014; Nielsen, Cologna, et al., 2021; Frederiks, Stenner, Hobman, and Fischle, 2016).

Rigorously analysing the impacts of such behaviour change apps, both in the short and the long term, is therefore a still open research issue, with relevant practical policy implications, that is worth being taken.

1.4 Research goals and questions

Against this background, in this dissertation I aim at contributing to the literature dealing with the households' role within the energy and climate transition, and focus on households' energy consumption practices that are performed at home, within the broader socio-technical contextual factors they are part of.

I take advantage of three policy interventions that were performed in Switzerland between 2016 and 2022, for each of which a specific smartphone app targeting a reduction in electricity consumption in households was designed and field tested with voluntary households. The apps were designed before and outside the work I performed for this dissertation. The main goal of the dissertation is thus to assess the effectiveness of such app-based policy interventions aimed at reducing home-related energy consumption in households. Considering the above limitations, my research aims at understanding whether they are really worth the effort; and, if so, which features can produce larger energy saving —and hence CO_2 saving— effects. The evidence I collected from the analysis of the energy and carbon saving effects in those three cases resulted in recommendations for future research, with the final aim of tangibly supporting the transition to a low-carbon society and delivering a concrete and long-lasting impact.

Based on the three case studies, I specifically tackle the following overarching research questions (RQs):

- RQ1: Can the use of behaviour change apps produce a reduction in residential energy consumption and related CO_2 emissions of households?
- RQ2: If app use is found to produce a reduction in energy consumption and CO_2 emissions during or immediately after the intervention, is such a reduction also observed long after the end of the intervention?
- RQ3: Are the effects on energy consumption and CO_2 emissions constant, on varying observable characteristics of households? Or does heterogeneity in observed characteristics of households lead to heterogeneous effects as well?
- RQ4: Which app features can foster higher user engagement, thus providing greater support to the reduction of energy consumption and CO_2 emissions?

1.5 Three app-based interventions

The three policy interventions I analyse were respectively developed within the enCompass project, funded by the European Union under the Horizon 2020 research programme, the Social Power project funded by the Swiss Gebert Rüf Foundation, and the Social Power Plus project funded by the Swiss Federal Office of Energy. I selected them partly since I expect the variety of their configurations can provide useful insights to better understand the actual potential and relevance of persuasive apps in the framework of the energy and low-carbon transitions. Undeniably, however, their selection was also performed for convenience reasons: I was part of the projects' research teams —though in different ways and with different roles I indicated in the related chapters— and have therefore access to relevant data generated during interventions and to the most relevant intervention-related materials, that I will exploit to enrich the analysis.

The country where the interventions are located (Switzerland), however, requires specific considerations. The carbon footprint of the Swiss electricity consumption is quite low, compared to other European countries that are still highly dependent on fossil fuels: about 60% of the average Swiss electricity mix is in fact made by hydroelectric power, and 30 % is made of nuclear power (Bundesamt für Energie, 2021), which are both carbon free, at least in terms of direct combustion emissions. By accounting for exchanges on the electricity market, the carbon intensity of electricity consumption (average amount of CO_2 emissions per consumed kWh) was in fact estimated at 128 g for Switzerland (Krebs and Frischknecht, 2021), which is definitely low compared with other "WEIRD" (Western, Educated, Industrial, Rich and Democratic" Henrich et al., 2010) countries of the Global North, such as for instance Germany, where the average carbon intensity of electricity consumption was estimated at 366 g of CO_2 emissions per kWh (Icha and Kuhs, 2021). In Germany in fact fossil fuel sources are still responsible for more than 40% of the electricity mix, despite the relevant contribution by new renewables, which overall account for more than 30% (IEA, 2021a). If the three Swiss interventions I analyse will prove to be effective in decreasing households' electricity consumption, and thus the related CO_2 emissions, they would be even more relevant and useful in other countries where the carbon intensity of electricity consumption is higher. The treatment effect in terms of reduction of absolute CO_2 emissions I will find in the three case studies —if any— has therefore to be regarded as a minimum treatment effect. The percentage effect on the reduction of electricity consumption being equal, the absolute effect on the reduction of CO_2 emissions would in fact be higher in other WEIRD Global North countries.

The three interventions share two key characteristics:

- they aim at reducing electricity (and gas, in one case) consumption in households, and thus the related CO_2 emissions, by exploiting the provision of eco-feedback and other motivational affordances through smartphone apps directly connected with smart meters measuring the household's consumption;
- their effectiveness has been tested in quasi-experimental settings based on a before-after design (panel data modelling approach), by recruiting voluntary households that were engaged via public communication campaigns hosted by the utility company providing them with electricity (and gas).

Due to the small size of the recruited samples of households, in all cases no random assignment of the treatment has been performed and all the recruited voluntary households have been attributed to treatment groups. In order to assess the average effect on households that received the treatment, comparable control groups of households have been identified on a later stage through random stratification based on a few observed characteristics. Thus, the opt-in, self-selection framework that characterises the three interventions will not allow me to draw conclusions on the short and long-term effect of behaviour change apps on the average population: in none of the cases I analyse, in fact, the sample of households involved in field activities is representative of the population. The three quasi-experiments assess the effect of the treatment on households with intrinsic motivations to try app use (possibly for pro-environmental attitudes, expectations of monetary saving reasons, or maybe interest in new technologies and digital devices), against average households. Considering the fast evolving societal context, in which the effects of climate change are increasingly visible to the population on a daily basis even at our latitudes (droughts, extreme weather events, avalanches, etc.), and the high monetary cost of energy supply induced by the Russian war in Ukraine is pushing novel and larger population groups towards energy poverty conditions, I expect that the segment of intrinsically motivated population will become larger and larger in the near future, thus making this analysis even more relevant and insightful.

Besides the above common key characteristics, each policy intervention relies on different theoretical backgrounds, exploits different persuasive motivational affordances, focuses on different types of domestic uses of electricity, and operationalises the provision of feedback and other persuasive affordances in different ways. To allow for comparability between the three apps, I will outline their features by referring to a common set of evaluation criteria, which allows to frame them by both the theoretical background and the persuasive principles and techniques they exploit. In short, a key difference between the three cases is that, apart for a comparison with other households in the "leaderboard" section, all the features offered by the enCompass app focus on the single household and do not leverage any type of social interaction among the community of its users. The Social Power and Social Power Plus apps, instead, largely exploit persuasive motivational affordances that draw on the social dimension and on the creation of an interactive and lively virtual community between their users.

Even though the analysis of just three cases will not allow me to draw conclusions of general value about the short and long-term effect of any type of behaviour change apps targeting domestic energy consumption, the broad variety of characteristics of such three interventions allowed me to explore different aspects of app-based behavioural interventions and, combined with the lessons learnt from other similar interventions that are available in the scientific literature, to collect valuable insights to inform policy-making.

1.5.1 Case one: enCompass

The enCompass intervention took place between 2018 and 2019 in three European countries (Germany, Greece, and Switzerland), within the enCompass project funded by the European Union Horizon 2020 research programme (`https://www.encompass-project.eu`, last accessed on January, 27 2023). The project developed an app-based persuasive platform aimed at reducing energy consumption in households, public buildings, and schools, and field-tested it in the three regions. For each type of target users (household members, pupils in schools, and employees in public offices), a different version of the app was developed. In all cases, the app accounted for any type of energy consumption associated with the buildings where the interventions was taking place, provided that smart meters were available to automatically access energy consumption data, and provided consumption (eco-)feedback, offered goal setting opportunities and provided customised tips and recommendations to save energy, under a global gamified approach (the use of game elements in non-game contexts, as it was first defined by Deterding et al., 2011. Here I focus on the app version targeting households, and specifically on the field intervention that was performed in the Italian-speaking part of Switzerland, in the small municipality of Contone. At the time of the project, I was part of the consortium team that was responsible for management of the field intervention and interaction with project participants living in Contone.

In the municipality of Contone, only electricity smart meters were available, under direct management by the local utility company. The intervention however Included households equipped with heat pumps and/or hot water boilers, in order to assess the effectiveness of the enCompass app in reducing electricity consumption for both heating (rooms and/or hot water) and non-heating purposes (use of appliances and lighting). Besides smart meter data, also a “before” and “after” set of questionnaires was developed, which allowed to collect additional information about energy behaviour and practices at home.

The enCompass intervention was devised under a quasi-experimental approach, with a self-selected treatment group of households (n = 75 at the start of the project, recruited via a local communication campaign) that was treated with the enCompass app for a full year, from June 2018 to May 2019, and asked to answer the questionnaires. For the control group, during the enCompass project a sample of households living in nearby villages was considered (n = 25): sample members were the respondents to the questionnaires aimed at integrating consumption data collected via the smart meters —again, self-selected. For both treatment and control groups, two full years of electricity consumption data were collected (June 2017 - May 2018 and June 2018 - May 2019), to be respectively used as baseline consumption and consumption during the treatment period, in order to assess the average treatment effect on the treated (ATT) via a Difference-in-Differences (DiD) estimator. However, as a result of the poor sample sizes, especially the one by the control group, no statistically significant treatment effects were obtained, as reported in project deliverable 7.4 “Final overall validation and impact report’ (at the time of writing still under embargo to allow for scientific publications) and in Koroleva et al. (2019).

To deal with my research questions, I was however interested in clearly identifying the impacts of the enCompass app. For this reason, I revisited the enCompass case’, with the aim of performing an in-depth exploration of the intervention’s effects. For this purpose, I contacted the local utility company and managed to obtain electricity consumption data for a larger sample of households, namely all the 230 households living in the municipality of Contone, to be used as the control group. Further, I obtained their electricity consumptions also for the periods June 2019 - May 2020 and June 2020 - May 2021, for all households living in Contone, thus including those of the treatment and control group.

With the new data-set, larger in terms of both sample size and number of available years of electricity consumption, I performed a panel regression analysis aimed at identifying the average treatment effect of the enCompass intervention on the sample of treated households, both in the short term (during the intervention itself), and in the longer term (respectively, one and two years after the end of the intervention). Doing so, I could address the research questions RQ1, RQ2, and RQ3 presented in Section 1.4. Available information on observable characteristics of the households of the two samples also allowed me to estimate the average effect of the enCompass treatment on the treated households (ATT) (this time, obtaining statistically significant results) and to perform an analysis of heterogeneity of the effects among the different sub-group of households. Finally, I also searched for possible differences in the average treatment effect due to the level of app use, by performing a cluster analysis on the treated households based on their level of app use, automatically provided by the app's internal analytics system.

1.5.2 Case two: Social Power

The Social Power intervention instead place in Spring 2016 in two cities, respectively located in the Italian (municipality of Massagno) and German (municipality of Winterthur) speaking-part of Switzerland. At the time of the project, I was part of the consortium team that was responsible for the design and assessment of the field intervention, and for interaction with project participants living in Massagno. In this case, for a period of three months n= 54 self-selected households in each city took part in a field intervention aimed at reducing their electricity consumption via the Social Power app, again automatically connected to smart meters to provide customised and detailed consumption feedback. In this case, only consumption for electric appliances and lighting was considered, and households with heat pumps or boilers were not allowed to join the field intervention.

Social Power adopted a gamified approach as well, and exploited a number of persuasive features supporting electricity saving. Central to its approach was the idea of leveraging social norms and the collective dimension, through the creation of teams of households, invited to collectively save electricity over a period of three months. Indeed, two partially different versions of the Social Power app were developed and tested on the field, which differed for the gamified structure they adopted: a collaborative gamified structure, inviting household teams to reach an overall electricity-saving target of 10%, and a competitive gamified structure, putting two teams of households into direct competition and inviting each team member to save as much electricity as possible together with their fellow team-members, in order to beat the other team. In both cases, percentage electricity savings were computed with respect to previous electricity savings collected for the same households over a comparable period (baseline).

The two gamified structures differed in the type of social feedback offered to the user, which was designed in order to respectively enhance the underlying collaborative or competitive principles. No other differences, instead, concerned the other app features. To prompt households to engage in energy-saving activities, Social Power in fact offered a number of additional features, including individual feedback, challenges, and tips, providing households with hands-on learning opportunities to support self-efficacy in reducing their electricity consumption.

Besides the automatic collection of consumption data via the smart meters, the Social Power project developed a "before" and "after" questionnaire targeting the treated households and arranged individual interviews with a sub-sample of such households, with the aim of collecting additional information to asses the effectiveness of the intervention. A quasi-experimental approach was used also in Social Power to estimate the average treatment effect on households treated by the Social Power app. For this purpose, control groups of households were identified in each city and three three-month electricity monitoring periods were considered: before the intervention (baseline), during the intervention, and one year after the end of the intervention (followup). This research design, again based on a Difference-in-Differences estimator, allowed to estimate the average treatment effect on the treated by the Social Power app, also differentiating between the collaborative and competitive gamified structures, both in the short-term (Wemyss, Castri, et al., 2018), namely during the intervention itself, and in the long-term (Wemyss, Cellina, Lobsiger-Kägi, et al., 2019), namely one year after the end of the intervention.

The previous analyses, however, did not explore the heterogeneity of the effects among the treated households, which could instead offer interesting insights, since information is available about the type of household and home, in addition to information on the location (city) where the intervention took place. For this dissertation, therefore, I revisited the Social Power case and, by exploiting exactly the same data collected during the Social Power project, I performed novel analyses on the electricity consumption data available for the treatment and control groups in the two regions, by adopting a panel data regression modelling approach and by looking at the heterogeneity of treatment effects, both in the short- and in the long-term, on sub-groups of households identified based on available characteristics. Doing so, I gained deeper knowledge and a richer understanding on the effects of the Social Power app, which provided me with novel elements to tackle RQ1-RQ3 presented in Section 1.4.

1.5.3 Case three: Social Power Plus

The third case study refers instead to the ongoing Social Power Plus project, which I am currently leading (`http://www.socialpower.ch/`, last accessed on January, 27 2023). The project started in Fall 2020 and will be concluded in Summer 2023. It involves about 200 households located in the three German-speaking Swiss regions of Schaffhausen, Wil and Winterthur, again selected via a voluntary recruitment process coordinated by the local utility companies, and deals with both heating and non-heating energy consumptions by the households. Heating consumptions are monitored via gas smart meters in the region of Will and by electricity smart meters in the regions of Winterthur and Schaffhausen, where only households equipped with heat pumps were eligible to join the project.

The Social Power Plus app can be regarded as a “followup” of the “Social Power” project, since it largely leverages motivational affordances on the collective dimension, and has the aim of creating a community of households collectively engaged for the energy transition. Namely, in this app the “social” dimension is dominant. The Social Power Plus app was in fact first of all co-designed with a group of interested citizens, who were engaged in a set of participatory workshops in a living lab framework (Pallot et al., 2010). Doing so, it was expected to be more appealing to its future users, citizens, and thus to tackle the limitations identified by Strengers (2014), about these apps having been designed for “Resource Men” interested in numbers, figures, and technological optimisation, instead of dealing with the relevant factors and practices that drive energy consumption in households. Furthermore, it includes social comparison and competition features, aimed at strengthening the feeling or urgency for a collective change, and especially it offers peer-to-peer social learning opportunities, which are delivered by means of an in-app forum (the Social Power Plus “pinboard”), on which app users are invited to share their energy saving tips and experiences —positive or negative ones— and monthly online meetings open to all participants to openly discuss about the evolution of the intervention and about changes in households’ energy consuming practices, including the difficulties and opportunities they entail. The app also offers biweekly individual challenges aimed at re-crafting eight specific energy related practices (heating, showering, washing, cleaning, cooking, dishwashing, studying and working, recreation), also supporting those processes via (non-customised) tips. Finally, the app also offers individual feedback features informing on the daily and weekly evolution of energy consumption and on the energy saving compared to a consumption baseline, and provides frequent weekly notifications to prompt energy saving.

Overall, the main aim of Social Power Plus is to favour the creation of novel social norms, social relations, and competences around household energy consumption through peer-to-peer learning opportunities, thus triggering households’ collective engagement in reconfiguring their energy consumption. The households that voluntary decided to join the intervention were treated with the Social Power Plus app for three months, from January to April 2022. After that period, the Social Power Plus app remained available to them until the end of year 2022, for nine additional months; however, no new challenges were released, no competitions were activated, and definitely less frequent notifications were sent to app users. The project’s goal is to assess whether an effect can be found in the short term, namely during the “high interaction intensity” three-month intervention period, and whether it persists over time, namely at the end of the “low interaction intensity” nine-month intervention period. For this purpose, also in this case comparable control groups are identified with the help of local utility companies. The average treatment effect on the treated will again be estimated based on a comparison between the electricity consumption during a baseline period and the two intervention periods, according to a Difference-in-Differences estimator. Additional insights on the outcome of the intervention, together with an evaluation of the app's features, are collected via a "before" and "after" questionnaire delivered to all households of the treatment groups.

The time-schedule of this dissertation does not allow me to quantitatively investigate the long-term effect of the Social Power app, since the "low-intensity" period closes in the same months in which the dissertation is due. Furthermore, due to different reading periods of the baseline and control group energy consumptions, analysis of the short-term effect during the "high-interaction intensity" period, turned out to be only possible for subsets of households of the regions of Wil and Winterthur, and precluded for the region of Schaffhausen, where only annual electricity consumption data over a full calendar year (January to December) is available for both the baseline and the intervention periods for the control groups. For these reasons, in the dissertation I do not perform quantitative estimates of the treatment effect, and instead focus on the qualitative and quantitative analysis of the two questionnaires and on the evaluation of the app's user experience, namely the evaluation of its features by its users, which allows me to address RQ4 presented in Section 1.4. I also perform a qualitative analysis of the materials that have been posted in the "pinboard" section of the app, to verify if a social learning process has actually occurred and, ultimately, to assess if the ideas of challenging specific energy consuming practices and enabling novel interaction possibilities within the community of app users can actually contribute to shaping the evolution of social norms and competences around energy consumption in households, and thus ultimately supporting the needed transition.

1.6 Thesis structure

The dissertation is organised in three main parts: in the first one (Chapters 2 and 3), I introduce the overarching concepts I refer to throughout the document, as well as theories and previous applications of behaviour change interventions available in the scientific literature. In the second part (Chapters 5, 4, 6) I present each of the three case studies, by introducing their goals, outlining the methodologies of intervention and of analysis, presenting their results and discussing them, by also cross-comparing them with the available literature. The three cases are not directly interconnected with each other, apart for a few connections between Social Power and Social Power Plus, as also suggested by their names. This is why the order in which I chose to present them is not particularly relevant. In the third part (Chapters 7 and 8) I provide an overall discussion based on the learnings from the three cases, highlight the contribution and limitations of my research, and conclude by suggesting recommendations for applied future research aiming at tangibly supporting the energy and climate transitions.

2 Conceptual framework

> *Buy, buy, says the sign in the shop window.*
> *Why, why, says the junk in the yard.*
>
> — **Paul McCartney**
> (Song-writer)

In this chapter I present the key concepts that I refer to throughout the dissertation. In order to position and clarify the foundations of my work, I start from basic concepts about *socio-technical system*, *transitions*, and *sustainability*. I then present the concepts of energy efficiency and energy sufficiency, that are at the basis of the three app-based interventions I analyse as my case studies. I then focus on the concepts of *behaviour* and *behaviour change*, which the whole dissertation deals with, by also introducing the main theories that were developed to inform behaviour change interventions and explain their outcomes. Specifically, I introduce the theories underlying the majority of energy saving apps: the *Theory of Planned Behaviour*, the *Value-Belief-Norm Theory*, the *Transtheoretical Model of behaviour change*, *Nudge* and behavioural economics theories, the *Self-Determination Theory*, and *Captology*. Then I present and discuss the concept of *gamification*, which is exploited by most apps. Finally, I introduce the concept of *social practice* and the related body of theories that have been proposed in the last couple of decades to achieve sustainability transitions, in antithesis to behaviour change approaches. I conclude by providing a conceptualisation of *households*, the main unit of analysis I refer to in this work: I show the nuances and different interpretations that have emerged in recent literature for such an apparently unambiguous concept and clarify the specific definition I refer to in this dissertation.

2.1 Socio-technical systems and their transitions

According to Urry (2010), climate change due to human activities can be regarded as one of the contradictions stemming from contemporary excessive consumption capitalism, which, in the quest for increased growth, generated high carbon, path-dependent and locked-in economic and social institutions. The concepts of path-dependency and economic and social lock-in are central to the work by Geels, Sovacool, et al. (2017), who argue that deep decarbonisation requires the transition of entire socio-technical systems, conceptualised as the “interlinked mix of technologies, infrastructures, organisations, markets, regulations, and user practices that together deliver societal functions” (p. 1242). Such a transition can be enacted by contrasting the path-dependencies and alliances that have emerged and consolidated throughout the co-existence and co-evolution of all such elements, and by re-aligning them in a different direction. This conceptualisation is grounded in the overarching theoretical framework offered by the Multi-Level Perspective (Geels, 2004; Geels and Schot, 2007), according to which socio-technical transitions can occur when three mutually reinforcing processes occur: the emergence of innovations in protected *niche* spaces, the weakening of existing dominant configurations in *regime* conditions, and the emergence of exogenous pressures among the *landscape* factors. When all niches, regimes and landscapes align towards novel directions, they can create windows of opportunities for socio-technical transitions to emerge and settle, thus replacing previous system configurations. The process of learning, co-evolution and adaptation at multiple levels results in multiple innovations, such as "investment in new infrastructures, establishment of new markets, development of social preferences, and adjustment of user practices" (Geels, Sovacool, et al., 2017, p. 1242).

Low-carbon transitions are therefore not just driven by technological innovation, carbon pricing, or financial incentive policies: they also require widespread social change by both citizens and companies, that are requested to modify their behaviours, beliefs, cultural conventions, skills, and purchase decisions, as well as to create new business models and systems through which goods and services are produced and distributed. Ultimately, if the aim is to implement a sustainable consumption paradigm (Akenji, 2014), a shift is needed in whole socio-technical systems of provision. Socio-technical perspectives emphasise that such a shift occurs through interactions between technology and society, that co-shape each other (Raven, Jolivet, et al., 2009). Particularly, Geels, Sovacool, et al. (2017) argue that positive discourses are needed about the economic, social, and cultural benefits of low carbon innovations, and that these can be achieved by means of experimentation, bottom-up learning, and stakeholder engagement and empowerment processes. On the same page, Ivanova et al. (2020) call for challenging broad patterns of consumption and societal dynamics, by means of critical appraisal of infrastructural, institutional, and behavioural lock-ins. For this purpose, Hajer et al. (2015) call for the mobilisation of novel agents of change, such as businesses, cities and civil society.

The transition to a low-carbon society thus requires transformative, radical changes and innovations in the systems of provision underpinning all energy consumption domains, such as industry, commerce, transport, and buildings, as well as in the demand for energy and in the related behaviours and practices enacted by social actors.

2.2 Weak and strong sustainability

Scholars such as Akenji (2014), Lorek and Spangenberg (2014), and Spangenberg (2014) have called Western societies for the radical abandon of the economic growth paradigm and for a shift to a lower consumption paradigm: only if people consume less, less raw materials are used and less waste, including carbon emissions, is produced. Their call fits in a "strong sustainability" approach, which is based on social justice, lower consumption and sufficiency principles (conceptualised as the reduction in resource use in absolute terms), and it regards individual and collective well-being as disconnected from economic growth. This differs from the “weak sustainability" approach, which is instead embedded in the economic growth paradigm and relies on economic efficiency driven by technological and market-based solutions (Lorek and Fuchs, 2013). According to the latter authors, only a strong sustainability and sufficiency approach can support achievement of international sustainability goals.

Similarly, Spangenberg (2014) have conceptualised strong sustainability as a shift towards “*better but less* for affluent groups” and “*enough and better* for those still living in poverty” (p. 62). This would allow to overcome socially unsustainable under-consumption (namely, meet both physiological and psychic needs, that guarantee a dignified life that also offers opportunities for participation in social life), while at the same time phasing-out environmentally unsustainable over-consumption. Such a shift requires both consumer involvement and strong distributive justice policies and regulations aimed at re-balancing income and wealth. The authors posit that strong sustainability needs a coherent and overarching reorganisation of the social, economic, and institutional fabric of societies and economies, tackling production, distribution, and consumption patterns. Such a reorganisation should neither be fully delegated to the consumers’ responsibility nor their involvement and role should be neglected (Lorek and Spangenberg, 2019).

Likewise, O’Rourke and Lollo (2015) have dealt with the concept of “strong consumption sustainability”, calling for a broad shift involving individual behaviour as well as collective norms, values, and systems of provision enabled by structures such as markets, institutions, policies, and regulations. Arguing that so far efficiency approaches have not produced evidence of reduction in absolute climate and environmental impacts, and that they still imply problems of equity and wealth distribution among the population, such scholars call for a system level structural change, capable to “decouple human well-being from the destruction of nature” (O’Rourke and Lollo, 2015, p. 240), by means of broad behavioural, cultural, institutional, and structural system change, within a systems framework for learning, iteration, and scaling. Since human well-being is determined by more than just income and economic purchasing power (and could for instance be measured in terms of life expectancy, literacy rates, and subjective measures of life satisfaction, as proposed by J. T. Roberts, Steinberger, et al., 2020), implementation of such a socio-technical change is expected to bring about a double dividend of environmental and social benefits (T. Jackson, 2009).

2.3 Energy efficiency

Having provided a broad overview of principles behind sustainability transitions, I now focus on key concepts that specifically refer to the energy and climate transition. In the energy domain, the promotion of energy efficiency (the reduction of the ratio between the primary energy input and the energy service outputs) is usually regarded as the way forward to promote a low-carbon society, since it allows to reduce the amount of consumed energy —and thus carbon emissions— required for the provision of a given energy service. However, the demand for energy services has kept to increase worldwide and the rate of world primary energy consumption has been steadily growing since 1850 (Sorrell, 2015; IEA, 2022) —and greenhouse gases emissions alike (Shukla et al., 2022). This is due to the increase in types of energy services, such as for instance motive power, lighting, or thermal comfort, that in developed countries are considered as "normal" or "necessary", and to the increase in the amount of the world population accessing them. Moreover, the important energy efficiency gains that have been achieved for specific goods or services thanks to technological progress, directly result in lower purchasing costs for such goods or services, which in turn lead to an increase in their demand and drive further consumption by a larger number of individuals. Once such efficiency gains have been achieved, it is also possible that additional consumption in other goods and services is made with the money that was saved through efficiency gains.

For instance, focusing on individuals and their lifestyle, Notter et al. (2013) have remarked that energy savings due to the adoption of energy efficient technologies in buildings and mobility seem to be compensated by an increase in the heated area per individual, by higher indoor temperature setting, by purchase of more equipment, or by higher demand for transport services. Namely, part of the potential energy and carbon savings resulting from energy efficiency measures are offset by direct and indirect effects stemming from them. Such effects are neither anticipated nor intended, though they can significantly reduce the size of the savings achieved. Overall, such phenomena are conceptualised as *rebound* effects (Sorrell and Dimitropoulos, 2008; Sorrell, Dimitropoulos, and Sommerville, 2009; Gillingham et al., 2020) and in economics they are usually referred to as the Khazzoom-Brookes postulate (Saunders, 1992). Even though empirical evidence on the size of rebound effects (especially, the indirect ones) has not yet been fully produced (Druckman, Chitnis, et al., 2011), these phenomena are widely recognised as relevant. Combined with the worldwide increase in population accessing use of energy services and products, rebound effects explain why, despite huge efficiency gains in many energy services and industry sectors, overall energy consumption is still increasing.

2.4 Energy sufficiency

In the face of such efficiency shortcomings, the concept of "energy sufficiency" has recently gained momentum. While energy efficiency deals with decrease of consumption in relative terms compared to some output, the concept of energy sufficiency deals with the reduction of overall amounts of consumed energy. Coherently with strong sustainability approaches, it aims at ensuring possibilities for long-term use of energy resources and at guaranteeing well-being for everybody.

Energy sufficiency fits in the broader concept of "sufficiency", that was inspired by ecological economists such as Georgescu-Roegen (1975) and Daly (1991) and conceptualised at the beginning of the 1990s by W. Sachs (1993) as a way to establish "the right measure": having enough to meet one's (not only purely material) needs (Schneidewind and Zahrnt, 2014). More recently, Princen (2005) conceptualised sufficiency as a set of organising principles for social life aimed at addressing overuse of resources: “a necessary condition, a set of decision criteria, a set of principles critical for reversing the biophysical trends and re-organising society for sustainable resource use” (p. 355).

Energy efficient consumption processes or activities might still imply high energy consumption levels in absolute terms; energy sufficient consumption processes or activities are instead characterised by low energy consumption in absolute terms. Furthermore, while energy efficiency aims at offering the same energy service output by optimising the amount of required energy input, energy sufficiency aims at delivering different energy service outputs, either in quality or quantity (Thomas et al., 2019). Namely, energy sufficiency aims at changing consumption practices —and for this purpose it requires action on both individual behaviour, daily routines and practices— and at changing infrastructures and institutions that support and drive them (Jungell-Michelsson and Heikkurinen, 2022). In practical terms, Schneidewind and Zahrnt (2014) have identified four founding principles around sufficiency (the “Four Lessens”): less speed, less distance, less clutter, and less market. Similarly, Spangenberg and Lorek (2019) have presented sufficiency as “the antithesis to the ‘faster, further, more’ orientation of the consumer society” (p. 1071). Sandberg (2021) has identified four types of consumption change that can be related with sufficiency: absolute reduction in consumption patterns (such as travelling shorter distances), shift from one mode of consumption to one that has a lower environmental impact (such as shifting from private car use to public transport), product longevity (extending product lifespan) and sharing products among individuals (such as car-sharing services). Specific examples of energy sufficient behaviours for households have for instance been listed by Moser et al. (2015) and Seidl et al. (2017): line-drying laundry instead of using a tumble dryer, eating vegetarian food instead of meat, commuting by bike instead of by private car, reducing living space per person by moving into smaller houses, or reducing room temperature and wearing warmer sweaters and socks in Winter.

2.4.1 Energy sufficiency, equity and human well-being

Muller (2009) argues that energy sufficiency might become a moral duty of liberal societies, which are called to ensure social justice and avoid external impacts from energy consumption that are harmful to other people. If normative principles of “enoughness” and sufficient levels of consumption get spread among the population, they can be accessed by a large number of individuals, who then benefit by an improvement in their well-being and increased chances for living a ”Good Life“ (Schneidewind and Zahrnt, 2014). On a similar same page, Druckman and T. Jackson (2010) have dealt with the idea of a “decent life”, that can “provide [people] with food and shelter for themselves and their families but also [allow them] to participate effectively in the life of society. [...] It is more than just food, clothes and shelter. It is about having what you need in order to have the opportunities and choices necessary to participate in society” (p. 1794).

Namely, the authors conceptualise a sufficient society that fulfils needs and eliminates large shares of current unnecessary, discretionary consumption.

More specifically, energy sufficiency has been conceptualised by Darby and Fawcett (2019) as an organising principle for living within ecological limits and providing social equity: it is "a state in which people's basic needs for energy services are met equitably and ecological limits are respected" (p. 362). Implementing such a principle guarantees that all people's *needs* are met, while some of their *wants* are not. Such a conceptualisation implies that, in order to guarantee equity and well-being for everybody, for some people energy consumption should *decrease* while for some people it should *increase*.

Energy sufficiency thus corresponds to achieving a level of energy consumption that allows to meet everybody's needs and thus equitably guarantee human well-being and distributive justice. Consuming more than such a threshold level would produce environmental and climate harm, not increase individual well-being, and ultimately make everybody worse-off (Burke, 2020). As recently argued by a collective of German scholars in response to the European energy-supply crisis brought about by the Russian war in Ukraine (Best et al., 2022), consuming less than the threshold would instead also be beneficial in terms of energy security and autonomy at the country level.

Early examples of such thresholds, however identified when the concept of sufficiency was still in its infancy and therefore also largely inspired by efficiency principles, where for instance the "one kilowatt per capita" estimated by Goldemberg et al. (1985) or the "2000 Watt per capita" estimated by Spreng (2005) and Notter et al. (2013), which then was formally turned into the "2000 Watt and 1 CO_2 ton Society" adopted by many Swiss cities and by the Swiss Confederation as a policy goal to be achieved by 2050. A recent review of policies for achieving carbon neutrality however concluded by calling for an "energy conservation and sufficiency revolution" (Bertoldi, 2022), grounded on pilot policy testing: energy sufficiency principles have still a long way ahead, before they become as mainstream as energy efficiency.

2.4.2 Socio-technical transitions towards energy sufficiency

There is general agreement that the shift towards an energy sufficient society can only stem from a multi-faceted reflection on human needs and wants, energy services, urban structures, social norms, consumption habits, and the related policies and interventions (Toulouse et al., 2019). Implementing the transition to an energy sufficient society in fact requires much more than progress in technologies —which is, however, in most cases a necessary precondition (Spengler, 2016; Millward-Hopkins et al., 2020). On the other hand, the transition cannot be simply achieved by a change in individual behaviour and practices or social innovation processes —which are however necessary as well (Schneidewind and Zahrnt, 2014; Samadi et al., 2017). It is also widely acknowledged that human needs and well-being, rather than technology fixes, should be the starting point for sufficiency (Burke, 2020).

In the last decade the practical potential for transition to a sufficient society has been explored by a number of scholars, with partially contrasting outcomes. Modelling and simulations of future energy scenarios by Wachsmuth and Duscha (2019) have for instance shown that complementing energy efficiency interventions with sufficiency measures could ensure the reduction of energy demand in absolute terms that is required to achieve the Paris Agreement's target of limiting average temperature increase well below 2 °C until 2100, at best even 1.5 °C. Previous works by Alcott (2008) and Figge et al. (2014) have instead warned that sufficiency measures at the individual level might also lead to rebound effects, just like efficiency measures.

In particular, Sorrell, Gatersleben, et al. (2020) have recently detailed a number of reasons why energy-sufficient consumer behaviours might lead to an increase in energy consumption: if, by consuming less, people save money and time, they may invest in other energy-intensive goods and services. Furthermore, by enacting sufficiency behaviours in some domains, individuals may feel they fulfilled their moral obligations, feeling licensed to indulge themselves in other energy intensive, materialistic consumption activities. Depending on the energy intensity of such activities, this may entirely or partially erode the energy and emission savings directly stemming from sufficiency behaviours. Extreme situations might even occur, when, as an indirect consequence of of sufficiency measures, energy consumption and carbon emissions ultimately increase (*backfire* effect).

These phenomena were for instance modelled by Druckman, Chitnis, et al. (2011), who analysed three sufficiency measures in households: lowering the heating thermostat by 1 °C, reducing food purchase by one third to eliminate food waste, and walking or cycling for trips less than two miles instead of using the car. Overall, they estimated an average 34% reduction in the anticipated decrease of carbon emissions by such measures, due to rebounds. If generalisation of their results were possible, this would imply that about two thirds of the carbon emission reductions expected by performing energy sufficiency measures could actually be achieved, while one third would be lost due to rebounds.

These estimates confirm the need for accompanying the change in individual behaviours with a broader change at the institutional and collective level. Overall, these processes would induce a socio-technical transition to a novel society, in which rebound would have no opportunities to occur, as energy consumption beyond sufficiency levels would neither be needed nor desired. Specifically for households, Spangenberg and Lorek (2019) envision a society characterised by radical changes in household needs and the way they are satisfied: needs should be satisfied with less consumption of new material goods and with more consumption of immaterial social and collective goods. Overall, the authors argue that "needs are to be satisfied in a different, more sustainable way, while conspicuous consumption is to be avoided" (p. 1071). Scholars such as Toulouse et al. (2019) have therefore called for future research activities aimed at questioning collective conventions about what basic needs are (such as for instance a reasonable level of comfort at home or frequent week-end short-breaks to other countries), at showcasing alternatives, and at negotiating new conventions. In the resulting novel societal context, the time, energy, and personal resources no longer devoted to the satisfaction of materialistic needs would provide opportunities for rewarding social interactions and for inner personal fulfilment, thus resulting in an increase in quality of life and personal well-being.

2.5 Behaviour

Having clarified that the energy and climate transition needs policy interventions at a variety of levels and in a variety of domains, since this dissertation focuses on persuasive apps which aim at changing behaviour of households, I now move to the specific concepts on which my work is grounded: behaviour, behavioural theories, gamification, social practices and households.

A broad and descriptive definition of behaviour is offered by the Cambridge English Dictionary, according to which behaviour represents "the way a person acts in a particular way in a particular situation or under particular conditions" (McIntosh, 2013). In their Dictionary of sociology, Lawson and Garrod (2012) also attempt some explanations about how and why a given behaviour is performed: they present behaviour as "the events that individuals engage in, which may or may not be intended and planned. Behaviour thus has several sources, from emotions, through instinct, to rationality. It does not have to involve purpose in the consciously planned sense".

Indeed, Kaufman et al. (2021) have presented behaviour as an "interdisciplinary boundary concept" (p. 599), which operates at the interface between psychological concepts of attitude, belief and social norms, sociological concepts of practice, agency and structure, cognitive science concepts of conscious and un-conscious activities, routines and habits, as well as broader natural and technical sciences. Based on an extensive review of behavioural research within the domain of sustainability transitions, the authors have identified four key perspectives according to which behaviour is conceptualised, arguing that policy interventions frequently adopt a combination between them:

- *Reflective behaviour*: behaviour is the result of conscious deliberations and activities by individuals, which are determined by internal factors (e.g. attitudes), individual-level context (e.g. resources), socio-cultural context (e.g. education or religious beliefs), and techno-economic context (e.g. government policies). From this perspective, changes in behaviour first require the identification of critical barriers precluding their performance, and then their manipulation, so that individuals, are lead to change their behaviour by conscious deliberation. Everyday actions of individuals and households are frequently the subject of reflective behaviour change approaches, which usually address multiple determinants, by acting on individuals' knowledge and beliefs, for instance through education, incentives, emotional or moral appeals, persuasive communication, or social norms.
- *Automatic behaviour*: behaviour is the outcome of unconscious, intuitive and non-deliberate actions of individuals, which are repeated over time and usually

are referred to as *habits* (Maréchal, 2010; Whitmarsh et al., 2021) or *routines* (Breukers et al., 2015). Automatic behaviours are thus produced by recurring cues or constraints on an individual's psychological or structural context: routines are initially deliberately formed but they rapidly become less or non-conscious, provided that the context remains stable. Behaviour change can therefore originate from both intentional manipulation and unintentional changes in those cues or constraints.

- *Everyday behaviour*: from this perspective, the focus of analysis and possible intervention moves from individuals (or households) to the social practices that are performed in a given spatial and temporal contexts. Here, behaviour is conceptualised as a cluster of actions (the social practices) that emerge from the relationships between individuals in their contexts. Such an approach, which characterises studies grounded in Social Practice Theories, is particularly suitable to energy consumption. In fact, it is not an end in itself, but it is performed for getting some services, such as cooking, lighting, water and space heating, or personal and freight transport (Goldemberg et al., 1985; Jonsson et al., 2011; Hess, I. Schubert, et al., 2022). Performance of a given practice is thus the result of constantly reproduced capabilities, materials, and negotiated meanings about that specific practice and the other social practices that are somehow inter-twined with it. The evolution of social practices can therefore be influenced by the evolution of such factors.
- *Strategic behaviour*: behaviour is an intentional action, undertaken for getting specific benefits in a competitive context by specific actors, which can for instance be either government agencies, established firms, or even emerging niche innovators. From this perspective, actors are conceptualised as collective entities, that operate at the meso-level within socio-technical transitions.

Persuasive apps mostly draw from the perspectives of reflective and automatic behaviour. This is the case for the enCompass and Social Power apps and the related interventions I analyse in the next chapters of this dissertation. Indeed, the reflective perspective is at the centre of many behaviour change theories developed within social psychology domains, which I briefly introduce in the next section.

In this work, however, I will also marginally refer to the everyday behaviour perspective, which is relevant for the case of the Social Power Plus app: the specific behaviour change features dealing with household routines that the app builds upon have in fact some connections with Social Practice Theories, which I introduce in the last dissertation section. Before entering into the theories, however, I provide a few more details about routines and habits, which are key to both automatic and everyday behaviour perspectives.

2.5.1 Routines and habits

With reference to the everyday perspective, Hess, I. Schubert, et al. (2022) posit that "practices can be perceived as routinised behaviours embedded in existing infrastructure/technology, preferences/social conventions, and knowledge" (p. 3). From the automatic behaviour perspective, instead, habits can be conceptualised as "repeated, automated, and identity-expressing actions, performed without much conscious thought", which are "stabilised by external contextual and structural factors" (Hess, I. Schubert, et al., 2022, p. 3). This perspective emphasises that, as energy costs are small, largely invisible and usually affected by misperceptions and selective attention (Steg and Vlek, 2009), energy consumption tends to be the outcome of unreflective, highly inertial habitual routines, that are performed without a conscious deliberation of alternatives (Whitmarsh et al., 2021) and in which energy costs are given lower priority and attention than other factors such as convenience and symbolism (Sorrell, 2015). When outcomes are satisfactory, in fact, habitual routines tend to be automatically repeated, triggered by cognitive processes that are learned, stored in memory, and then retrieved when individuals perceive a given situation is taking place (Steg and Vlek, 2009). This process is highly beneficial to the individual, as it allows to free up cognitive resources for other decision-making process that are perceived to be more difficult or important and thus require active and conscious deliberation. However, it is highly critical for behaviour change processes.

Whitmarsh et al. (2021) have argued that habits are responsible for behavioural lock-ins and are among the strongest obstacles to change. According to Breukers et al. (2015), unconscious routines can be changed by first making them conscious. However, the authors also acknowledge that time is needed for activating intentional behaviour change and for the changed behaviour to become a new established routine. Further, they also remark that, specifically in the energy domain, individuals may be less open to adopt new routines, for the fear of a reduction in comfort or convenience (which for example might stem from reduced heating levels in Winter or air conditioning in Summer). Alternative approaches to change routines are grounded in "nudge theories" developed within behavioural economics (Thaler and Sunstein, 2009), and focus instead on accompanying individuals to unconscious change (see Section 2.6.4). Anyway, as habits are cued by stable contexts, it is usually easier to re-orient them when significant changes in the context occur (Steg and Vlek, 2009). For this reason, Whitmarsh et al. (2021) have suggested that behaviour change interventions preferably target individuals that are experiencing significant changes in their lives, such as motiving to a different home or having a child.

A quite different strategy to break routines and evolve them was suggested by Schwanen et al. (2012) and then backed-up by Sorrell (2015). It consists in addressing the collective determinants of habitual behaviour, rather than the deliberate decisions of single individuals. This is especially aligned with the everyday behaviour perspective codified by Kaufman et al. (2021), according to which "practices rely on the availability of physical affordances for the actions, relevant meanings associated with the actions, and competence (or skills to perform the actions), with relational interaction between these elements over time and space" (p. 598). I provide some examples of how this can be done in Section 2.8 about Social Practice Theories.

2.6 Behavioural theories

Favouring a change in behaviours requires understanding how behaviours are, directly or indirectly, shaped, influenced and constrained. A huge body of research has attempted to tackle this challenge and to identify the antecedent factors of behaviour, by addressing it from different domains, such as psychology, economics, sociology, marketing, information technology, energy, and consumption studies. Many theories and models have been developed and some of them have also been empirically tested, however neither widespread agreement has been achieved nor univocal evidence have been found and the research challenge to get to a unified theory of behaviour is this still open.

Anyway, there is general agreement that several interrelated variables jointly influence and predict individual differences in household energy consumption. Abrahamse, Steg, et al. (2005) argued that energy consumption behaviour is driven by two types of factors:

- *macro-level* ones, such as the TEDIC factors (Technological developments, Economic growth, Demographic factors, Institutional factors, and Cultural developments);
- and the *micro-level* ones, such as motivational factors (preferences, attitudes), abilities, and opportunities.

Similarly, T. Jackson (2005) has proposed two broad categories of theories explaining behaviour: those that are mostly *internal* to individuals and those that are mostly *external* to them. Internal ones focus on behavioural antecedents such as attitudes, values, habits and personal norms. External ones instead focus on fiscal and regulatory incentives, institutional concerns and social norms, and coincide with the contextual category. With respect to the central debate in the social sciences on the role of agency (human action) and structure (the social institutions that constitute the framework for human action) in the configuration of social systems, T. Jackson (2005) has argued that "internal" theories maintain the primacy of individual agency: individuals are free to choose a given behaviour, provided they have appropriate beliefs, attitudes, or capabilities. "External" theories instead maintain the primacy of structures: individuals are locked-in in certain behaviours, due to a number of external conditions, such as for instance economic necessity, social expectations and institutional arrangements.

In an influential review of theories and models aimed at explaining pro-environmental behaviour[1], Steg and Vlek (2009) have further detailed previous dichotomous classifications of key behavioural determinants, by keeping the external category (contextual factors) and splitting the internal one in two categories. They have in fact proposed a threefold classification:

- *motivational factors*, such as attitudes, beliefs, values, intentions. These theories assume that behaviours are the outcome of individuals' deliberate and reasoned choices and are strictly related with their motivation to engage in a given behaviour. Motivation in turn may depend on the perceived costs and benefits of the behaviour

[1] Behaviour that consciously seeks to minimise the negative impacts of one's actions on the natural and built world (Kollmuss and Agyeman, 2002, p. 240).

and its alternatives, on moral and normative concerns driven by personal and social norms and dominant cultural values, or on affective and emotional factors;

- *habitual factors*, such as routines and habits. These theories assume that individual behaviour is habitual and guided by automated cognitive processes, rather than being preceded by elaborate reasoning;
- *contextual factors*, such as physical infrastructures, technical facilities, availability of products, price, fiscal and regulatory conditions, social and cultural norms. These theories assume that technical, economical and organisational factors determine behaviour indirectly, by shaping and constraining the social and institutional context within which behaviour is performed.

Again keeping the external category, also Frederiks, Stenner, and Hobman (2015b) have more recently proposed another threefold classification of behavioural determinants. They identified *socio-demographic* factors (e.g. education, income, household and house size, phase in family life cycle), *psychological* factors (e.g. values, attitudes, beliefs, motivations, personal and social norms), and external *situational and contextual* factors (e.g. socio-cultural, economic, political, legal, institutional forces). They have remarked that external factors have a key role in driving behaviour, as they may prevent households from performing given behaviours (or force them to do so), for instance through regulations or infrastructural availability, independently of socio-demographic and psychological orientations.

Though most of the behavioural theories developed so far can mostly be fitted in only one of such categories, many scholars (for instance T. Jackson, 2005; Steg and Vlek, 2009; Spangenberg and Lorek, 2019) have argued about the importance of engaging with both the social context that shapes and constrains social action and with individual choice and motivational factors, including short-cuts and heuristics that drive habitual, routinely behaviour. An attempt to bridge the agency-structure dichotomy has been performed in various propositions of Social Practice Theories, which can all be traced back to the Structuration Theory by Giddens (1984). According to this theory, social structure is both the medium and the outcome of people's social practices. Human agency takes place through the repetitive, routine practices of daily life, which are performed within the opportunities and constraints offered by social structures, such as language, rules, norms or meanings. Following Turner (2006), the Duality of structure theory by Giddens posits that "there is an ongoing reciprocal relationship between structure and agency. Structural circumstances provide the means to reproduce social practices, but when social practices are reproduced they perpetuate the structure, making it a social reality in a new historical moment" (p. 16). From this perspective and the later conceptualisation by Schatzki (2021), structural properties of social systems are constantly produced and reproduced by the action of individuals (*practices-as-performance*), who daily act as carriers of *social practices*, thus conditioning the maintenance or evolution over time of the practices themselves (*practices-as-entity*). An evolution of practices can therefore be fostered by policies that focus on the specific elements that drive the way a given practice is performed by the individuals. I dedicate Section 2.8 to a broader introduction to the concept of social practice and the related theories.

Another attempt to account for both the social context and individual action has for instance been proposed by Stern (2000). Starting from internal/motivational theories, he argued that behaviour theories should ideally account for: *motivations*, *attitudes* and *values*, *contextual* or *situational factors*, *social influence*, *personal capabilities*, and *habits*. Attempts to develop integrative theories, though including only some of these elements, have indeed been performed by the integrated Attitude-Behaviour-Context model by Stern (2000), by the Theory of Inter-personal Behaviour by Triandis (1977), by the comprehensive Model of Consumer Action by Bagozzi et al. (2002), or by the COM-B behaviour model by Michie, Van Stralen, et al. (2011).

The latter, COM-B, is particularly simple, and thus potentially well-suited to support behaviour change interventions. It models behaviour as the outcome of bidirectional interactions with and between *capability*, *opportunity*, and *motivation* factors (Figure 2.1). Specifically, following Michie, Van Stralen, et al. (2011), it distinguishes between:

- physical and psychological capabilities (i.e. "the capacity to engage in thought processes such as comprehension and reasoning", p.4);
- physical (i.e. "afforded by the environment") and social (i.e. "afforded by the cultural milieu that dictates the way we think about things", p. 4) opportunities;
- and motivation induced by reflective processes (i.e. "involving evaluation and plans") and automatic processes (i.e. "involving emotions and impulses that arise from associative learning and/or innate dispositions", p. 4).

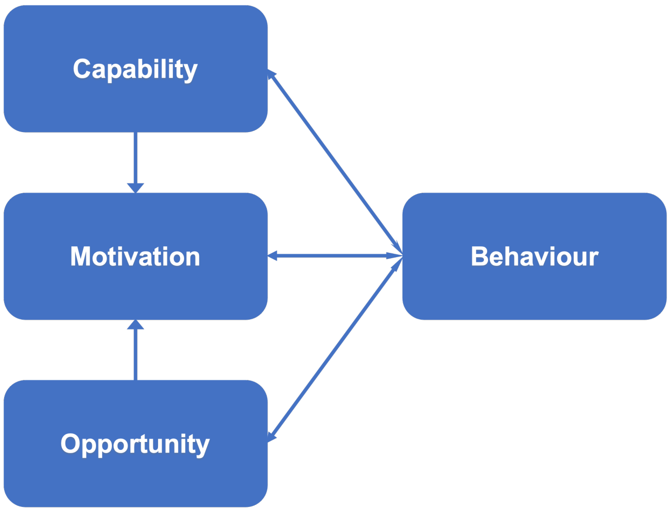

Figure 2.1: Schematic representation of the COM-B model.

Despite their theoretical relevance, these integrative theories have not been widely exploited yet in recent applied research, which instead is mostly grounded on internal/motivational theories. Whitmarsh et al. (2021) have in fact identified in the Theory of Planned Behaviour (TPB, Ajzen, 1991), the Value-Belief-Norm theory (VBN, Stern, Dietz, et al., 1999), and the Transtheoretical Model (TTM Prochaska and Velicer, 1997) the three most common behavioural theories used to inform climate change mitigation interventions. For this reason, in the next sections I briefly introduce them.

I also introduce the nudge theory (Thaler and Sunstein, 2009), which was developed in the context of behavioural economics and is increasingly used also in digital-based interventions to deal with habitual factors driving behaviour. I then introduce the Self-Determination Theory (SDT, Deci and Ryan, 2008) and the strictly-related concepts of intrinsic and extrinsic motivation, extensively used in the framework of gamification processes, which in turn are widely used by digital apps on which my work focuses. I conclude this section by introducing the Captology and persuasive technology framework (Fogg, 2003), that provides the theoretical grounds for many behaviour change apps.

2.6.1 Theory of Planned Behaviour

The Theory of Planned Behaviour (TPB) (Ajzen and Madden, 1986; Ajzen, 1991), originally developed in the social psychology domain, argues that behaviours are intentionally performed by individuals: they are the outcome of reflected, deliberate processes. According to TPB, an individual's intention to perform a given behaviour is the antecedent of behaviour itself, and it is predicted by three factors (Figure 2.2):

- the individual's *attitudes* towards the behaviour: people behave according to their beliefs and expectations on the outcome of their behaviour, which also influence their attitudes towards the behaviour itself. More specifically, beliefs and expectation of outcome lead to an attitude towards the given behaviour, which then affects the intention to perform a behaviour. Note that, if affective/emotional and moral aspects are included among such attitudes, TPB can also account for affective and moral antecedents of behaviour;
- the individual's *subjective norms*[2], namely the individual's perception about the particular behaviour, which is influenced by the judgment of significant others (i.e. the perception that people who are important to the individual think she should or should not perform the behaviour in question). Through this factor, TPB acknowledges social influences on personal behaviour;
- the individual's *perceived behavioural control*: defined as "the person's belief as to how easy or difficult performance of the behaviour is likely to be" (Ajzen and Madden, 1986, p. 457), this factor is directly related with resources and opportunities available to the individual to perform a given behaviour. Thus, even though indirectly and partially, TPB allows to also account for contextual factors[3].

TPB has been tested in a large number of applied research interventions, which confirmed its validity. However, most studies limited themselves to testing correlations between

[2]The construct of "subjective norms" has been referred to as "personal norms" in most of the later literature.
[3]The construct of perceived behavioural control is closely related with the construct of self-efficacy by Bandura (1977), which he defines as "the conviction that one can successfully execute the behaviour required to produce the outcomes" (p. 193).

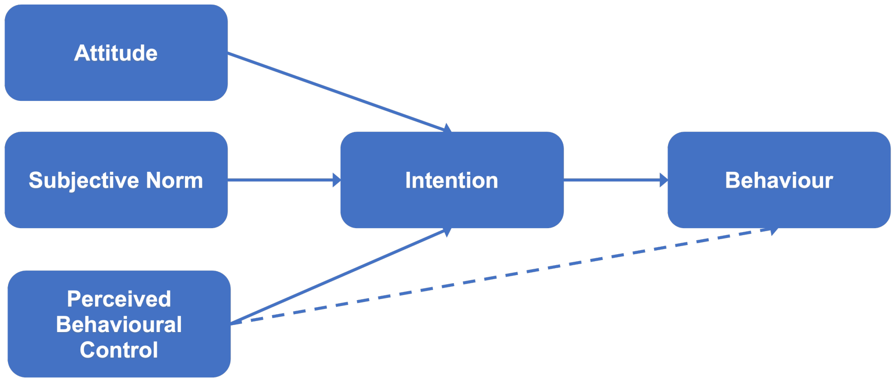

Figure 2.2: Schematic representation of the Theory of Planned Behaviour (TPB).

the three above factors and intention to perform the target behaviour —instead of also verifying correlations between behavioural intentions and actually performed behaviours. There is wide agreement among scholars that field interventions should aim at measuring actual target behaviours as much as possible.

2.6.2 Value-Belief-Norm Theory

The Value-Belief-Norm (VBN) theory of environmentalism by Stern, Dietz, et al. (1999) postulates that values influence behaviour, via pro-environmental beliefs and personal norms. The theory is informed by previous theories explaining behaviour (and particularly pro-environmental behaviour) as the outcome of individual feelings of strong moral obligation to engage in a given behaviour, such as for instance the Norm-Activation Model[4](NAM, Schwartz, 1977; Schwartz and Howard, 1981). Also, it is inspired by normative models of pro-environmental behaviour, according to which pro-environmental behaviour depends on pro-social, biospheric values by the individual.

According to VBN, depending on her value orientation (egoistic, altruistic or biospheric), the individual is characterised by different levels of acceptance of the "new ecological paradigm", which basically lists a set of core "biospheric values" (Dunlap and Van Liere, 1978) that reflect respect to natural planetary and resource limits and the importance of preserving the balance and integrity of nature. Individuals with strong biospheric values are expected to focus on the consequences of their behaviour on the environment; individuals with strong altruistic values are expected to focus on the consequences on other people; individuals with strong egoistic values are expected to focus on the consequences on their well-being. These value orientations affect people's general beliefs about the relationship between humans and the environment and, more generally,

[4]NAM, which provides a framework to understand altruistic behaviour, assumes that altruistic behaviour is driven by personal norms, namely by feelings of personal obligation to act in a particular way in specific situations. In turn, two factors act as determinants of personal norms: the awareness of consequences of one's actions and the acceptance of personal responsibility that one holds for such consequences. Namely, when an individual gets aware of the harmful consequences of her behaviour and also accepts personal responsibility for causing this harm, this may elicit negative feelings, such as guilt. These negative feelings thus activate an individuals' personal norm, namely the obligation to behave more in line with personally important moral standards.

individual's ecological worldviews. In turn, ecological worldviews are supposed to influence pro-environmental behaviour through awareness of adverse consequences of one's behaviour on the environment. Finally, such an awareness is expected to activate the individual's ascription of responsibility (ability and willingness to assume responsibilities for those consequences), which ultimately triggers personal norms, by activating a moral obligation to act pro-environmentally (Figure 2.3).

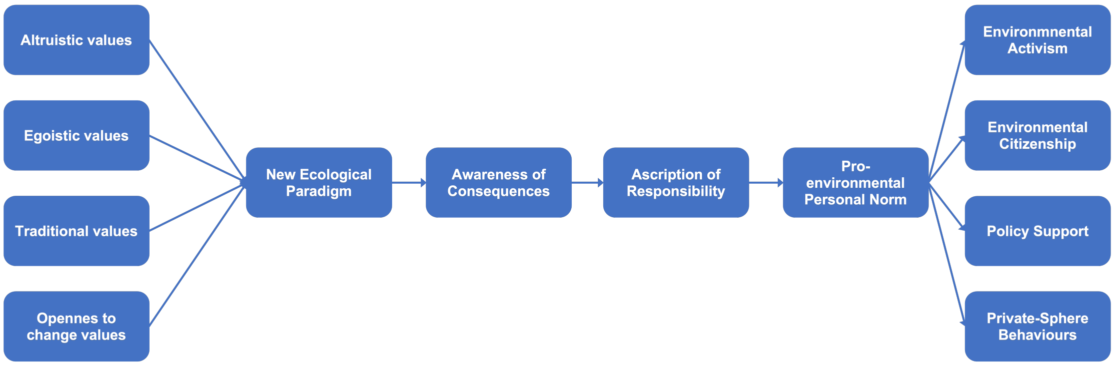

Figure 2.3: Schematic representation of the Value-Belief-Norm Theory (VBN).

2.6.3 The Transtheoretical Model of behaviour change

The Transtheoretical Model of behaviour change (TTM) by Prochaska and Velicer (1997) was originally developed in the health domain. It differs from the theories I presented so far, as it does not aim at identifying the antecedents of a given behaviour. Rather, it aims at dealing with the process through which single individuals change their behaviour. The authors in fact remark that behaviour change does not happen instantly, as a single event. To the contrary, change occurs throughout a time-ordered process, during which the individual progresses along a series of cognitive and motivational stages, from pre-contemplation to maintenance of change. In *pre-contemplation*, the individual has no intention to take action for change in the foreseeable future. When she develops the awareness that a change may be needed, she enters the *contemplation* stage, which is followed by *preparation*, during which she forms the intention to take specific actions in the immediate future. When she starts acting, she enters the *action* stage, during which she actually changes her behaviour. The process is not concluded yet, however, as a long *maintenance* phase is then needed, during which the individual has to keep performing the new behaviour and prevent possible relapse back to the contemplation or even pre-contemplation stage (Figure 2.4).

Figure 2.4: Schematic representation of the Transtheoretical Model of behaviour change (TTM).

The TTM is not the only stage model of behaviour change that has been developed. Other stage models are for instance the Model of Action Phases (MAP) proposed by Heckhausen and Gollwitzer (1987) and Gollwitzer (1990), which includes four stages (*pre-decision*, *pre-action*, *action*, *post-action*) and has then further been enriched by other scholars who attempted to also include behavioural antecedent factors grounded in the Norm-Activation Model and the Theory of Planned Behaviour (Bamberg, 2013; Ohnmacht et al., 2017). Or, the stage model by Li et al. (2011), that only distinguishes between two stages: *discovery* and *maintenance* of the new behaviour. The TTM is however probably the most used stage model of behaviour change in applied research, not only in the health domain, but also in energy and climate transition processes. It is particularly appreciated as, besides identifying behaviour change stages, for each stage it also suggests which specific intervention strategies and techniques can be activated in order to favour progress to the next stage, thus providing practical and operative suggestions on how to spur change.

2.6.4 Nudge and behavioural economics theories

The above theories and models rely on typical economical assumptions of deliberative, and mostly rational, behaviour: according to them, when individuals are faced with a choice, they undergo complex cognitive processes, which lead them to develop "behavioural intentions", that are then followed by the performance of the chosen behaviour. Assumptions behind rational choices have however been questioned since the 1950s by cognitive and social psychology as well as by behavioural economics (Lehner et al., 2016). Individual actions are in fact constrained by the difficulty of processing information, understanding a situation, identifying the consequences of possible alternative actions and choosing between them. Indeed, human mind is characterised by limits to cognitive capacity: individuals have "bounded rationality", choices are made under incomplete information settings, and human actions are affected by cognitive constraints and biases, which for instance drive us to focus on some things and ignore others, to make decisions based on rules of thumb, and sometimes even not to make deliberate choices at all, but to operate under the effect of habits and mental shortcuts (Byerly et al., 2018).

In this framework, the "Dual process theory" developed in behavioural economics (Wason and J. Evans, 1974) posits that human decision-making occurs through two different processes, which are frequently referred to as "systems":

- *System One* drives automatic-thinking processes, namely simple, intuitive, emotional, automatic and fast decisions;
- *System Two* drives reflective-thinking process, namely slower, more supervised and more effortful decisions.

Following Kahneman (2011), behavioural economics acknowledges that many everyday decisions are performed by System One, which converts familiar tasks into automated routines, that are based on heuristics and affected by cognitive biases. This leaves room and cognitive resources for System Two to deal with more rare and complex decisions, that are taken through deliberate, reflective processes. Further, behavioural economics accepts that individuals may not always operate under strict self-interest assumptions and that they may even act under the effect of emotions. Namely, it acknowledges that often behaviour is definitely other than intentional or planned: it is impulsive, habitual, and emotional (John et al., 2009; Avineri, 2012; Dolan et al., 2012; Lehner et al., 2016; Byerly et al., 2018; Wee et al., 2021).

Accepting such principles, behavioural economics has attempted to develop theories on how human behaviour unfolds, which overcome the simplifications by economic models grounded on rational decisions, and whose insights can directly inform policy-making. The Nudge theory developed by Thaler and Sunstein (2009) has in particular become well established and it is increasingly used to support changes in behaviour that are beneficial to both society and the individual as well. Using the words of its creators, the concept of nudge refers to "any aspect of the choice architecture that alters people's behaviour in a predictable way without forbidding any options or significantly changing their economic incentives" (p. 6). In the above dual-system context, the role of a nudge is to counteract heuristics and cognitive biases, favouring unconscious and automatic actions to evolve into more socially and environmental desirable actions. This is done by acting on "choice architecture", namely on the context in which individuals make choices, with the aim of encouraging socially desirable behaviours and discouraging socially undesirable ones. For instance, this entails *changing default options* that are offered to individuals (by pre-selecting the socially and environmentally preferable options), *simplifying the provision of complex information* (e.g. through labels and product standards), or *providing feedback* on one's *own* or other *peers'* behaviour (which would lead individuals to reflect on their own actions and thus break "System One routines").

Nudges thus influence people's choice of action without limiting their options or enforcing rules and regulations: they guide people in a desired direction when making decisions, without introducing coercion. This is why Thaler and Sunstein have also introduced the concept of "libertarian paternalism" to explain how nudges work. These ideas have activated an intense debate, as many scholars have perceived nudges as ways to manipulate people, that are inherently immoral and non-ethical (Hansen and Jespersen, 2013; Mols et al., 2015; Hansen, 2016; C. Schubert, 2017): who should be entitled to decide which behaviour to nudge?

For instance, Hausman and Welch (2010) firmly argued that "Systematically exploiting non-rational factors that influence human decision-making, whether on the part of the government or other agents, threatens liberty, broadly conceived, notwithstanding the fact that some nudges are justified. Publicity, competition and limits to human abilities to influence choices limit the threat. But once the character of the paternalism in Thaler and Sunstein's 'libertarian paternalism' has been clarified, its risks to an agent's control over her own deliberation are evident" (p. 136). Indeed, I do not consider nudges as unethical and follow Hansen and Jespersen (2013), who argue that certain types of nudges might even result in empowering individuals. This occurs when nudge interventions do not aim at changing individual behaviour via hidden psychological manipulation and, to the opposite, rather promote actions that are consistent with individual preferences. This happens when nudges aim at making features, actions or consequences salient to the individual, such as through feedback or commitment setting.

Nudge techniques have been experimented to promote a broad range of pro-environmental and sustainable consumption behaviours. For example, the review by Lehner et al. (2016) accounts for the use of nudging in the domains of household energy consumption in buildings, mobility and food; the review by Byerly et al. (2018) explores use of nudges for family planning, land management, meat consumption, transportation choices, waste production and water use; the review by Wee et al. (2021) examines use of nudges to promote green purchases, recycling and the promotion of healthy consumptions. Indeed, nudge is a collective term for different policy tools and, as the concept is still relatively new, there is no universal agreement between the scholars on what fits within the nudge definition and what, though promoting a change in behaviour, should not be referred to as a nudge (Hansen and Jespersen, 2013; Hansen, 2016; Berger et al., 2022). Table 2.1 summarises the types of nudges that were considered in two of the above-mentioned literature reviews. Even though very similar concepts emerge, there is no full agreement on their definition and on the specific list of techniques to be considered as nudges.

Table 2.1: Examples of nudge techniques, as proposed in two extensive literature reviews of nudge interventions.

Techniques listed by Lehner et al. (2016)		Techniques listed by Byerly et al. (2018)	
Name	Description	Name	Description
Changes in the choice architecture	Introduce changes in the environment that guide and enable individuals to make different choices.	Priming	Provide subconscious information and sensory cues.
Simplification and framing of information.	Offer feedback on energy consumption (through informative energy bills, metering, or displays) or energy labelling of appliances and buildings.	Messenger Salience	Effects of a behaviour change information depend on who conveys the information. Use reminders and message framings that capture attention.
Changes to the default option	Offer opt-out strategies for green electricity products or for engagement in smart-grid trials, where technologies control and manage consumption.	Defaults	Introduce automatic settings or baseline reference points.
Use of descriptive social norms	Provide social comparison feedback on the consumption of similar households.	Norms	Provide information on the behaviours and expectations by others.
—	—	Commitments	Introduce explicit goals, pledges and promises to change behaviours.

Though the concept of nudge is very broad, insights supporting behaviour change emerging from behavioural economics do not limit to nudges. A number of behavioural researchers have in fact suggested how to support behaviour change by addressing automatic, fast, unconscious and affective processes that drive individual actions — instead of the reflective processes that, driven by education and information provision, lead individuals to make decisions based on deliberative, cognitive assessments of costs and benefits of available alternatives. For instance, Frederiks, Stenner, and Hobman (2015a) have identified eleven key cognitive biases that affect individual behaviour in the energy domain, by also providing useful recommendations to address them[5]. Similarly, Dolan et al. (2012) have developed the "MINDSPACE" framework, which identifies nine principles that can be applied in any behavioural domains. Such principles, which are all encompassed in the framework's acronym (Messenger, Incentives, Norms, Defaults, Salience, Priming, Affect, Commitments, Ego), are integrally reported in Table 2.2. For many of such principles, as well as for the recommendations by Frederiks, Stenner, and Hobman (2015a), overlapping with practical features of nudge interventions is very high. As from the perspective of practical intervention implementation distinguishing between them is not fully relevant, in the next chapter I broadly refer to nudges to also account for elements arising from behavioural economics in general.

Table 2.2: The behavioural-economics inspired MINDSPACE principles. Integrally reported from Dolan et al. (2012).

Principle	Behavioural aspect that is addressed
Messenger	We are heavily influenced by who communicates information to us.
Incentives	Our responses to incentives are shaped by predictable mental shortcuts such as strongly avoiding losses.
Norms	We are strongly influenced by what others do.
Defaults	We "go with the flow" of preset options.
Salience	Our attention is drawn to what is novel and seems relevant to us.
Priming	Our acts are often influenced by sub-conscious cues.
Affect	Our emotional associations can powerfully shape our actions.
Commitments	We seek to be consistent with our public promises and reciprocate acts.
Ego	We act in ways that make us feel better about ourselves.

2.6.5 Self-Determination Theory

The Self-Determination Theory (SDT) (see for instance Ryan and Deci, 2000b, Deci and Ryan, 2004 or Deci and Ryan, 2008) is a macro-theory of human motivation, which focuses on the conditions that enable or hinder the motivation to perform a given behaviour. SDT acknowledges that individuals' motivation to enact a given behaviour is driven by a number of factors. Basically, either individuals spontaneously perform a given behaviour without any kind of conditioning, because they find it interesting, novel, challenging, just for the mere pleasure of carrying it out (*intrinsic motivation*), or they perform the behaviour because they are pressured to act by factors that are external to

[5] I reported the full description of such cognitive biases and related recommendations in Table 3.3.

themselves (*extrinsic motivation*). In the latter case, they perform the behaviour in order to achieve some separable outcome. Intrinsic motivation is associated with spontaneous interest and exploration, vitality and enjoyment of life, and it is the manifestation of a human tendency towards learning and creativity.

Before SDT was fully theorised, Geller et al. (1990) argued that enduring, long-term behaviour change is more likely to result from interventions that minimise use of extrinsic motivational factors: "powerful extrinsic motivators are assumed to inhibit individuals from gaining an internal justification for performing the target behaviour after the external controls are withdrawn" (p. 127). However, much of what people do is however not intrinsically motivated. Indeed, Ryan and Deci (2000a) have argued that motivation to perform a given behaviour can range from "amotivation" or unwillingness, to passive compliance, up to active personal commitment. Namely, the authors have developed a "taxonomy of human motivation" (Figure 2.5), which identifies different levels of extrinsic motivation, that can stem from either personal endorsement and feelings of choice or from compliance with external regulations.

In particular, Ryan and Deci have conceptualised most of extrinsically motivated behaviours as "externally regulated" behaviours, that are performed to satisfy an external demand or reward contingency (e.g. vouchers or discounts for local shops, or punishments). However, they have also argued that some extrinsically regulated behaviours could be "internally regulated" (with varying degrees of autonomy, ranging from "introjected" to "integrated" regulated[6]), depending on how much individuals internalise regulations and assimilate them to the self. Among the extrinsic motivational factors, Ryan and Deci (2000a) present the external regulation ones as "impoverished forms of motivation", while the integrated regulation ones as "active, agentic states", which tend to have comparable motivational effect as intrinsic factors (p. 55). Namely, the higher the autonomy perceived by an individual, the more the external motivation factors promote internalisation and integration of the regulation (self-determination), and thus the closer extrinsic motivation factors get to intrinsic motivation.

Indeed, the different types of motivation lie along a continuum of autonomy. Ryan and Deci cite a number of previous studies that, in the educational domain, found that more autonomous extrinsic motivation is associated with more engagement, better performances, lower dropout, and higher quality learning. More generally, advantages of greater integration also include more behavioural effectiveness, greater volitional

[6]According to Ryan and Deci, introjected regulation occurs for instance when behaviours are performed to avoid guilt or anxiety or to attain ego enhancements, such as pride and feelings of worth. Regulation through identification occurs when individuals consciously value a behavioural goal or regulation as personally important. Integrated regulation occurs when regulations are fully assimilated to the self, namely they have been "evaluated and brought into congruence with one's other values and needs" (Ryan and Deci, 2000b, p. 73). This can for instance be fostered by creating opportunities for goal setting or for feelings of community between individuals, by prompting collaboration activities between them. Anyway, the related behaviours are still extrinsic, as they are performed in order to get separate outcomes, and not because they are inherently enjoyable.

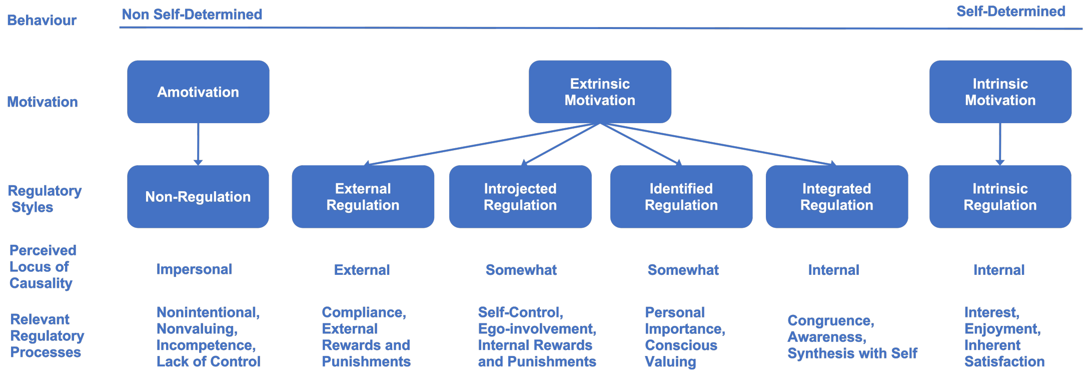

Figure 2.5: Schematic representation of the Taxonomy of Human Motivation by Ryan and Deci.

persistence, enhanced subjective well-being, better assimilation of the individual in her social group, and more constructive social development (Ryan and Deci, 2000b).

Based on extensive experimental research conducted by its authors over at least two decades, SDT argues that integration of motivation to perform a given behaviour can occur when three innate psychological needs are satisfied (Ryan and Deci, 2000b):

- *autonomy*: "being the perceived origin or source of one's behaviour" (p. 7);
- *competence*: "feeling effective in one's own interactions with the social environment and experiencing opportunities to exercise and express one's capacities" (p. 8);
- *relatedness*: "feeling connected to others, to caring for and being cared for by those others, to having a sense of belongingness both with other individuals and with one's community" (p. 9).

Through experimental tests in laboratory followed by studies on the field, Ryan and Deci have proven that social-contextual events such as the provision of feedback, communication, and rewards that produce feelings of competence, can increase integrated motivation for an action. Particularly, they have found that positive performance feedback increases intrinsic motivation, while negative performance feedback decreases it, through the mediation of perceived competence (Ryan and Deci, 2000b). Furthermore, they have also found that, to increase intrinsic motivation, feelings of competence also need to be accompanied by feelings of autonomy: people must not only experience competence or self-efficacy, they must also perceive their behaviour as self-determined. When opportunities for self-determination are offered, intrinsic motivation is enhanced, as a consequence of greater feelings of autonomy (Ryan and Deci, 2000b). Also, they found that, when relation bases with other individuals are secure and available, intrinsic motivation is enhanced as well. To the opposite, they found that tangible rewards, imposed deadlines, directives, evaluations, or goals decrease intrinsic motivation. Ryan and Deci (2000b) explain that this occurs when the individual perceives the behaviour as if it were due to an external *locus of causality*.

Finally, Ryan and Deci found that if extrinsic rewards are made contingent on task performance, they can undermine intrinsic motivation. This is also coherent with the

Motivation crowding theory by Frey and Jegen (2001), according to which use of external regulation motivational factors, and particularly of tangible prizes, increases individuals' intrinsic motivation and self-determination to change only if individuals perceive them as "supportive" —namely, if individuals perceive they foster their self-esteem and at the same time strengthen their perception of being free to act, instead of being controlled. Otherwise, external regulation motivational factors only have short-term effects, which disappear when the reward is no longer in place.

Encouraging individuals to perform activities that are uninteresting for them, namely activities for which they are not intrinsically motivated, thus requires to promote autonomous regulation for extrinsically motivated behaviours. Ryan and Deci suggest this can for instance occur when behaviours are prompted or modelled by significant others to whom individuals feel attached or related —provided that individuals have the capability to perform such behaviours. Namely, relatedness is crucially important for the integration of regulation. However, integration of extrinsically motivated activities also requires an increase in perceived competence. For these reasons, Ryan and Deci suggest that social contexts supportive of autonomy, competence, and relatedness can foster greater integration and regulation of non-intrinsically motivated behaviours, towards the generation of motivation and active commitment to perform them. Practically speaking, this can for instance consist in the introduction of rewards for activities that the individual feels competent to comply with, in the endorsement of specific behaviours by relevant references groups, or in the provision of choices and freedom to act, which enhance perceptions of volition and autonomy. Only when people feel competent enough to perform a given behaviour and when the social context provides support for autonomy, the integration of the relevant regulation is likely to occur, thus providing the foundation for subsequent self-determined behaviours.

2.6.6 Captology and persuasion

The discipline of persuasion, which can be defined as "an attempt to change attitudes or behaviours (or both) without using coercion or deception" (Fogg, 2003, p. 15), plunges its roots in ancient times, with the art of rhetorics by Greek philosophers. Since then, it has been widely investigated and used in the domains of marketing and psychology, mediated by printed materials or, more recently, mass media such as radio and television. The diffusion of computer facilities opened up novel possibilities for persuasion: computers and the Internet can enable new ways to spread persuasive messages to specific target groups and also offer novel, interactive persuasive features.

A new research stream was in particular created a couple of decades ago within the broader domain of Human-Computer-Interaction (HCI), with the very aim of exploring the persuasive potential by computing devices. A new concept was even introduced under the name of *captology*, namely the study of Computers As Persuasive Technologies (CAPT, Fogg, 1998; Fogg, 2003). Captology investigates if and how people can be motivated to change when interacting *with* computing products —and not simply interacting *through* them, which would instead be related with studying computer-mediated communication (the way people communicate or interact with each other through a computer device). In HCI and captology, computing products are regarded as active participants to the mediation and, to all intents and purposes, can be source of persuasion themselves.

Fogg (2003) also defined the concept of *persuasive technology*, namely "any interactive computing system designed to change people's attitudes or behaviours" (p. 1). The key behaviour change added value offered by computer tools lies in their interactivity, which allows persuasive technologies to customise, modify and evolve their persuasion techniques based on user input, feedback, needs, and contexts. Fogg has argued that persuasive technologies also offer additional benefits, such as persistence over time, easy scalability across a large number of users, large data-sets storage and processing capacity, ubiquity and embeddedness in everyday objects and environments. Provided that strict criteria are adopted, persuasive technologies also offer anonymity conditions.

Besides providing conceptual grounds for persuasive technologies, Fogg has also developed a behaviour model to inform their design, which in literature is currently referred to as "Fogg's behaviour model" (Fogg, 2009, Figure 2.6). According to such a model, a given behaviour can be performed if an individual has sufficient *motivation* and *ability*. If these are lacking, they can be favoured by effective *prompts* (originally called *triggers* and renamed by Fogg himself in 2017). The model accepts compensation between motivation and ability: a given behaviour can be performed by an individual with high motivation and low capacity, as well as by an individual with high capacity and low motivation. Proper (and different) triggers need however be provided: Fogg calls those aimed at increasing capacity as *facilitators* (e.g. increasing knowledge on how to perform a behaviour, for instance through the provision of information), and those aimed at increasing motivation as *sparks* (e.g. emphasising self-salient goals).

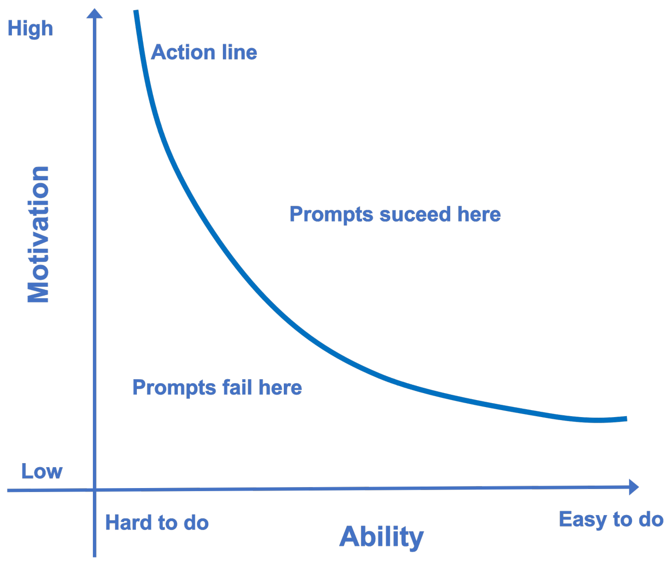

Figure 2.6: Schematic representation of Fogg's behaviour model (source: `http://www.behaviormodel.org/`, accessed online on November, 30 2022).

2.7 Gamification

Gamification is a concept introduced at the end of the 2000s, to refer to use of "gameful" activities to motivate performance of given activities or behaviours. Closely related with the idea of developing games to train, educate, and persuade people to perform given activities in their real life (which are usually referred to as "serious games" or "games with a purpose"), gamification does not rely on full-fledged games. Rather, it exploits game design elements such as points, leaderboards, levels, narratives or time constraints, and game mechanics such as competition, cooperation, assignments and goals (Weiser et al., 2015; Krath et al., 2021), and applies them into real-life contexts, products or services to motivate a desired behaviour (Deterding et al., 2011; Hamari, Koivisto, and Sarsa, 2014). Differently than games, gamification has no purely entertainment purposes; however, it aims at making performance of a given behaviour more entertaining. The desired behaviour is expected to emerge as a result of positive, intrinsically motivating gameful experiences made possible by the presence of novel motivational affordances implemented into the gamified process.

Adopting the definition proposed in a seminal paper by Deterding et al. (2011), gamification is usually referred to as "the use of game design elements in non-game contexts" (p. 1). Alternative definitions have also been proposed, such as the one by Hamari (2019), who focused on the novel motivational affordances that gamification brings about, defining gamification as "transforming activities, systems, services, products or organisational structures to afford gameful experiences". Other scholars provided definitions that focus more closely on the real-life impact expected by gamification processes. Zichermann and Cunningham (2011) defined gamification as "the process of game-thinking and game mechanics to engage and solve problems" (p. XIV) and Kapp (2013) defined gamification as "the use of game-based mechanics, aesthetics, and game-thinking to engage people, motivate action, promote learning, and solve problems" (p. 54). A definition attempting to include all such aspects has also been proposed by Seaborn and D. I. Fels (2015), as "the intentional use of game elements for a gameful experience of non-game tasks and context" (p. 17).

Gamification mechanics, such as rewards and loyalty programmes in marketing and grades in schools, have been used long before the emergence of the gamification concept and the related research domain. It was however thanks to the development of ICTs, the diffusion of mobile technology, the availability of cheap sensors to track human behaviour and everyday activities in smart cities, and a growing cultural openness to video-games and digital entertainment systems (Deterding, 2012), that the idea of gamification gained momentum. Namely, since the creation of the concept, the term "gamification" has nearly always been referred to digitally-enabled processes and activities. First, it spread among companies and commercial settings, then among researchers in the domains of Human-Computer-Interaction, computer science, game studies, and psychology, and finally also among policy-makers for the promotion of public health, education, or civic engagement.

Gamification received a few critiques from scholars blaming it for the "pointsfication" of behavioural processes, due to its widespread use of points, badges and leaderboards, which are attributed at the individual level and are perceived to be empty and sterile motivational factors, compared to community-based strategies that encourage individuals' continued involvement (Antin, 2012; Kapp, 2013). Nonetheless, applications of gamification have spread fast also in the sustainability domain to encourage behaviours such as reducing the amount of resources used, investing in recycling initiatives and renewable forms of energy, reusing materials, decreasing car use and increasing active mobility, or reducing energy consumptions in buildings.

As remarked by Seaborn and D. I. Fels (2015), gamification first started in business and marketing settings, and then entered the research domain. This is why early gamification works are characterised by the lack of standards of practice for design and implementation. Indeed, as Nacke and Deterding (2017) noted, first empirical gamification research was mostly aimed at understanding whether gamification interventions were effective in achieving their goals, and was only marginally grounded in theories. Reference to theory would however have allowed to both explain the interventions' outcome and to better design them, grounding their mechanics and components into the cognitive, emotional and motivational mechanisms that favour achievement of impact. Later research, just a handful of years after, started to engage with theories, with the aim of understanding why gamification works and which intervention types and components are more effective, in order to provide practical support to policy-makers and the industry as well. Though the concept of gamification is still in its infancy, as it has been formalised only less than a decade ago, Nacke and Deterding (2017) in fact noted that the transition from "theory-less" to "theory-driven" empirical research is already an ongoing process. Current research, in particular, is attempting to understand which psychological and social processes and contextual conditions play a role in driving outcomes of gamification interventions.

According to an extensive review by Seaborn and D. I. Fels (2015), the key theoretical reference adopted in empirical gamification research is the Self-Determination Theory. Use of points, badges and leaderboards, accompanied by goal setting, information feedback on progress made, and comparison with other members of the community engaged in the gamification process, is expected to activate an integrated regulation experience in the user, that builds on feelings of competence, autonomy, and relatedness. Use of gamified elements has also been grounded into the Theory of planned behaviour and in the Transtheoretical modell, as well as in Fogg's behaviour model, even though in less cases. A recent systematic review of literature reviews (meta-review) on gamification, serious games and game-based learning empirical interventions, has for instance found use of 118 different theories, which originate from a variety of research streams, including cognitive psychology, social psychology, and human-computer interaction (Krath et al., 2021). This analysis has however confirmed SDT as the most popular theory, used by 82 studies. This is probably because SDT is a macro-theory of human motivation, a broad framework that can be applied to different contexts. Findings by Krath et al. (2021) also show that most of the used theories have explicit conceptual connections between each other: properly identifying and systematising them would allow to identify the key principles about how gamification works and to clarify its empirical potential. A preliminary attempt to summarise the commonalities among such theories is offered by the authors themselves, who identify a list of ten key principles why gamification works (Table 2.3), suggesting to start using them in the design of gamified processes, in serious games and in game-based learning.

Table 2.3: Key theoretical principles that help explain the effects of gamification, according to Krath et al. (2021) (integrally reported from their work).

Theoretical principle	Explanation
P1 Clear and relevant goals	Gamification can transparently illustrate goals and their relevance.
P2 Individual goals	Gamification can allow users to set their own goals.
P3 Immediate feedback	Gamification can provide users with direct feedback on their actions.
P4 Positive reinforcement	Gamification can reward users for their performance and communicate the relevance of their achievements.
P5 Social comparisons	Gamification can allow users to see their peers' performance.
P6 Social norming	Gamification can connect users to support each other and work towards a common goal.
P7 Adaptive content	Gamification can adapt tasks and complexity to the abilities and knowledge of the user.
P8 Guided paths	Gamification can nudge users towards the actions necessary for achieving the goals.
P9 Multiple choices	Gamification can allow users to choose between several different options to achieve a certain goal.
P10 Simplified user experience	Gamification systems are usually easy to use and can simplify content.

Other recommendations for applied research are provided by Seaborn and D. I. Fels (2015), who, again referring to SDT, suggest to invest in intrinsic or internally driven motivation, beyond relying on extrinsic or externally regulated motivators such as points and rewards. They suggest to cater to intrinsic values of end-users and, for this purpose, recommend to adopt user-centred approaches (Norman, 2013), which allow to focus design of gamified systems and apps on the needs and desires of potential users. A similar recommendation is also made by Morganti et al. (2017). User-centred design, in fact, favours the identification of integrated regulation motivators for specific target groups of users and specific contexts, as well as the identification of their favoured customisation and personalisation possibilities, thus increasing chances for effectiveness and impact. Finally, as noted by Klock et al. (2020), who performed a large literature review on the way gamification contents and the related user experience are tailored to different user groups in different contexts, there is a clear trend towards customising gamification to individual needs and preferences (e.g. customisation of contents based on individual taste, adaptation to different users in the same context, provision of customised recommendations for user and context combinations). Such customisation processes are however seldom grounded in theories and would benefit by higher engagement with them.

Similarly, also Nicholson (2015) suggests that gamification should go beyond points, leaderboards and badges, and proposes his "RECIPE" for meaningful gamification processes, namely a list of six elements that would allow to increase autonomy, competence and relatedness, thus triggering an increase in integrated motivation for change. The recipe by Nicholson is composed by the following ingredients (p. 5):

- *Reflection*: assist participants in finding other interests and past experiences that can deepen engagement and learning;
- *Engagement*: encourage participants to discover and learn from others interested in the real-world setting;
- *Choice*: develop systems that put the power in the hands of the participants;
- *Information*: use game design and game display concepts to allow participants to learn more about the real-world context;
- *Play*: facilitate the freedom to explore and fail within boundaries;
- *Exposition*: create stories for participants that are integrated with the real-world setting and allow them to create their own.

2.8 Social Practice

In the last decade, inspired by the seminal article by Shove (2010a), many scholars have criticised the above behavioural approaches as they too narrowly focus on individual attitudes and choices, failing to consider the broader social obligations and structural factors that condition daily behaviours. Further, behaviour change approaches have been said to conceptualise energy demand as a stable quantity, that has to be satisfied, independently on the mediating infrastructures, technologies, and social practices underlying it (Royston, Shove, et al., 2017). Scholars advancing such critiques have argued that "consumers needs have histories and futures that are not fixed, not natural, and not inevitable either" (Royston, Selby, et al., 2018, p. 133). From their perspective, to support the energy and climate transition focus should be moved from energy and individuals to *what energy is used for*, namely to the *services* provided by energy (Wilhite et al., 2000) and to the *social practices* that create energy needs (Butler et al., 2018).

Indeed, individuals do not use energy for its own sake; rather, energy consumption emerges as a by-product of everyday activities (Shove and G. Walker, 2014). Everyday practices should thus be considered as the starting point of any analyses or interventions aimed at reducing energy consumption, to capture the variety of ways in which normality is made and reproduced (Wilhite et al., 2000). Energy demand is in fact inseparable from "what people and their machines do in their homes, at work, in leisure time, and in moving around" (Royston, Selby, et al., 2018, p. 128).

These authors, who ground their reasoning in Social Practice Theories, call for approaches capable to understand and transform the social and political organisation of society (the "social organisation of normality", Shove, 2003) that configures the amounts and patterns of energy demand. They posit that fundamentally challenging existing collective conventions, such as the average indoor temperature in households, is needed to "radically reconfigure the ways in which energy demands are enmeshed in the social, institutional and material fabric of society" (Royston, Selby, et al., 2018, p. 129). Societal institutions are called to act in a systemic way, to "normalise an alternative set of practices, values, beliefs, stories, and symbols and to serve as a guide for new lifestyles and infrastructures" (O'Rourke and Lollo, 2015, p. 247). And for this purpose, institutions should account to social, cultural and collective phenomena, rather than only focusing on the individual level (Wilhite et al., 2000). In fact, through their social interactions and interaction with available infrastructures, people create cultural conventions about how to use energy services, by either normalising and reinforcing current practices, or by stigmatising and challenging them (Hargreaves and Middlemiss, 2020).

Grounded in the Structuration Theory by Giddens (1984) that I have introduced in Section 2.6 and later conceptualisations by Schatzki (2021), Reckwitz (2002), and Warde (2005) among the others, Social Practice Theories (SPTs) posit that the social practices that drive energy demand have social, cultural and material histories. Understanding energy consumption thus requires to understand the spatial, temporal and social distribution of social practices that affect and shape it. One of the most influential conceptualisations of social practices is the one proposed by Shove, Pantzar, et al. (2012). It considers households' energy demand as the outcome of routinised activities, which in turn result from the combination of i) available materials (such as equipment or infrastructure), competences (namely skills and know-how), shared meanings (such as images, symbols, and collective conventions of normality setting social norms), ii) the way these components interact with each other, and iii) the way each practice interacts with other existing practices (Figure 2.7). Social practices thus result from the intertwining of such elements and cannot be reduced to any single component (Nash et al., 2017).

Social practices are namely the outcome of a set of shared social norms, material affordances, and individual and collective competences: depending on the way such elements are organised, they can either foster the daily reproduction of carbon intensive behaviours or, to the opposite, drive performance of low carbon behaviours. Particularly, practices are not static: they evolve over time and space and have their own "life", as well as social, cultural, and material histories. Therefore, in a given historical time, practice might differ depending on the specific context where they are carried on. However, practices are robust and resilient and therefore they cannot be changed quickly (Jalas et al., 2017).

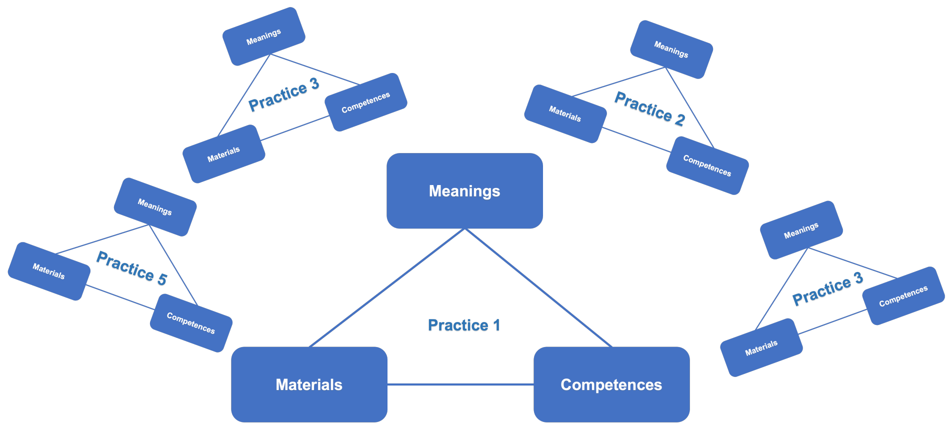

Figure 2.7: Schematic representation of social practices according to the conceptualisation by Shove, Pantzar, et al. (2012).

In this conceptualisation, individuals play a role as carriers of social practices, as they move through everyday life. The SPT literature in fact focuses on analysing and understanding processes of "recruitment" and "defection" of users to/from practices, namely the processes through which practices capture their carriers or how individuals defect from possible alternative practices (Shove, 2010b). Overall, individuals are crucial to the existence and characterisation of practices: if they learn new skills, respond to social meanings, or use novel infrastructures, they contribute to the evolution of practices, by performing them in different ways (Nash et al., 2017).

Although individuals are active agents in change processes, SPTs suggest to directly target and challenge practices themselves and the everyday situations that shape —and are shaped by— individual behaviours (Jensen et al., 2019). Active policy interventions to change social practices performed in a specific time and context can either aim at "recrafting" the elements of which practices are made, at "substituting" one practice with another one, or at changing how practices interlock or connect with one another (Spurling et al., 2013). In particular, Hargreaves and Middlemiss (2020) remark the importance of social relations, that drive the reproduction of practices by individuals. The authors notice that relationships with friends and family shape the way people engage with energy, particularly regarding when, where and how much energy to use.

For instance, by referring to the practice of cooking, Nash et al. (2017) have remarked that a change in practices does not simply correspond to a change in how some individuals cook their food. Rather, changes in practices imply that broader shifts in societal organisation and understandings of the concept of cooking take place. Remarking the importance of the network of social relations that every individual is engaged into, Hargreaves and Middlemiss (2020) have highlighted the importance of acting on such relations, which are essential to shape the way people "sustain or challenge cultural conventions of normal energy use, as well as how they respond to interventions" (p. 198).

In the last decade, policy interventions inspired by Social Practice Theories have been identified by many scholars as the way forward to address the climate crisis and favour the transition to a more sustainable society (Røpke, 2009; Shove, 2010a; Hargreaves, 2011; D. S. Evans et al., 2012; Moloney and Strengers, 2014; Sahakian and Wilhite, 2014; Shove, 2014; Kuijer and Bakker, 2015; Hampton and R. Adams, 2018; Labanca and Bertoldi, 2018; Valentine et al., 2019; Labanca, Pereira, et al., 2020; Della Valle and Bertoldi, 2022). However, as noted by Sahakian, Rau, et al. (2021), large-scale change initiatives grounded in social practice perspectives are still scarce: SPTs have in fact still mostly been used to understand a given system configuration, by identifying how practices have emerged, evolved and settled themselves across time and space. Innovative interventions are nonetheless being performed, mostly within participatory processes, during which first collective reflections on given social practices are performed, by identifying the practice components and possible change points to reconfigure them, and then attempts to implement such changes in real life take place, with the aim of learning from both successes and failures. Sahakian, Rau, et al. (2021) offer an overview of interventions informed by SPTs. So far, outcomes of such interventions have mostly been assessed through qualitative approaches, focusing on small sample sizes. When quantitative approaches were adopted, descriptive statistics have mostly been used, usually in before-after research designs and nearly always without control groups.

Scholars relying on social theories of practice argue that the former are incommensurably different from behavioural theories (the two theoretical approaches are like "chalk and cheese", Shove, 2010a) and reject the possibility to fruitfully integrate them. However, integrations have been attempted, such as for instance the one by Hess, I. Schubert, et al. (2022). Inspired by Spurling et al. (2013), the authors performed a quantitative analysis on the outcomes of a behaviour change randomised controlled trial addressing household routinised behaviours on washing, standby, and cooking, by estimating its energy saving effects on each social practice component. The Social Power Plus app, that constitutes one of my case studies, consists in another attempt to integrate behavioural and social practice theories. I will come back to SPTs in the related chapter as well as in the Discussion chapter.

2.9 Households

Since the Seventies, when energy conservation goals started to appear in research, policy-making, and more broadly in the public discourse, households have been identified as a key target group, due to their contribution to the consumption of energy (heating of spaces and of water, air conditioning, refrigeration, lighting, cooking, use of computers and digital appliances, etc.) and of natural resources for the satisfaction of their daily needs. Indeed, since the last decade, households have been increasingly been regarded as active participants to the energy and climate transition and are called to a responsibility to act as change agents (Naus et al., 2015): they are seen as a source of innovation and as active transition co-managers.

Nevertheless, a recent analysis on households' contribution to sustainability transitions has remarked that the conceptualisation of households is mostly implicit and taken for granted. Raven, Reynolds, et al. (2021) have in fact acknowledged that the meaning of "household" differs depending on the overarching theoretical framework and approach used. They have identified two main conceptualisation categories, respectively referring to them as the "closed-box" and the "open-box" approach.

The closed-box approach, usually adopted by research grounded in techno-economic domains, considers the household as a physical unit where infrastructures operate, for which data on energy use/carbon emissions can be collected and for which only aggregate attributes are available, mostly collected via surveys. In this case, intra-household dynamics and the role of material settings and social context are rarely considered. Rather, analyses consider energy and resource needs by the households, though they usually overlook the processes that generate them, which can either be internal to the household itself or be related with interactions between the household and its social and material context.

The open-box approach, instead, considers the household as a social unit, within which internal dynamics between its members occur (possibly also including conflicts). The open-box household also widely interacts with the context in which it is embedded, that is both an enabler and a constraint to the behaviour of the household itself. This approach highly values relationships within and between households, such as possibilities to share experiences and help each other to reduce consumption. It acknowledges that the core social relations that households undertake with family and friends, which are based on emotions, care, intimacy, love and friendship, have important implications for energy demand (Hargreaves and Middlemiss, 2020). This approach mostly performs qualitative analyses, focusing on the discourses and interactions within the household members and between them and contextual elements such as other households, infrastructures, or other inter-linked practices.

The analysis by Raven, Reynolds, et al. (2021) concludes with a call for future research to explore the combination between open- and closed-box approaches, which could support an improved understanding of household's agency in energy transitions.

2.10 Conclusions

In this chapter I introduced key concepts that are relevant for framing, guiding, and discussing the analyses I perform on the three persuasive apps aimed at supporting the energy and climate transitions. The three app cases are all grounded in a broad socio-technical approach and operate at the micro-level of individual households. By adopting the conceptualisation of transitions by the Multi-Level Perspective, they could be regarded as enablers of niche spaces in which groups of individuals get support towards the transition. They however have no ambition to trigger and sustain the broad system level change that is required for sustainability transitions to occur. Rather, they have to be considered a supportive element, that can potentially favour, speed up, and enhance the transition.

The three cases of persuasive apps have also been designed with a strong sustainability approach in mind, as their primary goal is the reduction in the overall amount of energy consumed by households (which also corresponds to reducing the related carbon emissions), in an energy sufficiency framework. The reduction in consumption stemming from app use also implies a reduction in the monetary costs that households meet to satisfy their energy needs, thus also producing social benefits.

Considering the Global North context in which these apps are developed and tested, the risk of critical social consequences is rather limited: at least in the case of Switzerland, situations of energy poverty, which might be enhanced by use of such apps, are currently rather limited. The trend is about over-heating and over-consumption of energy in residential buildings (Maxim et al., 2016; Bertho et al., 2021) —which is why these persuasive apps have been developed. This implies that the strong sustainability approach adopted by the three persuasive apps is not in conflict with society-related sustainability factors —to the opposite, apps are potentially beneficial both to the environment and to society.

Regarding the conceptualisation of behaviour, as I mentioned in Section 2.5, the three apps refer to the perspectives of reflective and automatic behaviour, and their features are mostly grounded in the Theory of Planned Behaviour, the Transtheoretical model of behaviour change, and the Self-Determination Theory, within the broader Captology framework, as they allow administration of behaviour change interventions through Human-Computer-Interactions. According to the Theory of Planned Behaviour, the apps support the increase in perceived behavioural control through the provision of recommendations, tips, and feedback on individual's progress towards change. They also support the evolution of individual's subjective norms, as they strengthen the belief that key societal actors (those who initiate and promote app's use) call for a change in consumption. Furthermore, following the Transtheoretical model, the apps support progress from one behaviour change stage to the next one, by enabling activation of the specific processes identified by authors of the Transtheoretical model themselves. In the chapters focusing on each app case I will specify which processes are leveraged and which app features and techniques are exploited for this purpose.

All three apps also adopt a gamification approach and leverage typical game design elements — though with differences and peculiarities, that I will introduce in the chapters devoted to each app case. Gamification approaches are grounded on the Self-Determination Theory and aim at increasing feelings of autonomy, competence, and relatedness. Through individual goal setting and feedback on progress towards goal achievement, which are accompanied by badges and possibly also by tangible rewards, the apps' features enhance the creation of feelings of competence and autonomy, that support integrated regulation and the increase in individual's motivation for change.

Similarly, features favouring interaction with other peers or comparisons with them increase the feeling of relatedness, which further enhance integrated regulation.

Besides these key theoretical references, the Social Power Plus app also marginally draws from the everyday behaviour perspective, which has connections with Social Practice Theories. By offering challenges aimed at changing the way key energy-consuming household practices are daily reproduced, it promotes a re-crafting of their meaning, material, and competence components.

Finally, regarding households, the closed-box conceptualisation is definitely well-suited to the approaches by the enCompass and Social Power app cases. For Social Power Plus, instead, a "mixed" conceptualisation would fit better, as in fact the Social Power Plus app tackles the call by Raven, Reynolds, et al. (2021), by adding a social-sharing component to the individual-based behaviour change features that are considered by the other two apps. Even in the case of Social Power Plus, however, only interactions external to the household are considered, namely interactions with other peer households of the same community of app users. Due to the way it has been designed, analysis of internal dynamics within the household are not possible for Social Power Plus either: even in this third case, the "household box" is still partially closed.

3 Related Work

> *As for the future, your task is not to foresee it, but to enable it.*
>
> — **Antoine de Saint Exupery**
> Writer

This chapter offers an overview of policy interventions that have been proposed to trigger and support behaviour change processes, focusing on empirical work aimed at reducing residential energy consumption and the related carbon emissions in households. I perform an in-depth analysis of policy interventions leveraging smart metering devices that allow high frequency, automatic collection of energy consumption data at the household level and digital tools (mostly apps, but also web-platforms), that also allow bi-directional interaction possibilities with their users, thus enabling novel intervention possibilities. Specifically, I perform a broad overview of empirical behaviour change research that, over the last three decades, has emerged in the domains of social psychology, behavioural economics, and computer science.

I first explore behaviour change intervention techniques, by introducing the taxonomies that have been developed to classify them and then by summarising their effects as well as recommendations for practical interventions that have been offered by previous literature. I then focus on specific interventions providing information feedback, using social influence techniques, leveraging digital nudges and, more broadly, persuasive apps. These types of intervention are in fact key to the three case studies I address in this dissertation.

I summarise the learnings I draw from these analyses by means of a narrative literature review, which allows me to identify the effectiveness of previous interventions, as well as the main limitations affecting them and the still open questions to be addressed in novel research. The chapter concludes with methodological insights about empirical research performed so far, that supported me in the design of the policy evaluation methodologies to adopt in each of the three case studies of the dissertation, with the aim of obtaining scientifically strict and reliable results.

3.1 Behaviour change interventions

Energy conservation in households has been a research topic since the seventies of the last century: while in the beginning it was related with the oil crisis and price shocks, since the late noughties it has been inspired by global challenges such as the climate crisis.

In both cases, it aims at understanding how residential energy consumption behaviour change can fruitfully be promoted.

The broad body of research performed so far shows that interventions aimed at reducing energy consumption in households can either target the context in which individual decisions are made (e.g. implementing infrastructural and technology developments or introducing new regulations or financial incentives) or aim at a voluntary change in households' individual decisions. Specifically, Steg and Vlek (2009), have classified interventions aimed at fostering pro-environmental behaviours as either *informational* strategies, which aim at changing motivations, perceptions, attitudes, habits and norms, or *structural* strategies, which aim at changing the circumstances under which behavioural choices are made, by acting on physical infrastructure, technical facilities, availability of products or services with given characteristics or at given prices.

The provision of information can both make individuals aware of the need and possible ways to reduce household energy use and increase their motivation to conserve energy, overcoming possible negative implications on status, comfort and effort that this may entail (Steg, 2008). In interventions aimed at providing information, Steg (2008) suggests to explicitly integrate normative factors: if a given energy saving behaviour is only performed for hedonic or cost reasons, as soon as it loses attractiveness or cost-effectiveness, individuals would stop performing it. Instead, if a given behaviour is performed for normative reasons, behaviour is expected to be more robust against contextual changes. Furthermore, Steg (2008) argues about the need for structural interventions, as without them feasible alternatives aimed at performing a given behaviour might be precluded to the individual. Acting on contextual factors ensures that the needed products and services, infrastructures, economic factors, and cultural norms, are available and can enable implementation of a given behaviour. Or, if such elements are already available, improving them helps to reduce the perception of effort, cost, or discomfort that is usually associated with sustainable behaviour.

More recently, focusing on climate change mitigation, Nielsen, Linden, et al. (2020) have further detailed the classification of policy interventions, by identifying four categories: interventions can be aimed at *altering the decision environments* (manipulation of the choice architecture, as in nudge approaches), can *appeal to norms*, can *provide information*, and can aim at *improving the skills* required to perform the desired behaviour. The authors have also remarked that micro-level interventions focusing on individuals would benefit by being integrated with broader behavioural initiatives that aim at addressing the "cognitive heuristics and biases of choices, values and norms, individual habits, as well as individual, social or political processes" (p. 1613).

The need for accounting for the context and the constraints it imposes on available possibilities for action also emerges from the COM-B behaviour model that I briefly introduced in Section 2.6. To use the COM-B model to inform contents of behaviour change interventions, the authors suggest to first identify capability, opportunity and motivation factors with respect to the desired target behaviour. Then, their "behaviour change wheel" can be used to identify the specific set of intervention functions that can best affect those factors, by picking from the following list: *education*, *persuasion*, *incentivisation*, *coercion*, *training*, *restriction*, *environmental restructuring*, *modelling*, or *enablement* (Michie, Van Stralen, et al., 2011).

Another classification of interventions specifically aimed at supporting the transition to a low carbon society has been recently proposed by Whitmarsh et al. (2021). They identify four main categories of behaviour change interventions (*economic*, *structural*, *information provision*, and *social influence*), each of which can either:

- target individual decision-making (*downstream*) or the context in which decisions are made (*upstream*);
- provide/improve options (*pull*) or remove options (*push*);
- exploit automatic processes (*nudge* or *changing choice architectures*) or rely on intentional and deliberative processes (*citizen assemblies*).

The authors recommend that interventions combine multiple of the above approaches. Particularly, they argue that nudge interventions aimed at exploiting automatic processes may not be sufficient to produce the deep societal transformation required by the climate crisis, as they focus on single, specific behaviours only.

Indeed, scholars increasingly suggest to develop interventions that leverage different strategies, that can synergistically be beneficial to each other. A valuable framework that can operatively support the design of applied interventions is the one developed by White et al. (2019), who, based on an extensive review of 320 articles related with sustainable consumer behaviour, proposed the "SHIFT" acronym to encompass the broad set of interventions that can favour behaviour change: *Social influence*, *Habit formation*, *Individual self*, *Feelings and cognition*, and *Tangibility*. Depending on their specific goals, targets, and contexts, practitioners are invited to pick from these strategies to inform their behaviour change interventions. To provide practical support in the implementation of such strategies, taxonomies of behaviour change techniques have been developed.

3.1.1 Taxonomies of behaviour change techniques

In a paper dealing with traffic and use of safety belts, Geller et al. (1990) have proposed a "taxonomy for behaviour change interventions", which summarises a list of twenty-four different techniques to motivate behaviour change. About two-thirds of them can be classified as *antecedent* techniques, since they need to be implemented *before* the behaviour occurs (such as the provision of information or energy saving tips), and one third of them can be classified as *consequence* techniques, since they are activated *after* the behaviour has been performed, with the aim of influencing future behaviours (such as the provision of feedback on households' amounts of saved or consumed energy).

In terms of practical activities to be performed on the field, Schultz (1999) has summarised more than twenty years of behavioural interventions by social psychologists on recycling and waste. He reports that most frequent techniques consist in the *provision* *of information*, *prompts*, *public commitment*, *normative influence*, *goal setting*, *removing barriers*, *rewards*, and *feedback*. A few years later, based on an analysis of 38 studies aimed at increasing energy saving in households, Abrahamse, Steg, et al. (2005) have detailed the antecedent/consequent taxonomy by Geller et al. (1990). They have identified *commitment*, *goal setting*, *information*, and *modelling*, as the antecedent techniques aimed at influencing one or more underlying behaviour determinants prior to the performance of behaviour, and the provision of *feedback* and *incentives* or *rewards*, as the consequence techniques focusing on the positive or negative effects of a given behaviour (Table 3.1).

Table 3.1: Types of intervention techniques, according to the review of energy conservation interventions by Abrahamse, Steg, et al. (2005), informed by the taxonomy by Geller et al. (1990).

Type	Name	Definition	Behavioural determinant
Antecedent intervention technique	Commitment	An oral or written pledge or promise to change behaviour.	If the pledge is made to oneself: personal norm; if public: social norms.
	Goal setting	Giving households a reference point.	Personal norm and social-value individual orientation.
	Information	Provision of (tailored) information about energy-related problems and/or possible solutions.	Awareness and knowledge.
	Modelling	Provision of (tailored) examples of recommended behaviour.	Awareness and knowledge; Social norms.
Consequence intervention technique	Feedback	Provide information about one's energy consumption (individual feedback) or savings and/or the performances of others (comparative feedback).	If individual: associate outcomes with one's own behaviour; if comparative: competition, social comparison, social pressure (social norms).
	Incentive/ Reward	(Monetary) prizes, either direct or indirect, e.g. in the fashion of tax credits.	Extrinsic motivation, based on reinforcement of desired behaviours.

Later, based on an extensive review of experimental interventions to promote pro-environmental behaviour (87 articles presenting 253 interventions), Osbaldiston and Schott (2012) have identified ten behaviour change intervention techniques:

- *Making it easy*: changing situational conditions to make behaviour easier to do, such as providing low-flow shower heads to conserve water and energy;
- *Prompts*: providing non-informational reminders that focus on when the next specific action is performed, such as "turn the lights off when you leave the room";
- *Justifications*: providing the reasons for performing a specific behaviour (declarative information), such reducing energy consumption for fossil-based heating for climate reasons;
- *Instructions*: indicating how to perform a specific behaviour (procedural information), such as using the blinds to keep rooms cooler;

- *Feedback*: providing information about the extent to which a behaviour has been performed in an earlier time-frame, such as through monthly electricity bills, so that individuals learn what they did in the past and accordingly adjust their behaviour in the future;
- *Rewards (or incentives)*: offering people any kind of monetary gain for participating in an activity, such as cash coupons or lottery prizes;
- *Social modelling*: arranging demonstrations or discussions in which the initiators indicate that they personally engage in a certain behaviour;
- *Cognitive dissonance*: accessing pre-existing beliefs or attitudes, making them salient, and attempting to make people behave in ways that are consistent with those beliefs, to reduce the dissonance;
- *Commitments*: asking people to make verbal or written commitment to engage in a certain behaviour, such as by signing a pledge card;
- *Setting goals*: asking people to aim for a pre-determined goal, such as reducing their electricity consumption by 20%.

According to their meta-analysis, social modelling and commitment emerged as the most effective treatments for household energy conservation.

Also Allcott and Mullainathan (2010) have provided a specific classification of behavioural techniques aimed at reducing greenhouse gas emissions and at favouring energy efficiency. They have argued that non-price interventions can be as powerful in changing consumer choices as price-based interventions (e.g. taxes, emissions trading programmes, diffusion of new energy-efficient technologies). Considering their high cost-effectiveness, they have therefore suggested that non-price interventions become an integral component in climate change policies. Overall, Allcott and Mullainathan (2010) have identified the following categories of behavioural techniques:

- *Framing and psychological cues*: as the marketing industry has shown, psychological cues can be used to increase the demands for products and services. The way energy-consumption related information is presented to the consumers can affect the overall amounts they consume. A re-design of the information conveyed via the energy bill could, for instance, support energy saving by household customers;
- *Commitment devices*: individuals frequently acknowledge that performing certain actions would be useful for themselves, though they tend to postpone and delay. To counteract this phenomenon, individuals could be provided with support for locking themselves into performing actions they would be willing to perform but actually tend to procrastinate. This implies providing them with a guide and external entity, responsible for guaranteeing they stick to an agreed upon plan;
- *Default options*: in many choice areas, people rarely switch away from the default option they have been offered. This could be due to procrastination, could be an "endowment effect" (namely they prefer any option currently available to them) or be due to the (non-monetary) cost of changing option. In any case, setting the

default option as the most favourable one to the climate crisis (such as for instance automatically enrolling customers to the purchase of renewable-based electricity instead of the common electricity mix, while still offering an opt-out option by a simple pre-checked box) would produce tangible savings;

- *Social norms*: providing information on the behaviour of other people is a powerful behaviour change leverage. Individuals do not necessarily conform to other individuals' behaviour because they explicitly approve them; they might in fact simply feel reassured by such a conformity.
- *Implementation intentions*: while people's attitudes are quite easy to change, changing their actions is more difficult. An "attitude-behaviour gap" has in fact been clearly shown in energy and, more broadly, pro-environmental behaviour literature (Kollmuss and Agyeman, 2002; Peattie, 2010; Valkila and Saari, 2013; ElHaffar et al., 2020). A strategy to support concrete action could be to induce people to map out their intentions and the detailed plan on how to turn them into reality;
- *Exploiting nonlinear demand curves*: energy utility companies sometimes provide their customers with rebates for the purchase of energy efficient appliances. Since the demand curve for such appliances is non-linear, experiments could be used to get a precise understanding of the actual shape of the demand curve for energy efficient appliances, and the optimal amount of rebate that produces the largest purchase of efficient appliances. This would both guarantee an increase in energy efficiency and the effective expenditure of the utility's money.

A comprehensive and exhaustive theory-linked taxonomy of twenty-six behaviour change techniques was also proposed by Abraham and Michie (2008) with the aim of favouring effectiveness, standardisation and replicability of behaviour change interventions. Even though the above-mentioned more recent classifications of intervention strategies have been proposed, some of which also explicitly focus on the energy and climate transitions, I appreciate this taxonomy for being concise, clear, simple, and complete, which makes it well-suited as an operational guide to support the practical definition of interventions to be developed in real-life interventions. Therefore, I integrally report it here (Table 3.2) and will later refer to it in order to describe the characteristics of the app-based interventions I analyse in my three case studies.

Table 3.2: The taxonomy of behaviour change techniques used in interventions proposed by (Abraham and Michie, 2008) (integrally reported from their work).

Technique	Definition
1. Provide information about behaviour-health link	General information about behavioural risk, for example, susceptibility to poor health outcomes or mortality risk in relation to the behaviour.
2. Provide information on consequences	Information about the benefits and costs of action or inaction, focusing on what will happen if the person does or does not perform the behaviour.
3. Provide information about others' approval	Information about what others think about the person's behaviour and whether others will approve or disapprove of any proposed behaviour change.
4. Prompt intention formation	Encouraging the person to decide to act or set a general goal, for example, to make a behavioural resolution such as "I will take more exercise next week".
5. Prompt barrier identification	Identify barriers to performing the behaviour and plan ways of overcoming them.
6. Provide general encouragement	Praising or rewarding the person for effort or performance without this being contingent on specified behaviours or standards of performance.
7. Set graded tasks	Set easy tasks, and increase difficulty until target behaviour is performed.
8. Provide instruction	Telling the person how to perform a behaviour and/or preparatory behaviours.
9. Model or demonstrate the behaviour	An expert shows the person how to correctly perform a behaviour, for example, in class or on video.
10. Prompt specific goal setting	Involves detailed planning of what the person will do, including a definition of the behaviour specifying frequency, intensity, or duration and specification of at least one context, that is, where, when, how, or with whom.
11. Prompt review of behavioural goals	Review and/or reconsideration of previously set goals or intentions.
12. Prompt self-monitoring of behaviour	The person is asked to keep a record of specified behaviour(s) (e.g., in a diary).
13. Provide feedback on performance	Providing data about recorded behaviour or evaluating performance in relation to a set standard or others' performance, i.e., the person received feedback on their behaviour.
14. Provide contingent rewards	Praise, encouragement, or material rewards that are explicitly linked to the achievement of specified behaviours.
15. Teach to use prompts or cues	Teach the person to identify environmental cues that can be used to remind them to perform a behaviour, including times of day or elements of contexts.
16. Agree on behavioural contract	Agreement (e.g., signing) of a contract specifying behaviour to be performed so that there is a written record of the person's resolution witnessed by another.
17. Prompt practice	Prompt the person to rehearse and repeat the behaviour or preparatory behaviours.
18. Use follow-up prompts	Contacting the person again after the main part of the intervention is complete.
19. Provide opportunities for social comparison	Facilitate observation of non-expert others' performance for example, in a group class or using video or case study.
20. Plan social support or social change	Prompting consideration of how others could change their behaviour to offer the person help or social support, including "buddy" systems and/or providing social support.
21. Prompt identification as a role model	Indicating how the person may be an example to others and influence their behaviour or provide an opportunity for the person to set a good example.
22. Prompt self-talk	Encourage use of self-instruction and self-encouragement to support action.
23. Relapse prevention	Following initial change, help identify situations likely to result in readopting risk behaviours or failure to maintain new behaviours and help plan to avoid or manage these situations.
24. Stress management	May involve a variety of specific techniques (e.g., progressive relaxation) that do not target the behaviour but seek to reduce anxiety and stress.
25. Motivational interviewing	Prompting the person to provide self-motivating statements and evaluations of their own behaviour to minimise resistance to change.
26. Time management	Helping the person make time for the behaviour (e.g., to fit it into a daily schedule).

3.1.2 Effectiveness of behaviour change techniques

Many reviews of behavioural interventions in the energy, climate and, more broadly, environmental domain, have been performed in the last two decades, to explore effectiveness of intervention techniques and thus provide evidence-based suggestions for policy-making. Their outcomes, however, are not fully coherent with each other and, specifically for household energy consumption, there is still a lack of clarity on which behavioural intervention types and techniques are (most) effective in supporting change.

One of the first reviews about energy saving interventions in households is the one by Abrahamse, Steg, et al. (2005), who considered 38 energy-saving interventions targeting households. They found that the effect sizes[1] of early energy conservation interventions, measured through the Cohen's d index, are quite small —and in some cases they cannot even be computed, as the related studies did not report the needed information. Their results indicate that the provision of information does not necessarily lead to behaviour change, and that rewards are effective but short-lived. Feedback on consumption, instead, is effective, especially when it is frequent. Anyway, in general they found that effect sizes decrease with the rigour of the study, which suggests presence of methodological pitfalls in intervention evaluation processes.

Osbaldiston and Schott (2012) also performed a meta-analysis of 87 experimental interventions aimed at favouring pro-environmental behaviours, based on observed behaviours (instead of self-reported data). They found that social modelling and commitment emerged as the most effective treatments for household energy conservation.

Another rigorous review of behavioural interventions aimed at energy saving was performed by Andor and K. M. Fels (2018), who explicitly focused on interventions performed under strict experimental conditions, thus allowing to both estimate causal effects as well as to measure their effect sizes. Specifically, the authors analysed 44 studies either performed via randomised controlled trials, matching studies, Difference-in-Differences, instrumental variable estimation, and regression discontinuity design, or that controlled for self-selection with other methodologies. They grouped the related interventions in the following four categories: *social comparison*, *commitment*, *goal setting*, and *labelling* (which refers to the provision of information on the level of energy usage by appliances or on compliance with energy efficiency standards). Based on the outcomes of their systematic review, the authors concluded that all types of intervention have the potential to significantly reduce household energy consumption. However, while for social comparison the reduction results are more clear, many studies intervening through commitment and goal setting did not show significant effects. The authors assumed this might either be because the treatments actually have no effect or because the small samples that were used did not reach sufficient statistical power. Studies that indicated significant effects, however, reported conservation potentials of about 10%.

[1] Effect size represents the number of standard deviation units by which the intervention group outperforms the control group on a certain outcome variable (e.g. gas or electricity savings).

Also Nisa et al. (2019) have performed a large-scale meta-analysis of randomised controlled trials specifically aimed at promoting household action for the mitigation of climate change. They focused on interventions that do not involve economic (dis)incentives or regulations and adopt highly rigorous experimental designs to estimate their effectiveness: they only considered RCTs, excluding observational or quasi-experimental designs, based on measured changes in behaviour (rather than self-reported behavioural declarations or intentions). Overall, their analysis encompassed 83 interventions performed between 1976 and 2017 in six household domains: energy consumption at home, transportation, consumption of animal food, food waste, water consumption and recycling.

According to their results, behavioural interventions acting alone provide limited benefits to mitigate climate change, as households exhibit resistance when only targeted by behavioural interventions. They in fact estimate at 6.6% the mean probability of behavioural interventions to produce behaviour change that mitigates climate change. In more details, their analysis indicates that the most promising types of intervention in terms of effectiveness are choice architecture (nudges) and social comparison messages. Instead, even though they are the most common used strategies, information based interventions have very limited impact: their probability to produce positive behaviour changes is estimated to 3.4%. Nisa et al. (2019) therefore suggested that, instead of developing interventions based on behavioural strategies only, researchers should focus on interactions between behavioural and non-behavioural strategies, in an inter-disciplinary perspective aimed at understanding if, once integrated with other types of intervention, behavioural interventions can increase adoption of actions with a high potential to reduce carbon emissions. Furthermore, Nisa et al. (2019) found that, once the interventions have concluded, there is no evidence that they produce lasting changes: even though only a limited subset of the studies they analysed include follow-up data, the available data shows on average null effect.

Finally, also Khanna et al. (2021) have performed a meta-analysis on 122 intervention studies performed throughout the world between mid Seventies and 2020, with the aim of supporting behaviour change in residential building energy demand. They classified the interventions in five categories: *monetary incentives* (peak/seasonal/time-of-use pricing, rewards, rebates), *information* (home audits, tips, reminders), *feedback* (individual historical consumption), *social norms* (home energy reports and comparison with other households), and *motivation* (which includes commitment and goal setting as well as gamification). They assessed the relative change in energy consumption due to the interventions and found that monetary incentives have the highest average effect sizes; motivation and social norms instead have the lowest sizes. Moreover, they found that:

- interventions that include gamification or other commitment techniques, together with goal setting activities, report higher effect sizes on average;
- studies in which households were not required to opt into the intervention were characterised by lower effect sizes. This can be explained by considering that

households that self-select themselves into an energy conservation intervention have higher motivation than average households to save energy;

- interventions with longer treatment duration tend to find smaller effects on average —though the authors noted that the median duration of the analysed interventions is 12 weeks, and state the need for long-term trials and for repeated follow-up measures of impact after the end of the interventions, in order to assess the persistence of the effects after the treatment has ended.

Overall, based on their extensive meta-analysis, Khanna et al. (2021) concluded that synergistic packages of different, well-integrated intervention types should be favoured, in order to increase their energy and carbon saving impact.

3.1.3 Recommendations for behavioural interventions

Independently on the specific behavioural technique(s) chosen, Nisa et al. (2019) recommended that it is essential to focus interventions on behaviours with higher climate change mitigation potential. It is also important to enrol naive participants, instead of self-selected ones. In fact, it is likely that people who accept to take part in environmental studies are more interested in climate change mitigation than people that decline such invitations. Furthermore, the authors remark that participants who provide their consensus to join an intervention are more aware that their behaviour is monitored, which might imply they tend to conform more to the expected behaviour (Hawthorne effect[2]). Policy-making based on estimates of effects on self-selected participants to inform policy-makers might thus create wrong expectation of effects, that are not met when the interventions are extended to the whole population outside intervention settings.

Also Nielsen, Linden, et al. (2020) have identified key behaviourally informed principles that should be accounted for when designing public policies and private sector initiatives, independently on the specific technique used:

- address *non-financial barriers* to action, for instance by limiting use of financial incentives, which, as extrinsic motivational factors, may have the detrimental effect of reducing people's intrinsic motivation to act;
- exploit marketing opportunities via *informal social networks*;
- target *early adopters*, who can diffuse the innovation by example;
- provide information from *credible sources*, if useful by relying on intermediaries, at the times and places where choices are being made;
- ensure that interventions are *simple* and try to make low-emission alternative behaviours more convenient than baseline ones, by *setting the default choice*.

Regarding the specific characteristics of interventions, Nielsen, Linden, et al. (2020) have suggested to tailor them to:

[2]The Hawthorne effect is a "change in behaviour as a motivational response to the interest, care, or attention received through observation and assessment" (Sedgwick and Greenwood, 2015, p. 1).

- the *characteristics of the target behaviour* and the underlying choice processes: consider differences between *i)* behaviours that can be performed quickly and at low cost, which are driven by habits and tend to be automatic, less conscious processes, though need to be repeated over time to accumulate impact (such as lowering thermostat settings or carpooling to work), and *ii)* other high impact behaviours, which instead consist of single actions, that may require large financial investments, but that also have larger and long-term impact (such as the decisions to buy a new fridge or house retrofitting);
- the *context* in which the new behaviour has to be performed: consider that individuals differ in their opportunities and abilities to change their behaviour, depending on the available alternatives, their financial capability, and the existence of regulations hindering or favouring a given behaviour;
- the *specific actors* whose behaviour has to be changed: explicitly account for individuals' multiple and often conflicting goals, their cognitive or financial resources, their priorities, attitudes and values, as well as their norms.

Similar recommendations, though informed by different theoretical backgrounds, are proposed by Breukers et al. (2015). Specifically dealing with interventions targeting the transition to a decentralised energy system, and informed by socio-technical perspectives, the authors argued that households' wishes and needs are seldom taken into account in the design of the related interventions. They suggested that, instead of focusing on technology or price incentive mechanisms, households and their lifestyle characteristics should be put at the centre of the design of the interventions.

This suggests taking households as a starting point in the design of the intervention, by exploring their behaviours, needs, and motivations, and tailoring interventions aimed at reducing energy consumption on such factors. Furthermore, Breukers et al. (2015) remarked the need for accounting for the social and physical environments in which behaviours are embedded, instead of focusing uniquely on individual behaviour. Also, they argued that interventions mostly based on price signals, such as for example those relying on electricity dynamic pricing, critically ignore motivational factors other than the monetary ones, such pro-environmental or pro-social attitudes, or also comfort factors. Finally, by lacking reference to social norms, such interventions are at risk of only producing short-term effects. Or, if long-lasting results are obtained, the risk is that rebound effects appear, due to the increase in other energy-consuming activities, that are not constrained since no pro-social or environmental values are involved.

Frederiks, Stenner, and Hobman (2015b) and Andor and K. M. Fels (2018) have also remarked the need for documenting the costs of the interventions and discussing them in cost-benefit analyses on treatments' effectiveness, also in comparison with traditional policy instruments, such as taxes or bans: knowing about cost-effectiveness is in fact crucial to help policy-makers to choose between possible interventions aimed at supporting the energy and climate transitions.

The same authors also remarked the need for research designs that allow for causal inference. Similarly, Nisa et al. (2019) called for use of RCTs, interrupted time series, instrumental variables, regression discontinuity designs, Difference-in-Differences estimators, or matching techniques. In particular, instead of implementing a bundle of interventions in the same experiment, these authors called for the adoption of factorial designs, which allow to discover to what extent a particular intervention contributes to behaviour change. However, the same authors suggested that, once the specific effect of each intervention strategy has been clarified with proper research designs, it is more sensible to develop combinations of interventions, that leverage different strategies, which can synergistically be beneficial to each other. For instance, Nisa et al. (2019) suggested to start an intervention with a social comparison message, then add an information-based component and conclude with "an environmental appeal to save the planet" (p.10).

Having introduced the key intervention strategies that can be implemented in order to activate behaviour change and the related techniques, I now focus on specific literature dealing with the provision of information feedback and leveraging social influence, as these two strategies are fundamental within the app-based persuasive interventions that this dissertation focuses on.

3.2 Provision of information feedback

The provision of consumption feedback with the aim of changing behaviours in the environmental domain is usually referred to as "eco-feedback", namely the "feedback on individual or group behaviour with a goal of reducing environmental impact" (Froehlich et al., 2010, p. 1999). In the field of energy consumption, the effects of provision of information (eco-)feedback have first been explored in the late Seventies for electricity, under the rationale that the provision of information can modify or reinforce future actions to conserve electricity and thus respond to the oil crisis. As remarked for instance by Burgess and Nye (2008), Hargreaves, Nye, et al. (2010) or Karlin et al. (2015), energy consumption is in fact an abstract and non-sensory concept: people are not interested in energy use *per se*, but in the services it provides. Also, energy is invisible, silent and untouchable, and it cannot even be seen by one's peers. Further, it results from multiple behaviours and has low personal relevance to most individuals, since it is relatively inexpensive and its consumption causes no immediate personal harm.

Faruqui et al. (2017) cite early feedback studies performed during the 1970 oil crisis (Becker, 1978; McClelland and Cook, 1979; Seligman and Darley, 1977) and highlight key limitations that affected their overall effectiveness, such as small sample sizes and poor frequency and ways of providing feedback. Such interventions were supposed to enhance feelings of individual self-efficacy, namely the extent to which people feel capable to engage in a behaviour (Bandura, 1977). Later, a "Feedback Intervention Theory" (FIT) was also developed, according to which feedback interventions are conceptualised as "actions taken by (an) external agent(s) to provide information regarding some aspect(s) of one's task performance" (Kluger and DeNisi, 1996, p. 255). FIT posits that, if individuals acknowledge a discrepancy between performance on a task and an abstract standard (such as prior expectations, past performances, ideal goals, or personal or social norms), they might be lead to change their behaviour. By making such a discrepancy emerge, feedback "produces the motivation to eliminate the feedback-standard disparity" (Schultz, 1999, p. 4). Namely, feedback can influence behaviour by creating a connection between a given outcome, such as saving energy, and the behaviour that produces such outcome, such as turning the lights off.

Karlin et al. (2015) have however remarked that the simple provision of feedback is not enough to drive behaviour: the feedback must draw individuals' attention to a feedback-standard disparity that they have identified as self-relevant, for instance because it is related to a goal they have set for themselves. Also, feedback has to be provided at the right time, ideally when energy consumption occurs, and has to be customised to user characteristics, interests, and needs (Hermsen et al., 2016).

3.2.1 Feedback through smart meters and in-home-displays

At the time of early feedback interventions, feedback was provided manually, through printed reports, which were either mailed to recipients via traditional post or in some cases even manually door-to-door delivered. Novel feedback opportunities were made available by developments in information and communication technologies and by the availability of smart meters, namely electric meters that automatically provide the utility company with detailed information on energy consumption by the end-user, which can also be automatically processed and sent back to the end-users themselves via in-home-displays (IHDs, Bugden and Stedman, 2019).

Available since the early 2000s, and subject to large-scale roll-out plans since the late 2000s (Joachain and Klopfert, 2014), smart meters were expected to support the transition to renewable energies and enable more efficient, sufficient, and flexible energy consumption practices. They were in fact welcomed as promising devices to be integrated in behaviour change policy interventions, due their capability to inform and motivate change by the provision of high granularity and high frequency consumption feedback data, either directly or through in-home displays located in households' living rooms. Early studies by Darby et al. (2006) reported energy savings ranging between 5% and 15%. An extensive review by Fischer (2008) reported that feedback stimulates energy (and specifically, electricity) savings, usually between 5% and 12% (even though in some cases no savings were found).

Delmas et al. (2013) performed a meta-analysis of behavioural interventions conducted between 1975 and 2012, that specifically focused on information strategies targeting households. Some of those interventions already included use of smart meters, either directly or through in-home displays. The authors assumed that, due to a market failure, individuals lack the relevant information to engage in energy saving behaviours, and expected that the adoption of strategies aimed at providing information could support energy conservation behaviours: coherently with Feedback Information Theory, individual feedback was expected to help make energy consumption more salient and thus to increase the sense of relevance of taking energy conservation actions. Specifically, Delmas et al. (2013) classify information provision strategies in two broad categories:

- *price-based information strategies*, such as the provision of feedback on the cost of energy use, or the provision of monetary incentives, including direct monetary rewards for achievement of certain energy saving goals;
- *non-price information strategies*, such as the provision of energy saving tips, in-person home energy audits, individual feedback on one's own past energy consumption, or comparative feedback on the behaviour of other peer individuals.

Price-based strategies, such as the provision of feedback on the monetary costs of energy consumption, are activated under the assumption that, once households receive information on such costs, those who directly pay for their energy bill will enact energy conservation actions. However, the meta-analysis by Delmas et al. (2013) has shown that such strategies might lead to the opposite outcome: not only pecuniary feedback appears to have no effect on energy conservation behaviour; it also appears to drive the increase of energy consumption, resulting in a "rebound" effect. The authors explain the observed results by referring to a "licensing effect": if households learn that their energy expenditure is small, or that their potential monetary savings are small, they may feel entitled to fully benefit for energy use, since they are paying for it and energy use is perceived to increase comfort and convenience. According to Otto et al. (2014), rebound effects can only be avoided if people are intrinsically motivated to save energy and accept to give up the advantages they can obtain by spending elsewhere the money gained from savings. Moreover, in wealthy Western societies, monetary savings stemming from energy savings may be perceived as negligible and individuals may prefer to keep paying amounts that are still affordable to them, and to be guaranteed possibilities for consuming energy when they like.

Delmas et al. (2013) also found that strategies aimed at directly providing monetary incentives if households save energy, and at providing households with regular information about such incentives, might be counterproductive, since they might crowd-out altruistic or pro-social motivations. Such outcomes can be explained by the Motivation crowding theory by (Frey and Jegen, 2001), that I briefly introduced in Section 2.6.5: use of external regulation motivational factors such as tangible prizes favour an increase in individuals' intrinsic motivation and self-determination to change only if they are perceived to foster self-esteem and freedom to act.

3.2.2 Types of feedback

Karlin et al. (2015) remarked that the actual feedback impact depends on a number of factors, such as its frequency, granularity, the conveying medium, and the unit of measurement that is used. They performed a meta-analysis of 42 feedback intervention studies, exploring how effect sizes change on varying such factors:

- *frequency* was expected to help improve links between actions and consequences. However, the authors have not found evidence of increased effect when feedback frequency increases —and this may be because frequency of feedback provision does not coincide with frequency of feedback access by the target recipients (who may only access it occasionally, despite the high delivery frequency);
- *granularity*, namely the level of detail to which the feedback is offered. Provision of feedback that is broken-down at the level of appliances was expected to help connect energy consumption to specific actions, and therefore to be more effective. The authors however found no significant differences between different granularity strategies. They have argued that this may be as, even though users know where energy is consumed, they do not necessarily know how to reduce consumption, or have the needed competences to do it;
- *mediums* through which feedback is provided: feedback is more effective when it is provided through the most engaging mediums, such as computer softwares, rather than the least engaging ones, such as bills or paper cards;
- choice of the *units of measurement*: use of non-physical energy units, such as carbon emissions or financial spending, was expected to increase effect size, compared with physical units (kWh). Instead, the authors did not find this effect. They however remarked that the way in which feedback is presented to users can have an impact on the way it is perceived and interpreted, and thus can influence action.

Despite Fischer (2008) has argued that combining interventions risks to overload users with too much information, many meta-analyses have indicated that, to be effective, provision of consumption feedback must be accompanied by additional techniques, such as individual goal setting, historical comparison with one's own consumption, competition or collaboration in group challenges, or social comparison with peers (Darby, 2010; Allcott, 2011; Abrahamse and Steg, 2013). Particularly, the meta-analysis by Karlin et al. (2015) found that the combination of feedback with goal setting features increases the effectiveness of feedback. However, the authors also remarked that none of the analysed studies considered research designs allowing to detect the specific contribution by each single intervention strategy or technique, besides their combined effect.

3.2.3 Effectiveness of the provision of information feedback

A recent review of feedback-based interventions run between 1976 and 2019 indicates that realistic energy saving effects could range from 5% to 10%, on varying the characteristics of the households (especially, the age and the education level), the feedback frequency and provision type (via static energy bills or via automatic devices), the energy end-use (heating, electricity, or both), and the socio-economic context (Zangheri et al., 2019). Exploring differences in the methodologies adopted to estimate feedback effects, the meta-analysis by Delmas et al. (2013) shows that in less robust studies (those without control groups), average 10% energy savings were obtained, and in more robust studies (those with control groups or at least those controlling for household demographics and weather variables), 2% average savings were obtained. The authors thus argue that the 10% effects of less rigorous studies might be mostly due to self-selection bias or Hawthorne effect, which occurs when behaviour is changed due to the awareness of being observed. These results, much less interesting than those by the early feedback studies by Darby et al. (2006) and Fischer (2008), lead Delmas et al. (2013) to conclude that individual "informational feedback alone may be a necessary but not sufficient condition to produce conservation" (p. 735). Also, Delmas et al. (2013) conclude that low-involvement information-based strategies (e.g. tips) are not effective at reducing energy use, while high-involvement information-based strategies (e.g. energy audits) are effective. Similar conclusions are drawn by the rigorous meta-analysis by Karlin et al. (2015): feedback interventions produce statistically significant energy consumption reduction. However, the effect size is small: the average effect size of the interventions they analysed is 0.115 (or 0.071, if a random effect size model is considered).

Based on these findings, Buchanan et al. (2015) have brought the overall usefulness of feedback into question, remarking that feedback interventions were accompanied by limited evidence of effectiveness and arguing that such interventions failed to activate the user engagement that is needed for them to be effective. Their conclusions however are mostly based on IHDs, that present real-time, historical and cumulative consumption in energy and monetary terms. Specifically about IHDs, Wallenborn et al. (2011) also found that only households that were already interested in energy savings were interested in using IHDs. Other qualitative research by Pierce, Fan, et al. (2010) and Pierce, Schiano, et al. (2010) has also shown that households do not understand how IHDs work, the meaning of charts and figures they provide, and have limited time to interact with them. Also, users of IHDs were seen to lose interest in them after a few weeks of use: after an initial period of interest due to their novelty, feedback was seen to fade in the background of everyday life (Hargreaves, Nye, et al., 2010; Hargreaves, Nye, et al., 2013).

Furthermore, Buchanan et al. (2015) noted that, as feedback-induced energy savings were found to be small, also the related monetary savings are small. This implies that pre-existing households' energy saving motivation might be undermined by feedback, especially when energy services are highly valued by households and/or deeply embedded in everyday life (which corresponds to the above-mentioned licensing effect found by Delmas et al., 2013). Finally, the authors note that self-selection problems that characterised most of feedback studies so far might have led to over-estimate the effectiveness of feedback provided via IHDs (Hargreaves, Nye, et al., 2013): by running the same programmes on the wider population, effects might be even smaller.

For all these reasons, Buchanan et al. (2015) conclude that the provision of feedback via IHDs has no capability to reduce energy consumption by itself; rather, IHDs' success is dependent on user engagement. For this purpose, innovative forms of feedback need to be developed and tested, that allow actual engagement by households. This requires to go beyond the provision of cost and consumption information, which requires users's time and capability to be fully understood. According to Buchanan et al. (2015), effective feedback systems should allow users to relate the consumption information they provide with the energy consumption routines that are performed in the household and to identify concrete and viable energy saving actions that are available to household members. A similar conclusion is also drawn by Geelen et al. (2019), who analyse the outcome of the provision of energy consumption feedback collected via smart meters and provided to households via an app. In that case, the potential of apps to engage their users through interactive features was not fully exploited, as the app's features were mostly resembling those by IHDs. The authors in fact found no statistically significant behaviour change effects and called for the design of apps that, not only allow to monitor energy consumption, but can also guide their users with actionable and meaningful support, targeted to their own specific situation.

3.2.4 Feedback over time

Another open and controversial research issue refers to how long feedback should be provided for. Karlin et al. (2015) remark that, on the one hand, feedback is likely to be effective if it is only provided for a short duration (less than three months), namely until users still regard it as novel and interesting. On the other hand, providing feedback over a long time-span is likely to favour users becoming more familiar with it, learning how to reduce consumption, and ultimately activating automatic energy saving processes in response to feedback. Both phenomena can occur and therefore more research is needed to clarify which is predominant.

Another key open research issue regards whether feedback maintains it effect in the long-term or not. Only 5 of the 42 feedback studies considered by Karlin et al. (2015), in fact, tested for persistence of effects after the intervention had ceased. In those five studies, the effect size was higher during the intervention than during the follow-up period, however the difference between the two effect sizes was not statistically significant. Due to the limited amount of available evidence, Karlin et al. (2015) thus called for future research to collect data more often and for a longer period of time, both during and after the feedback intervention. The same recommendations stem from a systematic review of 72 interventions leveraging feedback from digital technology (smartphone apps, web-based platforms, In-Home-Displays or wearable devices) performed by Hermsen et al. (2016). Only a few of the studies they analysed consider longitudinal measurements and most of those them only measure the effects exactly at the end of the intervention period.

Among the exceptions, Ferraro et al. (2011) assessed the long-term effect of a residential water conservation intervention leveraging norm-based strategies: the provision of social comparisons produced a long-lasting impact until more than two years after the treatment was performed. In the energy domain, Hargreaves, Nye, et al. (2013) reported little long-term effect of a feedback electricity saving intervention leveraging households' IHDs. Encouraging long-term effects on electricity savings were instead Similar encouraging results were found in an energy saving intervention targeting university students living in a campus dormitory in South Korea, that were provided with normative feedback, delivered via email on a weekly basis, for a period of one full year (Anderson et al., 2017). Also Schleich et al. (2017), in a feedback intervention leveraging in-home electricity displays, found significant and time-persistent savings over an 11-month period.

3.3 Leveraging social influence

Research grounded in social psychology studies has widely explored if and how behaviour change, and especially change towards pro-environmental behaviours, can be motivated by leveraging social influence. As noted by Farrow et al. (2017), in fact, "it appears that what other people do and think matters a great deal to individuals" (p. 1).

Abrahamse and Steg (2013) have extensively studied interventions aimed at encouraging pro-environmental behaviours by means of social influence strategies. Following Forgas et al. (2001), they defined social influence as the phenomenon occurring when "an individual's thoughts, feelings or actions are influenced by other people or groups" (p. 1773). After Aarts and Dijksterhuis (2003), they have posited that social interactions, the observation of others, as well the provision of information about their behaviour, allow individuals to form opinions and beliefs about the way they should behave and about what is socially acceptable. As later specified by Farrow et al. (2017), social influence can in fact affect individuals in a variety of ways: people may wish to fit in (or stand out) social circles, avoid social disapproval, seek social esteem, or refer to the behaviour of others as a shortcut signalling the most effective way to behave. In any case, social influence tends to operate through fast, intuitive, and emotional mental heuristics.

Under this conceptual framework, Abrahamse and Steg (2013) have explored the potential mechanisms linking social influence to behaviour change and have identified three types of social influence principles: *social norms*, *social learning* and *social comparison*. Further detailing them, the same authors have identified six social influence techniques that can support behaviour change:

- *social norms information in feedback provision*: following Cialdini and Goldstein (2004), the authors broadly define social norms as "cues that help people make sense of social situations, especially those characterised by a high degree of uncertainty or ambiguity, in terms of how people are expected to behave" (p. 1774). People adhere to social norms to gain social approval or avoid social sanctions, so that other people will like them. Following the "Focus theory of normative conduct" (Cialdini, Reno, et al., 1990; Cialdini, Kallgren, et al., 1991), the authors argue as well that social norms can help guide behaviour, when they are made salient. It is important, however, to properly manage the saliency of a social norm, in order to obtain results in the desired direction. For instance, Abrahamse and Steg (2013) indicate that people have shown a tendency to litter in already littered contexts, even though the prevailing norm is against littering;
- *social comparison in feedback provision*: this technique consists in providing people with feedback about their own performance, compared with the performance by

other people. People in fact tend to compare themselves to others, in order to make sense of their own opinions and behaviours. Social comparison has been first theorised by Festinger (1954), as the process of considering information about one or more other people in relation to oneself. Two social comparison types are possible: an "upward social comparison" occurs when an individual compares herself with people doing better, to see herself in a more positive light. A "downward comparison" instead occurs when individuals compare with those who are worse off, to feel better about their own situation. In both cases, the "Social comparison theory" by Festinger states that social comparisons are more effective when they refer to individuals or groups that are perceived as more similar to the subject under analysis (similarity principle). Finally, note that such a social comparison feedback is different from feedback on social norms. The latter in fact does not include reference to the behaviour of the feedback recipient; rather, it simply refers to general social norms, such as stating that "75% of hotel guests reuse their towels";

- *feedback about group performance*: feedback can also be provided at the group level, for instance when individual contribution is not identified and feedback is only provided on total or average group performances. Such an approach is similar to the social norm one, since they both provide aggregate information about a group of people and what they are doing. However, group feedback includes the performances of the feedback recipient, while social norm feedback does not. Group feedback can be used to reflect a collective effort towards a shared goal, enhancing feelings of collective efficacy ("a group's shared belief in its conjoint capabilities to organise and execute the courses of action required to produce given levels of attainments", Bandura, 1977, p. 477) and also feeling of peer pressure to conform to the group norm. Indeed, group feedback can also provide support to the evolution of a social norm towards the performance of a target behaviour;
- *block leaders and social networks*: this technique exploits one or more volunteers (block leaders), who help spreading information about a given issue, through their social network. The stronger their ties with other members of their social network and the higher the perceived similarities between them and the social network members, the more likely the chance that block leaders can affect behaviour of other people. This technique thus builds on the capability to favour actual spreading of information through existing social networks, combined with linking and similarities principles;
- *modelling*: this technique refers to the use of confederates to show how a given behaviour should be performed. Namely, people engage in a new behaviour when they observe other people doing it. This thechnique builds on the "Social learning theory" (Bandura and Walters, 1977), which states that learning of new skills occurs in a social context. Note that an effective outcome of modelling can either be due to similarity, linking, or even to descriptive social norms;

- *public commitment making*: committing oneself to something refers to binding oneself to a certain opinion or behaviour (Kiesler, 1971). Once individuals make a pledge, then they are expected to stick to it, due to their need for consistency. Commitments that are made in public are especially effective, since individuals "are more likely to act in accordance with publicly held attitudes, as compared to privately held attitudes" (Abrahamse and Steg, 2013, p. 1774). Moreover, public commitment may also encourage behaviour change through social pressure to stick to the commitment, as individuals aim at protecting their own public image.

3.3.1 Social norms

Among the above techniques, the use of feedback to activate norms has been widely studied and offers opportunities for in-depth discussion. Norms are generally regarded as shared rules of conduct that are partly sustained by approval and disapproval by other people (Farrow et al., 2017). They are unwritten and implicit, which makes them different from regulations and other explicitly codified social frameworks. Indeed, norms can be either personal or social. Personal norms refer to the feeling of obligation to act in a particular way in specific situations (and they practically coincide with the concept of "subjective norm" used by Theory of Planned Behaviour). They were conceptualised by Schultz (1999) as "internalised self-expectations". Social norms instead refer to the behaviour of others and deal with the sets of beliefs about what other people are doing or what they approve or disapprove of doing (Cialdini, Reno, et al., 1990). More specifically, among the social norms Cialdini, Reno, et al. (1990) distinguish between:

- *descriptive social norms*: beliefs about what other people are doing;
- *injunctive social norms*: beliefs about what other people think should be done.

Injunctive norms are prescriptive concepts and they are thought to be effective as they signal the likelihood of obtaining social approval or disapproval. Descriptive norms are instead thought to be effective as payoff-maximising behaviour (Farrow et al., 2017) and also as indicators of injunctive norms (if there is uncertainty about them). People in fact frequently infer what ought to be done (injunctive norm) from what has been done (descriptive norm). Compliance with descriptive norms can therefore be regarded as an heuristic shortcut that reduces the effort to make a decision.

To stimulate a change in behaviour, any norm needs to be activated, namely it has to be "made salient in a particular setting" (Schultz, 1999, p. 4). According to the "Feedback intervention theory" (FIT) by Kluger and DeNisi (1996), providing feedback on social behaviour is one of the strategies to create such a salience. Also in this case, behavioural feedback is effective when it is discrepant from a behavioural standard, such as prior expectations, past performances, ideal goals, or norms.

To provide an example, I refer to the work by Schultz (1999), who dealt with an intervention aimed at increasing waste recycling by activating personal and social norms as standards in feedback interventions. Specifically, the author used descriptive and injunctive social norms to change a behaviour, by relying on the "Focus theory of normative conduct" (Cialdini, Reno, et al., 1990; Cialdini, Kallgren, et al., 1991) and on the theory of "Trans-situational influence of norms" (Reno et al., 1993). As recycling is a socially desirable behaviour, the provision of individual feedback stimulates the activation of personal norms: the feedback message puts pressure on the individual to act in accordance with the norm (as otherwise she would suffer a disutility caused by social disapproval). However, when the intervention stops and the individual assumes not to be observed any longer, it is possible that, no longer having a standard to compare against, the individual reverts to the previous behaviour. Experimental evidence collected by Schultz (1999), however, lead him to posit that the provision of group feedback about the recycling behaviour of neighbours can define a norm, which then remains active even when the feedback has stopped.

Also through later experimental research, Schultz et al. (2007) have argued that it is critical that feedback interventions manage to *define* descriptive social norms and the related standards, instead of just *activating* already existing personal norms. The behaviour standards set by descriptive social norms will in fact remain available also after the intervention has ended, and will keep motivating individuals to perform a given behaviour. Creating a new norm, however, requires more effort and time.

Another limitation to the behaviour change potential of social norms lies in the possible presence of ceiling effects: those who are already high in the desired behaviour might have less room for improvement than those who are low on such a behaviour (Schultz, 1999). Indeed, Farrow et al. (2017) argue that the extent to which use of social norms impacts actual behaviour depends on four key factors: *individual characteristics* (e.g. intrinsic motivation, socio-demographic characteristics, degree of familiarity with the specific behaviour), *characteristics of the norm evoked* (e.g. degree of difficulty to conform), *reference group* (e.g. size and geographical, temporal and social proximity), and *social and environmental contexts*. Again, individual, social, and contextual factors emerge as strictly inter-twined with one another.

3.3.2 Unintended boomerang effects

According to the Focus theory of normative conduct, feedback on others' behaviour provides evidence on which to base the social descriptive norm and thus creates a behavioural standard: group feedback provides individuals with a template for comparing not only their present behaviour but also their future behaviour, when the intervention will be over. According to this mechanism, even when the intervention is over, individuals treated with a group feedback still remind of the standard information on the behaviour of their neighbours, and such a standard might keep producing its effect over time, thus resulting in a long-lasting change in behaviour.

Note, however, that the Focus theory of normative conduct also warns that behaviour change is not the only possible outcome of social feedback provision. Schultz (1999) in fact notes that, in order to comply with the observed discrepancy, individuals might possibly modify or abandon the standard, or even reject the feedback message itself.

Furthermore, if the reference group's behaviour is in the opposite direction than the desired one, a "boomerang effect" may happen: response to behaviour-standard discrepancy might lead to decrease the targeted behaviour in those that are already high in the desired behaviour. The observation of the behaviour of other individuals that adopt lower standards might in fact reduce the perception of such a discrepancy, thus resulting in a reduction of the target behaviour itself.

Schultz et al. (2007) have explored in details the effect of providing feedback about social groups behaving against the desired behaviour and, through their theory on the "constructive and deconstructive power of social norms", have suggested how to overcome possible boomerang effects. Specifically, they have clarified that descriptive normative feedback has a different effect depending on whether feedback recipients are already doing better or worse than the average behaviour —namely if they are already performing the desired behaviour or not. If the household is consuming more than average households, then the provision of descriptive normative information about other households' consumption results in a decrease in the household's energy consumption. If instead the household has already lower consumption than average, the provision of the same descriptive feedback produces a "boomerang effect", namely an increase in the household's consumption. Research by Schultz et al. (2007) has shown that this phenomenon can be contrasted by providing households that are lower than average also with an injunctive message of approval, together with the descriptive norm feedback.

3.3.3 Effectiveness of leveraging social influence

Results by Schultz et al. (2007) have been confirmed by definitely large scale randomised experimental interventions aimed at assessing the effect of non-price strategies leveraging social norms for energy saving, that were performed a decade ago by Allcott and Mullainathan (2010), Allcott (2011), and Ayres et al. (2013). The authors analysed the effects on household electricity consumption due to the provision of Home Energy Reports (HERs), namely one-pager documents that were included in customers' electricity bills, which exploited descriptive social norms and offered information on the average consumption over the same billing period of similar neighbouring households (in terms of square footage and heating type of their home). Such reports were first offered to customers of the O'POWER utility company and then used by many utility companies in the US. Nowadays, similar contents are routinely offered also by utility companies' web- portals in Europe and in Switzerland as well. The experiment by Allcott involved more than 600'000 households across twelve US utilities between 2009 and 2011: half of them, randomly identified, received the HER treatment and half of them did not.

According to the above theory on "the constructive and deconstructive power of social norms" by Schultz et al. (2007), when provided with feedback about similar households' consumption, households consuming more energy than the average (namely, "the norm") were expected to decrease their consumption, while households consuming less than the average were expected to increase their consumption. Indeed, the latter phenomenon did not happen, thanks to the addition of an injunctive social norms message in the reports themselves: besides the social comparison feedback, households whose consumption was below the average were provided with a congratulations statement ("Great" or "Good", depending on the distance from average), accompanied by smiling emoticons; instead, households whose consumption was above the norm were provided with a "Below average" statement, accompanied by a frowning face emoticon.

Overall, Allcott (2011) found a 2% average treatment effect among all treated households (reduction in electricity consumption compared with a baseline consumption and measured against a control group). According to these results, similar non-price interventions leveraging social norms emerge therefore as promising strategies: if properly crafted, they succeed in producing energy saving outcomes, while at the same time guaranteeing highly cost-effective implementation at the large-scale. Estimates by Allcott and Mullainathan (2010) showed that the cost, from the utility perspective, of each kWh saved, was equal to 2.5 US dollar cents. A comparison against estimates of the average costs of other utility energy efficiency programmes run in the same years in the US, which were found to range between 1.6 to 3.3 US dollar cents per kWh (Friedrich et al., 2009) and between 5.5 to 6.4 US dollar cents (Arimura et al., 2012), confirmed the value of HER intervention strategies from the cost-effectiveness viewpoint.

The HER experiments by Allcott and Rogers (2014) also showed maintained energy savings over the long-term, which the authors explained as follows. Provided that the treatment is maintained over time, individuals gradually develop a "capital stock", which makes their energy consumption change to endure time. Specifically, as long as they react to the treatment, individuals develop new habits (consumption capital) and install new, energy-efficient appliances (physical capital), that ultimately support maintenance of lower energy consumption patterns. In the specific case of providing households with the HER, the latter acts as a cue, that temporarily reduces the marginal utility of energy consumption and induces households to save energy. As the cue is removed, households' energy consumption returns to its un-cued level: for instance, while the cue is active, household members remember to turn the lights off, but soon they lose motivation. However, thanks to the cue, they gradually invest in capital stock changes, which causes persistent effects, even when they go back to their un-cued behaviour and forget to turn the lights off. Based on these findings, therefore, Allcott and Rogers (2014) suggested that achieving long-term change is mostly a matter of repetition of the intervention over time, gradually reducing its frequency as long as the engaged households develop their consumption and physical capital towards energy saving.

A meta-analysis by Abrahamse and Steg (2013) has considered 29 resource use behaviour change interventions that involved social influence techniques, which were published between 1976 and 2013. The authors used their classification of six social influence approaches (see Section 3.3) as an entry-point. They concluded that interventions based on social influence are more effective in resource conservation than interventions that focus on the individual level. Particularly, they are more effective than the provision of individual information feedback and goal setting. They also showed, however, that all six social influence approaches are meaningful and can have larger impact if the target behaviour is observable and visible, especially in the public realm (Lapinski and Rimal, 2005). Again, the explanation behind the effectiveness of interventions leveraging social influence in a group is that, when social norms are made salient, people with high levels of group identification are more likely to act in accordance with the norm of the group.

Furthermore, according to their meta-analysis, the block leader approach emerged as the one characterised by the highest effect sizes, followed by public commitment making and modelling. The authors argued that these results were due to interpersonal and face-to-face interactions, that were in fact less prominent for interventions based on feedback and social norms. The exploitation of social networks and similarity factors can provide an additional explanation for higher effectiveness of block leader, public commitment and modelling strategies: personal communication may in fact make social norms and group identity more salient, thus favouring behaviour change. Based on these findings, in fact, Whitmarsh et al. (2021) have argued about the opportunity to leverage social influence processes to shape behaviour and promote climate action. Particularly, they have highlighted the importance of leadership in shaping social norms and in fostering collective effectiveness, as well as the key role by group discussion.

Early attempts in this direction had been for instance performed in the Nineties in The Netherlands, within the EcoTeam programme aimed at promoting pro-environmental behaviours. EcoTeam was inspired by the work by De Young (1996) and combined the provision of detailed procedural information and feedback on household's performance with a supportive social environment. The latter, in particular, consisted in the possibility for open discussions about possible behaviours in groups of peers, that allowed exchanging experiences, getting support, acknowledging the type of (evolving) group standards and norms, and also publicly committing (within the group) to perform a given behaviour. Results of a three-year longitudinal study on households participating in the EcoTeam Programme, performed in 1994-1996 by Staats et al. (2004), showed that, compared with a control group of similar untreated households, new pro-environmental behaviours, including electricity consumption at home, were sustained and maintained over time. These were documented by follow-up surveys based on self-reported behaviour and collection of quantitative consumption data over the two years following the intervention. Their approach appeared therefore as a promising avenue for future interventions. The authors themselves however remarked that, as group activities within the EcoTeam Programme were time-intensive for participants and required dedicated effort, already before the start of the intervention participating households were characterised by higher pro-environmental behaviours (as well as age, income, and education) than the average Dutch population. A comparison with national household survey indicated that only 20% of the population adopted the same level of pro-environmental behaviours as the participants. Staats et al. (2004) therefore concluded that specific efforts were needed to make those types of interventions appealing to broader segments of the population.

Later research has however dampened expectations on the effectiveness of the provision of social feedback. In their meta-analysis, Abrahamse and Steg (2013) found that social comparison interventions tend to have relatively low effect sizes —though, the authors noted, the low effect sizes they found might be due to presence of heterogeneous effects between sub-groups of study participants, which in turn also depend on their initial behaviours. The authors also noted that injunctive social norms tend to be effective at encouraging behaviour change, while descriptive social norms tend to have less consistent effects. The latter might also cause boomerang effects: individuals already implementing pro-environmental behaviour were found to worsen their performances when provided with feedback about average behaviour by other peers.

The above-mentioned meta-analysis by Delmas et al. (2013) has also shown that comparative feedback with respect to other peers is not a significant driver of energy conservation behaviour. However, the authors themselves have stated they could not derive robust conclusions, since the samples they considered were always quite small. Similarly, also the meta-analysis by Karlin et al. (2015) I cited above has found evidence that goal comparison feedback is most effective than social comparison feedback —nonetheless, also these authors remarked that research designs comparing the different feedback strategies would be needed to draw stronger conclusions. Current knowledge therefore suggests that, whether social comparison feedback is a valuable intervention strategy or not, is still as open research issue.

In their conclusions, Abrahamse and Steg (2013) stressed the need for further research aimed at not only assessing the interventions' effectiveness, but also at identifying the processes behind such effectiveness, and invited to account for the role of mediator factors such as similarity, social cohesion, or social identity, in future research. Furthermore, they remarked the need for long-term assessment of the interventions' effects, for exploration of potential positive spillover effects (the implementation of pro-environmental behaviours in different domains than those directly addressed by the intervention, see for instance Thøgersen and Ölander, 2003), as well as, following Allcott and Mullainathan (2010), for analysis of the interventions' overall cost effectiveness. In fact, despite interventions leveraging social influence factors could produce effects in the intended direction, this does not guarantee that they can be obtained in a cost-effective way.

3.4 Digital green nudging

As I have already remarked, behaviour change intervention strategies presented in the above sections have connections with intervention strategies developed within "nudging" theoretical frameworks, which acknowledge individuals' bounded rationality and presence of cognitive biases, and challenge the automated, routinely actions that individuals perform in non-conscious ways. A useful summary of key cognitive biases that affect household behaviour and of opportunities for nudging them towards change is offered by Frederiks, Stenner, and Hobman (2015a), who also provide examples on how to address cognitive biases in the context of energy saving interventions (Table 3.3).

Table 3.3: Cognitive biases affecting households' behaviour and policy suggestions to leverage them (Frederiks, Stenner, and Hobman, 2015a).

Cognitive bias	Policy suggestion	Example
Retaining the status quo, sticking to default settings, or deferring decision-making entirely (inertia).	Target practices that can easily and effortlessly be modified using default settings.	Set the dishwasher's default program to 'short-cycle' and/or to 'cold' water.
Satisficing: exerting only the effort to a satisfactory rather than an optimal result.	Simplify and keep communication short and simple, in order to reduce cognitive overload.	Avoid provision of too many energy saving tips; Automate and make the target action the default.
Loss aversion: losses are weighted more than equal gains.	Focus on the costs associated with energy-wasting practices.	Focus on time, effort, money saved, instead of benefits by saving energy.
Risk aversion when gain chances are high and loss probability are low (and risk-seeking in the opposite situation).	Focus on the low-risk of energy-saving practices that are safe, stable, and secure and offer relievers for financial, time, and effort risks.	Offer discounts, rebates, user-friendly operating instructions, helpful customer service.
Sunk cost effect: persist with an endeavour once valued resources have already been invested.	Reduce the salience of costs (time, effort, money) that have already been outlaid and draw attention to ongoing costs due to retaining inefficient items and wasteful energy practices.	Remark greater carbon emissions, higher electricity bills, and costs for repair and maintenance of outdated appliances.
Discounting: things are perceived as less valuable or significant if further away in time or space.	Reward for actions that yield benefits in the long-term.	Offer either intrinsic rewards (praise, recognition, etc.) and also extrinsic (in-kind gifts) rewards for actions with little immediate payoff.
Normative social influence: make social comparisons, follow the behaviour of others, conform to social norms.	Frame energy-saving practices as both common and socially desirable.	Advise consumers that people similar and close to them (e.g., peers, neighbours) are using less energy or taking certain energy-saving actions and convey social approval of such actions.
Motivation increases by rewards or incentives but, if extrinsic, only in the short-term.	Capitalise non-pecuniary rewards in order to produce durable behaviour change over the longer term.	Offer praise, public recognition and social approval, or in-kind gifts with suggestions to conserve energy.
Free riding and social loafing: lower contribution to the common good, if benefits are perceived to be available without paying for them.	Creating a shared and public group identity: contribution by every individual is important.	Publicly state the activities individuals are engaged into and the results obtained by them acting together whole community.
Using trust as a simple decision-making heuristic.	Information and incentives are more motivating if they stem from credible, trustworthy sources, in terms of both competence and integrity.	Send energy-saving messages from credible sources (public service commission) instead of low-credibility sources (local electric utility).
Availability bias: drawing on readily available information that springs to mind quickly.	Refer to actions that are easily available in consumers' memories and send simple prompts and reminders.	Refer to energy-saving behaviours that are topical or well-publicised in the media and draw on testimonials.

Nudges have already been extensively experimented and systematic reviews have been performed. The review by Byerly et al. (2018) has compared the effectiveness of nudge techniques against use of financial incentives (the provision monetary and non-monetary rewards or penalties) and education techniques (provision of facts, training and figures to increase knowledge). Even though nudge outcomes differ depending on the domain where nudges are implemented, and identification of the specific effects of single nudges is not always possible (frequently more nudges are combined in a single intervention), according to this review, commitments, changing default settings, messengers, and social norms emerge as the most promising nudge types, to be further explored. Also, there is a general agreement in literature that nudges should be regarded as complementary to traditional policy instruments rather than as integral substitutes for coercive (laws and regulations, which restrict freedom of choice) or economic (fiscal incentives, subsidies, taxes, fees) measures. Namely, nudges should be part of broader policy packages combining several elements (Lehner et al., 2016).

Particularly relevant to my work on apps is the use of nudges in the framework of digital settings. The specific concept of “digital nudging” has been introduced introduced in the recent years, as “the use of user-interface design elements to guide people’s behaviour in digital choice environments” (Weinmann et al., 2016, p. 433) and a growing scientific literature is exploring use and effectiveness of digital “green nudges” (C. Schubert, 2017). A very recent review by Beermann et al. (2022) has for instance identified six categories of digital nudge techniques that are frequently used in pro-environmental behaviour change processes: “nudges that structure the digital choice environment” (*default*), “nudges that signal non-personal or context information” (*priming, framing, social reference*), and “nudges that provide personal information or assistance” (*goal-setting, feedback*).

Another rich classification of the techniques that are most frequently used in digital nudging interventions for pro-environmental behaviour, which also provides examples about possible intervention content, is the one developed by Berger et al. (2022), which I entirely report in Table 3.4, with minor modifications only. A quick comparison with the techniques I presented in the previous sections shows there is a clear and direct overlapping between many of the intervention techniques that have been informed by (social) psychological theories and those that have been informed by behavioural economics and individuals’ cognitive biases.

A recent review by Zimmermann et al. (2021) analysed 43 articles reporting interventions aimed at digitally nudging pro-environmental behaviours, part of which specifically deal with the energy and climate transition. Based on such a review, the authors have concluded that the application of feedback in combination with social comparison is among the most effective digital nudges for actions that are repeated on a daily basis and that deal with energy and water consumption in households. However, the review has also shown that positive effects that were found in most of the analysed interventions were not sustained over time.

Table 3.4: Digital nudging techniques as they were summarised by Berger et al., 2022 (integrally reported from their work, with minor modifications only, to shorten texts).

Technique	When to apply	Definition	Example
Priming	Before action	A way of preparing people for their choice by simulating feelings and thoughts through specific topics, moods, or information like the consequence of a behaviour before it takes place.	A website banner emphasising the conscious collection of an online store, shown before customers access the shopping page.
Social norms	Before action	Individuals' beliefs about the typical and condoned behaviour in a given situation.	Displaying information on the donation willingness of past charity donors in an environmental charity webpage.
Goal setting	Before action	People are more likely to behave in line with their goal, if they commit beforehand to do so.	Committing to an energy-saving target (e.g. % savings relative to baseline consumption).
Default rules	During action	A situation where the preferred choice has been pre-selected and will remain so if the individual does nothing.	Default CO_2 compensation is enabled in flight booking portals.
Simplification	During action	Deliver complex (product) information or frame specific characteristics more noticeably.	Using logos (e.g. smiling world face) on sustainable products to nudge towards buying those labelled products.
Feedback	After action	Support people to reflect on whether their behaviour was good or improvable and point out to the consequences of the decision.	Providing detailed and customised feedback on energy consumption in terms of its costs and CO_2 emissions.
Social comparison	After action	A specific form of feedback, in which consumers receive information on their peers' behaviour, which is then compared with their own behaviour or consumption.	Receiving insights on the consumption of similar households (e.g. contrasting the household's consumption with average consumption by similar households, or displaying grades from A to G indicating the level of approval/disapproval for the household's consumption).
Framing	Before, during and after action	Use the bias of "anchoring": presenting the same information in different ways (frames) anchors people in different ways to their reference points, thus they decide differently.	Renaming the vegetarian food category on the menu: from "vegetarian main course" to "environmentally friendly main course for a happy planet".

3.5 Persuasive technologies

First developed in the field of Human-Computer-Interaction and inspired by the "Captology" theoretical framework I introduced in section 2.6.6, persuasive technologies were conceptualised as Behaviour Change Support Systems (BCSS), namely as "information systems designed to form, alter or reinforce attitudes, behaviours or an act of complying without using deception, coercion or inducements" (Oinas-Kukkonen, 2010, p. 6). Fogg (2003) identified seven intervention techniques that persuasive technologies can leverage to support changes in behaviour, as reported in Table 3.5.

Table 3.5: Intervention techniques that persuasive technologies can leverage (Fogg, 2003).

Technique	Definition
Reduction	Using computing technology to reduce complex behaviour to simple tasks increases the benefit/cost ratio of the behaviour, influencing users to perform it.
Tunnelling	Using computer technology to guide users through a process of experience provides opportunities to persuade along the way.
Tailoring	Information provided by computing technology will be more persuasive if it is tailored to the individual's needs, interests, personality, usage context, or other factors relevant to the individual.
Suggestion	A computing technology will have greater persuasive power if it offers suggestions at opportune moments.
Self-monitoring	Applying computing technology to eliminate the tedium of tracking performance or status helps people to achieve predetermined goals or outcomes.
Surveillance	Applying computing technology to observe others' behaviour increases the likelihood of achieving a desired outcome.
Conditioning	Computing technology can use positive reinforcement to shape complex behaviour or transform existing behaviours into habits.

Coherently with these basic techniques, inspired by Self-Determination Theory and stage models of behaviour change and considering later literature on persuasive technologies, Weiser et al. (2015) have recommended that persuasive technologies aimed at supporting behaviour change adopt the following "general design principles":

- *Offer meaningful suggestions*: persuasive systems should support their users in pursuing their goals. For this purpose, they first need to make users aware of behaviours that are conflictual with achievement of such goals, and then offer meaningful behavioural alternatives that align with the goals. Practical suggestions for alternatives can for instance be identified through artificial intelligence algorithms, which can learn users' behaviours from automatically collected data, integrated with user-declared or data-inferred behavioural constraints;
- *Support user choice*: in order to support the need for autonomy and fully empower users towards behaviour change, users should be given the opportunity to set

their own goals and progress towards them at their own pace. Particularly, goal possibilities offered by the persuasive system need to go beyond "either-or" options, in order to avoid that users feel patronised by the system;

- *Provide user guidance*: users should be guided through the process of acquiring the needed skills to perform the target behaviour. This may for instance be done by providing them with simple tasks, accompanied by clearly structured information. At the same time, the system should tolerate failures, in order to avoid frustration and favour that skill acquisition occurs with the needed autonomy;
- *Provide personalised experience*: persuasive systems should allow users to express their self-identity. This can be done by tailoring contents to specific user groups and by offering mechanisms that let them personalise their experience and interaction with the system, also accounting for the context in which they operate;
- *Design for every stage of behaviour change*: persuasive systems should provide information and, more broadly, features, that match the requirements of the different stages of behaviour change process. For example, with respect to the Transtheoretical model of behaviour change, during the pre-contemplation stage the system might provide ways to collect and reflect on behaviour-related data, with the aim of making the user aware of the problems associated with the behaviour she performs. Or, when the user is in the contemplation stage, the system might also provide the user with opportunities to set her goal for change. When the user is in the preparation stage, the system might provide information on alternatives suggesting how to change behaviour, thus triggering the increase in her ability to change. And when the user is in the action phase, the system might provide information feedback on the way the target behaviour has been performed and on progress towards goal achievement.

The above principles provide a general guidance for the features of persuasive technologies. From the operational point of view, the way the specific features are designed is often informed by the structured set of principles by the Framework for Persuasive Systems Design (Oinas-Kukkonen and Harjumaa, 2008; Oinas-Kukkonen and Harjumaa, 2009), which is grounded on a number of psychological theories (Oinas-Kukkonen, 2013). As the techniques for Persuasive Systems Design are simple, clear, and intuitive, and features of persuasive apps are directly informed by them, in the next chapters I will refer to them in order introduce the characteristics of the three persuasive apps I analyse in my three case studies. For this reason, I report them entirely in Table 3.6, by using exactly the same words as the authors who proposed them. Oinas-Kukkonen and Harjumaa (2009) organise the techniques in four categories: those aimed at *primary task support* (namely, help users to carry out their primary behavioural task), those aimed at *dialogue support* (namely, computer-human dialogue that helps users to move towards their goal or target behaviour), those aimed at guaranteeing *system credibility* (namely, how to design a system that is credible and persuasive) and those aimed at favouring *social influence* (how to design a system that motivates users based on social influence principles).

Table 3.6: Principles for Persuasive Systems Design by Oinas-Kukkonen and Harjumaa (2009).

Category	Principle	Definition
Primary task support	Reduction	A system that reduces complex behaviour into simple tasks helps users perform the target behaviour, and it may increase the benefit/cost ratio of a behaviour.
	Tunneling	Using the system to guide users through a process or experience provides opportunities to persuade along the way.
	Tailoring	Information provided by the system is more persuasive if it is tailored to the potential needs, interests, personality, usage context or other factors relevant to a user group.
	Personalisation	A system that offers personalised content or services has a greater capability for persuasion.
	Self-monitoring	A system that keeps track of one's own performance or status supports the user in achieving goals.
	Simulation	Systems that provide simulations can persuade by enabling users to observe immediately the link between cause and effect.
	Rehearsal	A system providing means with which to rehearse a behaviour can enable people to change their attitudes or behaviour in the real world.
Dialogue support	Praise	By offering praise, a system can make users more open to persuasion.
	Rewards	Systems that reward target behaviours may have great persuasive powers.
	Reminders	If a system reminds users of their target behaviour, they more likely achieve their goals.
	Suggestion	Systems offering fitting suggestions will have greater persuasive powers.
	Similarity	People are more readily persuaded through systems that remind them of themselves in meaningful ways.
	Liking	A system that is visually attractive for its users is likely to be more persuasive.
	Social role	If a system adopts a social role, users will more likely use it for persuasive purposes.
System credibility	Trustworthiness	A system that is viewed as trustworthy will have increased powers of persuasion.
	Expertise	A system that is viewed as incorporating expertise has increased powers of persuasion.
	Surface credibility	People make initial assessments of the system credibility based on a firsthand inspection.
	Real-world feel	A system that highlights people or organization behind its content or services has more credibility.
	Authority	A system that leverages roles of authority will have enhanced powers of persuasion.
	Third-party endorsements	Third-party endorsements, especially from well-known and respected sources, boost perceptions on system credibility.
	Verifiability	Credibility perceptions will be enhanced if a system makes it easy to verify the accuracy of site content via outside sources.
Social support	Social learning	A person is more motivated to perform a target behaviour if (s)he can use a system to observe others performing the behaviour.
	Social comparison	System users have greater motivation to perform the target behaviour if they can compare their performance with others.
	Normative influence	A system can leverage normative influence or peer pressure to increase the likelihood that a person will adopt a target behaviour.
	Social facilitation	System users are more likely to perform target behaviour if they discern via the system that others are performing the behaviour along with them.
	Cooperation	A system can motivate users to adopt a target attitude or behaviour by leveraging human beings' natural drive to co-operate.
	Competition	A system can motivate users to adopt a target attitude or behaviour by leveraging human beings' natural drive to compete.
	Recognition	By offering public recognition for an individual or group, a system can increase the likelihood that a person/group will adopt a target behaviour.

3.6 Persuasive gamified apps

The fast progress in ICTs and the related unprecedented diffusion of the smartphone has enabled novel intervention opportunities for persuasive technologies. Researchers from Human-Computer-Interaction fields have soon identified the large potentialities that could stem from coupling ICTs and smartphone apps to favour energy and low-carbon transitions. On-purpose developed smartphone apps in fact allow implementation of most of the behaviour change intervention techniques reported in the previous sections in a prompt, timely, and customised way — and hopefully also cost-effectively, as, once an app has been developed, its use can potentially be scaled to a large number of users. Through their design settings, apps can offer engaging and motivating user experiences, and, as they operate via smartphones, they are already integrated in users' everyday lives and activities. Therefore, they might result in higher and more durable energy saving outcomes, compared with intervention techniques aimed at providing static feedback via meters or In-Home-Displays.

Furthermore, smartphone apps also allow to track high granularity actions by the intervention target groups, to interact with them as needed during the intervention, via notifications or chats, and also to support bi-directional interaction between peer users involved in behaviour change interventions. Their features can target either the individual or the social level —or both— and, due to their dynamic and interactive nature, they are well-suited to support digital nudges and gamified approaches that unfold themselves in-between virtual and real-life. Also, automatic availability of in-app interaction data can provide insights on the users' level of engagement with different intervention techniques, thus offering novel opportunities for data collection processes that are not mediated by either the researcher's or the user's subjectivity.

Considering these characteristics and how easy to download and install a new app is and the carpet diffusion that smartphones have reached in the Global North countries (in 2020, the smartphone penetration rate was around 80% of the European and U.S. population, Newzoo, 2020), it is no surprise that smartphone apps have increasingly been adopted in real-world activities to support change, and specifically the energy and climate transition. Apps have been developed in both the private sector and research, with the aim of addressing a number of target groups (households, school communities, office employees, etc.), in a variety of domains, from energy consumption in buildings to transportation, as well as water or food consumption (Johnson et al., 2017; Mogles et al., 2017; Anagnostopoulou et al., 2018; A. Andersson et al., 2018; Cellina, Bucher, Veiga Simão, et al., 2019; Cellina, Bucher, Mangili, et al., 2019; Fraternali et al., 2019; Hedin et al., 2019; Spaiser et al., 2019; Cellina, Castri, Simão, et al., 2020; Suruliraj et al., 2020; Chatzigeorgiou and Andreou, 2021; Douglas and Brauer, 2021). Most of such apps rely on gameful approaches, as they frequently include gamified features (Weiser et al., 2015; Morganti et al., 2017; Shih and Jheng, 2017; Beck et al., 2019), and in some cases are even shaped as serious games (Wood et al., 2014; Baptista and Oliveira, 2019).

3.6.1 Techniques used by persuasive gamified apps

Usually, such apps either exploit connections with smart meters providing (nearly) real-time resource consumption data (such as for instance about energy or water consumption) or are equipped with data tracking systems (such as for instance about mobility or food consumption), which either automatically collect data or collect them via manual input by their users. Starting from such data, apps provide a number of persuasive features aimed at changing patterns of consumption and in most cases adopt gamification approaches. They can for instance provide users with feedback on consequences of their choices (e.g., in terms of energy consumption and CO_2 emissions), invite them to define personal goals for change, engage them in challenges, or favour active or passive interaction with other users, either through in-app systems or through external social networks. Typical elements that are included in app-based gamification processes have been identified by Hamari, Koivisto, and Sarsa (2014), who called them “motivational affordances”: points, leaderboards, badges/achievements, levels, stories/themes, goals, feedback, rewards, progress, and challenges.

Even though persuasive, gamified apps are increasingly widespread, their development is still a recent endeavour. At a general level, specific recommendations for persuasive, gamified apps are for instance offered by the works by Froehlich (2015) and Anagnostopoulou et al. (2018), who have also drawn on the behaviour change, nudging and persuasive technology literature I summarised in the previous sections. Practical recommendations for persuasive gamified apps stemming from their works can be summarised as follows:

- *Provide information*: information should refer to available behavioural options tailored to the individual’s needs, interests or context. It should be specifically related to her behaviour and be as timely as possible (close to the triggering cause, in both space and time), thus being easier to understand and remember;
- *Provide goal setting opportunities*: allowing individuals to select their own goals and targets for change can have powerful effects, since (provided that selected targets are really challenging for the individual), they create a self-competitive context leading the individual to strive for personal progress and mastery;
- *Provide feedback*: since individuals require information against which to assess their performance and progress over time, providing feedback is complementary to and essential for goal setting activities;
- *Provide rewards (incentives) or punishment (disincentives)*: these can be either tangible or intangible, expressed in monetary terms or in physical units, and need to be strictly related with individuals’ performances. Reward of good performances can reinforce individual motivation to adopt a certain behaviour, while punishment of poor performances can strengthen individual efforts —the latter has however to be handled carefully, since it can quickly demotivate (Foster et al., 2011).

- *Provide occasions for social comparison*: offer individuals the opportunity to compare their choices and performances against other people or groups they perceive as similar to themselves, such as members of the same community. This generates both peer pressure and a desire for imitation.

Additional recommendations for gamified interventions are also offered by Aparicio et al. (2012), who, informed by Self-Determination Theory, have identified the gamification elements that increase each of the three basic human needs identified by such a theory, which in turn contribute to increasing motivation to perform a given behaviour:

- to increase *autonomy*, they suggest to use: goal setting and commitment, configurable and customisable interface, choice between alternative activities, privacy control settings, notification control setting, customisation of profiles, avatars;
- to increase *competence*, they suggest to use: positive feedback, challenges, intuitive information feedback on progress, points, levels, leaderboards;
- to increase *relatedness*, they suggest to use groups challenges, message blogs, connection to social networks, chats, and leaderboards.

A few scholars (He et al., 2010; M. Z. Huber and Hilty, 2015) have taken a critical stance against typical point-attribution rules that usually characterise gamified apps, and, more broadly, "one size fits all" solutions for novel behaviours that are frequently suggested by persuasive apps. Doing so, app developers fail to acknowledge that there is no "one size fits all" solution in real-life, and that the effectiveness of a persuasive intervention tool is strictly dependent on individual baselines, viable alternatives, daily needs and constraints, besides individual attitudes and perceptions. Also, the dominant point-based approach has been criticised for its inherent technology patronising and elitist vision (Brynjarsdottir et al., 2012; M. Z. Huber and Hilty, 2015; Mols et al., 2015), according to which designers of persuasive systems apparently know what is always good and right, while ordinary people do not.

Scholars therefore recommended to rethink the currently dominant point-based reward systems, with the aim of giving app users as much freedom and customisation as possible. For instance, M. Z. Huber and Hilty (2015) and Froehlich (2015) clearly remarked that app users should be allowed to freely choose their own goal and target for change, by independently deciding if and how much they would like to change. Then, feedback and rewards provided by the app should explicitly be connected to progress regarding the target app users have autonomously set for themselves.

3.6.2 Effectiveness of persuasive gamified apps

Extensive reviews of key characteristics and limitations of persuasive apps have already been performed in the last years. For instance, Agnisarman et al. (2018) performed a systematic review of persuasive technologies for sustainable living based on 38 interventions dealing with electricity consumption, water consumption, transportation behaviour, and waste generation. Of the 38 studies, only four included a control group, and only three included post intervention evaluations to assess the long-term effect of the intervention. About half of them had shorter than one month observation periods, and sample sizes were small in most of the cases (for studies targeting households, sample size ranged from 1 to 30; while for studies targeting individuals, sample size ranged from 4 to 651). Therefore, Agnisarman et al. (2018) concluded that, despite current research suggests that persuasive technology has the potential to change user behaviour, stricter evaluation studies are needed, that account for larger sample sizes, longer duration of assessments, and controlled research designs.

Similar conclusions are also drawn by the systematic review by Adaji and Adisa (2022), which analysed use of persuasive technologies to influence sustainable behaviour by considering studies published between 2016 and 2021. The authors identified 16 intervention studies, which include mobile apps, serious games, web applications, virtual reality, and Internet of Things devices. Among them, 30% dealt with energy conservation, while others deal with food, waste reduction, urban mobility, water conservation, sustainable society overall, and climate change. For the analyses of effectiveness, 92% of the studies relied on self-reported data by users of the persuasive technologies (in many cases, collected via pre- and post-survey research designs), with very different study durations, in any case equal to a maximum of one year. Due to a lack of standard and rigorous procedures to evaluate the impact of the considered persuasive technologies, the authors could not offer conclusive assessments about their effectiveness, neither in the short-term nor in the long-term.

An extensive systematic review on residential energy feedback delivered through digital interfaces (smartphone or tablet apps and web platforms accessible via computer) recently performed by Chatzigeorgiou and Andreou (2021) has also explored characteristics, impacts, and limitations of major feedback studies performed in the last two decades. Despite presence of key methodological limitations in research designs (short study durations and small sample sizes which preclude generalisation), and the fact that different studies report wide varieties of effect sizes, the authors have concluded that the provision of digital feedback is a successful strategy to support the energy transition. As digital platforms, apps for tablets and mobile phones are increasingly integrated into everyday life, and therefore they have the potential for ubiquity interactions with their users, the authors consider them as promising platforms, on which further research is worth to be performed. Particularly, the authors suggest that research should explore the role of different feedback strategies for different target populations and contexts.

Turning to empirical digital interventions that leverage gamification and serious games in the energy domain, a systematic review by Johnson et al. (2017) analysed the effects of 25 digital interventions (mobile apps delivered via smartphone and non-mobile apps, delivered via personal computer). The authors could not perform a rigorous meta-analysis, as many of the studies they considered did not report sufficient statistical information to compute effect sizes; thus, they opted for a narrative description of intervention characteristics and their outcomes. Most of the studies relied on convenience sampling, identified through personal networks, and most of them adopted a questionnaire-based evaluation design, frequently investigating both quantitative and qualitative aspects, in a mixed method approach. Out of the 25 analysed studies, 19 reported only positive effects, 6 reported both positive and negative effects, and none reported only negative effects. However, Johnson et al. (2017) also remarked that the analysed studies are characterised by a number of shortcomings, such as small sample sizes, poor methodology description, absence of controls, use of descriptive statistics only, narrow data collection timeframes, and estimation of applied games' effects as a whole, instead of identifying the effect of different gamified elements and on different target groups.

Particularly, their analysis has shown that, when a follow-up data collection period was included in the studies, positive behaviour changes as a result of the interventions were not maintained. They have argued that this might be because most of the studies were lacking "meaningful gamification" elements (*play*, *exposition*, *choice*, *information*, *engagement* and *reflection*, as they were proposed by Nicholson, 2015) and rather simply leveraged reward-based gamification (points, levels, leaderboards, achievements, badges). The authors thus argued that, by acting on the principles of competence, autonomy, and relatedness identified by Self-Determination Theory (SDT), meaningful gamification elements would likely lead to an increase in integrated motivation, thus favouring long-term consolidated change, much more effectively than reward-based gamification, that instead only offers temporary effectiveness on the three SDT principles. Overall, Johnson et al. (2017) concluded that there is "encouraging initial evidence that applied games can have a positive influence in the domestic energy conservation domain" (p. 263). However, they also argued that further research is needed, in order to identify the effectiveness of applied games over a longer timeframe, across different user groups, and on varying the game elements that are exploited.

Also Beck et al. (2019) performed an extensive review of apps including at least one gamification component and directly targeting household behaviours in the broad domain of energy consumption (energy efficiency and conservation, solar or renewable energy, home energy efficiency upgrade, efficient transportation). Their review study, based on 57 apps that in 2017 were available on the US Apple App Store, did not try to estimate the app's effectiveness; rather, it analysed their components and features, in order to verify if and to what extent they were grounded in behavioural theories. Outcome of their analyses lead the authors to conclude that most of the gamified apps were adopting immature and not sufficiently theory-informed designs, which hints at their supposed limited capability to drive long-lasting behaviour change.

These outcomes are aligned with those by Morganti et al. (2017) and Rapp et al. (2019), who specifically focused on applied gamified interventions. Morganti et al. (2017) performed a systematic review of interventions leveraging gamification or serious games to foster energy efficiency behaviours. They were only able to find ten scientific articles dealing with this topic, which mostly lacked the needed scientific rigour for estimate of effectiveness. In a paper aimed at introducing a journal special issue specifically focusing on how gamification studies can be strengthened (theoretical grounding, methodological references, and applied intervention designs), Rapp et al. (2019), instead, did not perform novel reviews, but referred to previous literature. Anyhow, both groups of scholars called for long-term, theory-driven, rigorous designs, that allow to assess the specific effects of the different gamification components, the impact of contextual factors and individual differences, and also possible broader outcomes beyond the individual level, such as changes in social relations between individuals involved in gamified activities.

Finally, another recent systematic review analysed 29 interventions using serious games and gamified mobile apps to foster pro-environmental behaviours (Boncu et al., 2022). Specifically about interventions targeting reduction in energy consumption, the authors also identified a lack of consistency between the interventions' short and long term impacts: in some of the examined cases, reduction in energy consumption was obtained in the short- and in the long-term; in other cases it was not maintained in the long-term; and in some other cases, not even statistically significant reductions were found in the short-term. Boncu et al. (2022) have assumed that such a lack of coherency is partly due to the different features of the gamified persuasive apps used in each intervention and partly due to the small sample sizes that were used in order to assess the interventions' effectiveness: usually lower than 100, in some cases also split in sub-samples treated with different features. Small sample sizes, in particular, might explain the lack of statistically significant results. Further, Boncu et al. (2022) have remarked that attrition problems, which are reported in some of the analysed interventions, might have critically affected the long-term results, in terms of both their effect size and their statistical significance. Again, these results call for more rigour in research designs aimed at estimating interventions' effectiveness.

3.7 Methodological weaknesses

I conclude the literature review by summarising key methodological limitations and shortcomings that have emerged from previous interventions, as well as the recommendations that scholars have advanced to address them, in order to inform activities on the case studies that I tackle in the next chapters and guarantee their methodological rigour.

From the methodological point of view, early reviews of behaviour change interventions reported in literature highlighted a number of critical pitfalls or limitations. The most recent reviews instead indicate that policy interventions and the studies aimed at assessing their effectiveness are methodologically more thorough: their outcomes, therefore, provide more reliable and robust indications. Overall, however, ten key methodological limitations emerge from the literature that I reviewed in the previous sections:

- *poor identification* (and consideration in the analysis) *of the underlying determinants of energy consumption behaviour*, which stem from a generally poor theoretical understanding of how behaviour is formed and can change. This negatively affects the capability to understand why an intervention was effective or not, thus providing policy-making with limited support (Abrahamse, Steg, et al., 2005; Michie and Prestwich, 2010; Hermsen et al., 2016; Sunio and Schmöcker, 2017; Morganti et al., 2017; Michie, Carey, et al., 2018; Beck et al., 2019; Rapp et al., 2019; Nielsen, Clayton, et al., 2021);

- most assessments of effect *lack scientific rigour*, due to limited adoption of experimental (randomised controlled trials, the "gold standard" for interventions) and quasi-experimental designs (that *ex-post* consider control groups matched to treatment groups) or even for the lack of control groups at all; To date, in fact, most empirical research on residential energy consumption has involved non-experimental studies, which are inadequate for testing causal relationships and determining the direct effects of predictors on outcome variables, including the precise causal impact of various interventions on possibly observed changes in behaviour (Abrahamse, Steg, et al., 2005; Delmas et al., 2013; Hamari, Koivisto, and Pakkanen, 2014; Hamari, Koivisto, and Sarsa, 2014; Vine et al., 2014; Frederiks, Stenner, Hobman, and Fischle, 2016; Johnson et al., 2017; Andor and K. M. Fels, 2018; Nisa et al., 2019; Chatzigeorgiou and Andreou, 2021; Nielsen, Cologna, et al., 2021). Particularly, Allcott and Mullainathan (2010) recommend that more rigorous statistical analyses are performed, that for example include use of time series analyses (with consumption trend de-seasoning), of Difference-in-Differences estimators, of weather controls (for instance, through heating and cooling degree days[3]), and of demographic controls. They are however aware that, even though inclusion of a large number of household-level covariates would definitely be useful to improve precision of estimates of average treatment effect, accessibility to rich sets of covariates, such as for instance the ones used by Allcott (2011) (house-related covariates of energy consumption: year of construction of the house, type of heating, square-footage, single or multi-family dwelling type, renter or owner-occupied, presence of a fireplace or of a pool, number of bedrooms and of bathrooms; household-level covariates of energy consumption: number of residents, age of the household head, income), is usually precluded;
- *sample sizes are small* and there is a *lack of baseline measurements*. Furthermore, interventions are usually characterised by large within-group variance, which implies a reduction in statistical power, which in turn reduces chances to find statistically significant effects (Abrahamse, Steg, et al., 2005; Hamari, Koivisto, and Pakkanen, 2014; Hamari, Koivisto, and Sarsa, 2014; Vine et al., 2014; Frederiks, Stenner, Hobman, and Fischle, 2016; Johnson et al., 2017; Agnisarman et al., 2018; Andor and K. M. Fels, 2018; Nielsen, Cologna, et al., 2021; Chatzigeorgiou and Andreou, 2021; Boncu et al., 2022);
- *user engagement may critically decrease over time and attrition problems may occur* (Hamari, Koivisto, and Sarsa, 2014; Perski et al., 2017; Löschel et al., 2020; Boncu

[3]A brief and effective explanation of the concepts of heating and cooling degree days is provided by Allcott (2011) on p. 1086.

et al., 2022) —though this is not captured by the short-term duration of most of the interventions (during which user interest and engagement might have been higher just for the novelty by the intervention itself);

- there is a *lack of analysis of long-term effects of the interventions*: most of the studies only focus on short-term effects (less than three months) and in fact do not report whether behavioural changes were maintained, namely whether new energy-saving routines were formed, or whether energy use returned to baseline levels (Abrahamse, Steg, et al., 2005; Allcott and Mullainathan, 2010; Ferraro et al., 2011; Delmas et al., 2013; Karlin et al., 2015; Hermsen et al., 2016; Morganti et al., 2017; Agnisarman et al., 2018; Andor and K. M. Fels, 2018; Nisa et al., 2019; Chatzigeorgiou and Andreou, 2021; Adaji and Adisa, 2022);
- most of the studies *include more than one energy saving intervention strategy*. This implies that confounding effects between the different strategies can occur and precludes the possibility to identify which strategy is actually more effective than the others (Abrahamse, Steg, et al., 2005; Delmas et al., 2013; Hamari, Koivisto, and Sarsa, 2014; Karlin et al., 2015; Johnson et al., 2017; Morganti et al., 2017; Andor and K. M. Fels, 2018; Nisa et al., 2019; Rapp et al., 2019; Chatzigeorgiou and Andreou, 2021);
- voluntary households participating in interventions are *not representative of average population*: they tend to have higher intrinsic motivation, higher income, and higher education than average population, which implies that generalising results is critical and cannot always be performed (Hartman, 1988; A. Nilsson et al., 2014; Geelen et al., 2019; Nisa et al., 2019; Tiefenbeck et al., 2019; Cellina, Vittucci-Marzetti, et al., 2021). Particularly, if universal deployment of the behavioural intervention is envisioned after the field analyses aimed at estimating its effectiveness, according to Sergici and Faruqui (2011) both the treatment and control groups should be randomly selected among the population of interest (and random selection should then ideally be followed by the random allocation to treatment and control conditions). However, the authors indicate that, if a later large-scale deployment of the intervention is programmed via an opt-in framework, then it is preferable to adopt opt-in and voluntary sampling strategies also in the field analyses aimed at providing evidence of the intervention's effects;
- there is a *low level of granularity in the provision of feedback* on overall energy consumption, which would instead benefit by appliance level break-down (Delmas et al., 2013; Geelen et al., 2019; Chatzigeorgiou and Andreou, 2021);
- there is a tendency not to consider quantitative energy consumption data and to *rely instead on self-reported behaviours* (Steg and Vlek, 2009; Morganti et al., 2017; Andor and K. M. Fels, 2018; Nisa et al., 2019; Adaji and Adisa, 2022), which may be affected by social desirability biases. Failing to consider energy consumption data also reduces the capability to spot possible rebound effects, as not all behavioural changes necessarily result in energy savings. Particularly, households may have

bought energy-intensive appliances with the money saved through behavioural changes, thus ultimately increasing their overall energy use;

- interventions tend to focus on typical individual-level factors addressed in psychological research, such as attitudes and abilities, *overlooking macro-level factors* that drive and steer consumption, such as demographic or societal developments, e.g. the TEDIC factors I introduced in section 2.6, which however play a key role in shaping the physical and technical structures that condition behavioural choices and the related energy consumption (Abrahamse, Steg, et al., 2005; Karlin et al., 2015; Rapp et al., 2019).

3.8 Conclusions

In this Chapter I performed a narrative review of policy interventions aimed at favouring the energy and climate transition in households, through behavioural changes to energy sufficiency. The review shows that policy interventions aimed at saving energy in households have been extensively developed since at least five decades; in the last two decades, digitalisation and ICTs have favoured their implementation through smart devices and apps, which also enabled further interaction possibilities to motivate and support behaviour change processes. The review also suggests that, despite interventions originated from different disciplines and research strands (most of them being grounded in either social psychology, behavioural economics, or computer science), most of the intervention techniques, methodologies and practical tools are similar and strictly intertwined between such research strands. Also, digital tools and apps are now becoming the standard intervention devices across all the disciplines. Many review studies and meta-analyses have been developed in order to systematise knowledge of their effectiveness and, under an evidence-based policy-making approach, inform possible later large-scale deployment of the interventions themselves.

Table 3.7 summarises findings from the narrative reviews, systematic reviews, or meta-analyses of behavioural interventions that I considered in this chapter. The table shows that there is no agreement among the results of previous analyses in terms of which intervention strategies or techniques should be favoured. Some of the studies did not even manage to identify treatment effects, due to weaknesses in the interventions or in the accompanying data that were reported together with them. This limitation does not only characterise older studies, but was also found in the recent studies specifically focusing on digital nudging, gamification and persuasive apps. Overall, therefore, the evidence I collected indicates that much research is still needed to clarify if —and to what extent— persuasive apps can support the energy and climate transition.

Table 3.7: Effectiveness of behavioural interventions according to previous literature reviews.

Article	Study characteristics (range of publication years reported in brackets)	Most promising behaviour change intervention techniques
Abrahamse, Steg, et al. (2005)	Review of 38 interventions for household energy conservation (1977 - 2004).	Feedback on energy consumption (especially if frequent).
Allcott (2011)	Review of 17 experiments for electricity saving by OPOWER (2009 - 2011).	Social comparison feedback.
Osbaldiston and Schott (2012)	Meta-analysis of 87 experimental interventions targeting pro-environmental behaviour (1980 - 2009).	Social modelling, individual commitment.
Abrahamse and Steg (2013)	Meta-analysis of 29 (quasi-)experimental interventions for resource conservation involving social influence (1976 - 2013).	Block leader approach, public commitment making, modelling. Social comparison feedback has the lowest effect size. Social influence overall more effective than individual level interventions.
Delmas et al. (2013)	Meta-analysis of 156 experimental studies on energy conservation behaviour (1975 - 2021).	High-involvement information-based strategies, such as energy audits (social comparison feedback is no significant driver of energy conservation).
Karlin et al. (2015)	Meta-analysis of 42 studies on energy conservation feedback (1976 - 2010).	Individual consumption feedback and goal setting.
Hermsen et al. (2016)	Systematic review of 72 studies on digital feedback to support behaviour change (2004 - 2015).	Individual feedback generally effective in disrupting habitual behaviour.
Johnson et al. (2017)	Systematic review of 25 digital interventions leveraging gamification and serious games in the energy domain (2007 - 2015).	No single intervention technique identified, however most of the studies report positive behaviour change effects.
Morganti et al. (2017)	Review of 10 interventions leveraging gamification or serious games to foster energy efficiency behaviours (2009 - 2016).	No single intervention technique identified (studies lacking rigour to estimate effects of single intervention components.)
Agnisarman et al. (2018)	Systematic review of 38 persuasive technologies for sustainable living (2000 - 2016).	No single intervention technique identified (studies lacking rigour to estimate effects of single intervention components).
Andor and K. M. Fels (2018)	Systematic review of 44 experimental nudge interventions on energy conservation (1978 - 2017).	Social comparison feedback.
Nisa et al. (2019)	Meta-analysis of 83 RCTs to promote household action on climate change (1976 - 2017).	Choice architecture Social comparison.
Chatzigeorgiou and Andreou (2021)	Systematic review of 27 interventions on household energy consumption feedback delivered through digital interfaces (2007 - 2018).	Digital feedback is a successful strategy; study methodologies do not allow conclusions on the effectiveness of different ways to provide feedback.
Khanna et al. (2021)	Meta-analysis of 122 behaviour change interventions to reduce energy consumption in residential buildings (1975 - 2020).	Monetary incentives.
Zimmermann et al. (2021)	Systematic review of 43 digital nudging interventions for pro-environm. behaviour (2010 - 2020).	Social comparison feedback.
Adaji and Adisa (2022)	Review of 16 interventions using persuasive technologies to favour pro-environmental behaviour (2016 - 2021).	No single intervention technique identified (studies lacking rigour to estimate effects of single intervention components).
Boncu et al. (2022)	Systematic review of 29 interventions leveraging serious games and gamified mobile apps to foster pro-environmental behaviour (2004 - 2021).	No single intervention technique identified (studies lacking rigour to estimate effects of single intervention components.)

The literature review analysis I performed shows that insights about the key research question on how to achieve the long-term effectiveness of behavioural interventions are still lacking. In fact, the majority of the analyses on behaviour change interventions reported in literature, including those focusing on apps, limit themselves to assess short-term effects during or immediately after the intervention.

Also, the analysis has identified ten key methodological limitations that affected previous research on behavioural interventions. Overall, outcomes of this analysis show that the call for action that was performed a few years ago by Frederiks, Stenner, and Hobman (2015a) is still fully valid and open: research is still needed to perform large-scale empirical interventions characterised by strict evaluation procedures that guarantee greater reliability and generalisation possibilities, estimate of the durability of the effects over time, as well as easier scaling-up possibilities if the assessment of outcomes shows beneficial energy and climate results.

4

Case one: enCompass

> *If we achieve our sustainability targets and no one else follows, we will have failed.*
>
> — **Paul Polman**
> Business leader

The enCompass project[1], developed within a Horizon 2020 European Union (EU) research programme over the period 2016-2019 and led by Politecnico di Milano, designed, implemented and tested a behaviour change app targeting energy saving in households, schools, and public tertiary buildings. The enCompass app was automatically fed by electricity consumption data provided by smart meters, as well as by indoor and outdoor temperature, luminance, and humidity sensors. Such data was processed by context-aware, adaptive algorithms, which provided app users with gamified motivational stimuli for change and customised energy saving recommendations, tailored to the users' context, activity, comfort level and phase in the behaviour change process.

Three different versions of the enCompass app were developed, on varying its target users: one app for schools, one app for public tertiary buildings, one app for households. All such apps were pilot tested in three European countries (Germany, Greece, Switzerland) for a full year, between June 2018 until May 2019, by involving voluntary users. For a broad overview on the other project activities and their results, I refer the reader to the enCompass project website, which reports all the deliverables and scientific publications associated with the project (`https://www.encompass-project.eu/project-materials/`, last accessed on January, 27 2023).

Here I focus on the enCompass app targeting households and on the field test performed in Switzerland, in the small village of Contone, which is a hamlet of the municipality of Gambarogno in the region of Locarno. I open the chapter by introducing the characteristics of the enCompass app for households, then present the quasi experimental methodology I used to assess the effects of the enCompass intervention in Contone and to address the research questions RQ1 - RQ3 introduced in Section 1.4. I then present the results I obtained by estimating the average treatment effect on the treated (ATT) by the enCompass app, both in the short and in the long-term (up to two full years after the end of the intervention) and by accounting for heterogeneity between them, and conclude the chapter with a summary of such findings. I leave discussion for Chapter 7, which also deals with findings from the other case studies.

[1]"Collaborative Recommendations and Adaptive Control for Personalised Energy Saving", funded under call EE-07-2016-2017 "Behavioural change toward energy efficiency through ICT". Grant agreement No 723059.

With respect to previous analyses that were performed by the enCompass team during the Horizon 2020 project, my novel contribution specifically lies in:

- the clarification of the app's features from the theoretical perspective, by specifically fitting them within both stage models of behaviour change and principles for persuasive systems design;
- the identification of a different control group of households, which overcomes the lack of statistical significance of the estimated effects of the intervention;
- the use of a longer time period of analysis (two additional years), which allows to assess the persistence of the effects in the long-term;
- the use of panel regression models and the subsequent identification of possible heterogeneous effects of the intervention on varying the characteristics of the households and the level of interaction with the app by the treated households, grouped by means of clustering techniques.

4.1 The enCompass persuasive app

The enCompass app for households could in principle work with any type of energy consumption, such as for instance gas or electricity, provided that the energy source was centrally distributed and the related consumption data could be read by smart meters and automatically be sent to the enCompass back-end software. For the Swiss pilot, however, only electricity meters were available, since, by design of the EU project, the intervention had to be located in the region of Locarno supplied by the utility company "Società Elettrica Sopracenerina SA" (SES), which was among the project partners. In such a region, no gas grid is available, therefore the enCompass app could only deal with electricity consumption, provided by electricity smart meters. This did not preclude, however, the possibility to deal with energy consumption for water and space heating: in households equipped with electric boilers for hot water, electric heating systems, or heat pumps, treatment with the enCompass app addressed electricity consumption for both heating and non-heating purposes. Otherwise, the treatment with the enCompass app only addressed electricity consumption for non-heating purposes (lighting and electric appliances).

The app's developers designed its motivational features by taking inspiration from three key behavioural theories I already introduced in Chapter 3, namely the Theory of Planned Behaviour (TPB), the Norm-Activation model (NAM), and stage models of behavioural change. Specifically, as stated in Koroleva et al. (2019), the developers grounded the app's design into the Stage model of self-regulated behavioural change by Bamberg (2013), which is an attempt to integrate constructs from these two behaviour change theories with stage models of change.

For the sake of simplicity, here I introduce the app's features by framing them from the perspective of each behaviour change stage, using the five-stage classification proposed by the Transtheoretical model of behaviour change by Prochaska and Velicer (1997):

pre-contemplation (of change), *contemplation*, *preparation*, *action*, and *maintenance*. To facilitate comparison with other (app-based) behaviour change support systems, I also refer to the techniques listed in the taxonomy for behaviour change interventions by Abraham and Michie (2008) and to the principles by the framework for Persuasive Systems Design (PSD) by Oinas-Kukkonen and Harjumaa (2009).

Table 4.1 shows an overview of all the enCompass app's features, allowing to frame them in terms of both the theoretical background and the persuasive principles and techniques they exploit. Overall, apart for a comparison with other households in the "leaderboard" section, all the features offered by enCompass focus on the single household and do not leverage any social interactions among the community of its users.

4.1.1 Pre-contemplation stage

App users in the *pre-contemplation* stage have no motivation for reducing their energy consumption and do not intend to take action for change. This might be due to insufficient information about their possibilities for change or to a lack of trust in their ability to change. To support users towards change, among the processes suggested by Prochaska and Velicer (1997) at this stage, enCompass implements *consciousness raising*. The app increases the users' awareness about their amount of consumption by means of feedback on energy consumption, on the impact produced by their energy consumption, and on the overall level of comfort enjoyed at home. Feedback on energy consumption is provided via barplots and numerically shown in kilowatthours (Figure 4.1). Users interested in quantitative feedback can also choose between a daily or weekly visualisation and can select the specific period to visualise.

Figure 4.1: enCompass pages providing feedback on on daily and weekly energy consumption via barplots. App users can customise the time period and data granularity.

Table 4.1: Features of the enCompass app targeting households.

Stage	Process	Feature	Technique	PSD framework
Pre-contemplation	*Consciousness raising* Increasing awareness for causes, consequences and cues about a behaviour	Consumption and impact feedback (money, CO_2 emissions, hedonic factors)	2. Provide information on consequences 12. Prompt self-monitoring of behaviour	Self-monitoring Tailoring
Contemplation	*Self-reevaluation* Cognitively and affectively assessing one's self-image, with and without a particularly unhealthy habit	Default 20% saving goal Impact and comfort feedback	12. Prompt self-monitoring of behaviour	Self-monitoring Tunnelling
Preparation	*Self-liberation* Believing that one can change and committing to act on such a belief	Goal setting	4. Prompt intention formation 10. Prompt specific goal setting	Reduction Personalization
Action and Maintenance	*Counterconditioning* Learning of more sustainable behaviours that can substitute the less sustainable ones	Generic tips and customised recommendations	7. Set graded tasks 8. Provide instruction	Suggestion Personalization Reduction
		Goal setting	11. Prompt review of behavioural goals	Personalization
		Impact and comfort feedback	12. Prompt self-monitoring of behaviour 13. Provide feedback on performance	Self-monitoring Tailoring
	Contingency management Providing consequences (rewards) for taking steps in a particular direction	Points, badges and vouchers for real-life prizes	14. Provide contingent rewards	Praise Rewards
		Leaderboard	19. Provide opportunities for social comparison	Social comparison Recognition Competition
	Helping relationship Providing social support (caring, trust, general support) for new behaviour	Notification system to stimulate action maintenance	6. Provide general encouragement 17. Prompt practice	Reminder

The visualisation of impact is instead provided by three different pages, between which the users can freely switch. Impact is in fact directly measured in terms of the amount of saved electricity, numerically shown in kWh and represented via bar plots similar to those showing the amount of consumed electricity. Saved electricity is then translated into saved money (monetary impact), into saved CO_2 emissions (environmental and climate impact), and into amount of obtained rewards and gamified achievements (hedonic impact). As shown in Figure 4.2, saved money is represented by a set of piggy banks, saved CO_2 by a set of trees, and obtained achievements by a set of jars filled with balls and badges: the more the piggy banks, trees, and jars of treats, the more respectively the money and CO_2 emissions saved and the rewards and achievements obtained. Estimates of monetary and carbon savings are performed based on average values of electricity cost [CHF/kWh] and CO_2 emissions [CO_2/kWh] per unit of electricity consumption, which are provided by the local utility company. Namely, they do not reflect the exact costs and CO_2 emissions of the specific amounts of electricity consumed by each household in a given period, and are just approximate indications.

Overall, these self-monitoring techniques provide opportunities for the households to receive both general information about electricity consumption and information on its consequences, via the impact pages. They are also tailored to the app users, who can choose the most relevant visualisation of the impact according to their interest (money/CO_2/hedonic factors).

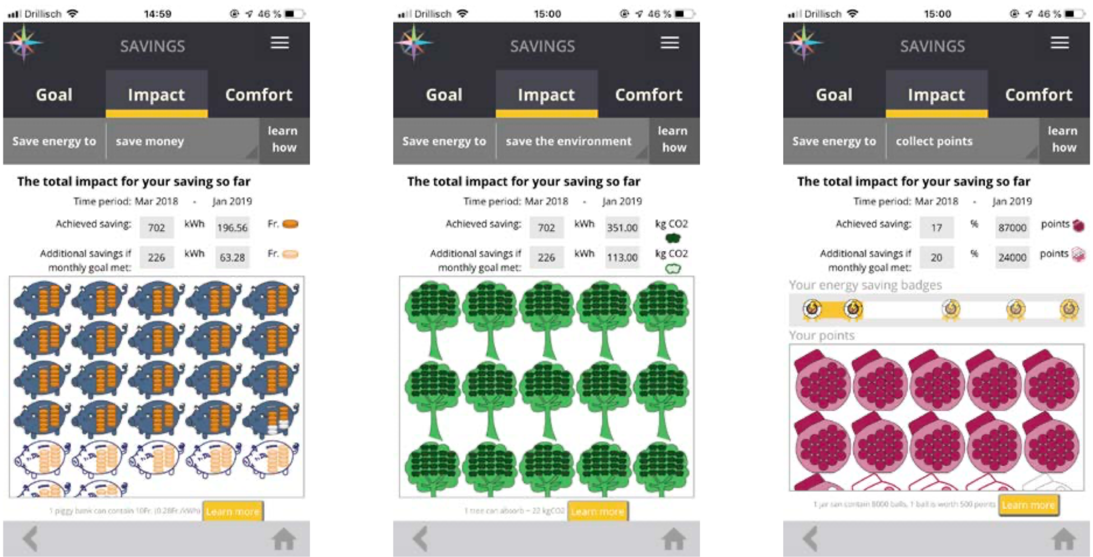

Figure 4.2: enCompass pages providing feedback on the financial, climate, and hedonic impact of energy saving activities, via the piggy bank, tree and jar of treats metaphors.

4.1.2 Contemplation stage

Contemplation is an intermediate stage before action, in which households could remain stuck for a long time in "chronic contemplation" of elements in favour and against a change in their consumption practices: it in fact refers to a stage in which users intend to change *within the next six months*. To support a short stay in such a stage, enCompass provides feedback on progress towards an energy saving goal by default set at 20% (percentage of energy savings relative to the household's baseline consumption in the same month of the previous year[2]). By making the 20% savings appear to be realistic and feasible to the user, enCompass directly supports a *self-reevaluation* process, namely a cognitive and affective assessment of one's self-image, by considering one's household in both configurations of energy consumption at the level of the baseline and energy consumption 20% lower than that.

To easily represent progress towards goal achievement, the metaphor of the battery is used: the more the energy is consumed, the lower the level of remaining available energy, which is visually shown in the battery representation (see Figure 4.3). If, considering the amounts of consumption reported by the smart meter, the household appears to be able to meet the energy saving goal, the battery level is shown in green colour and a green-coloured rewarding and reinforcing message is shown; warning messages are instead shown if the household is close to *not* meeting the goal (orange message and battery level) or has already failed the goal (red message and battery level).

Furthermore, enCompass aims at showing that energy savings can be obtained with no or limited decrease in comfort. For this purpose, it provides a screen which both shows the amount of electricity saved over a given period and the average perception of comfort by the app user over the same period of time. The latter is automatically elicited by the enCompass app. Comfort is in fact measured by means of data collected by indoor temperature, humidity and luminance sensors installed in the enCompass household, coupled with indications of perception of indoor comfort by the householders, captured via the enCompass app itself, through easy and frequent questions asking about how much comfortable householders feel. Average values of the sensor-collected data and user feedback on comfort feeling are thus represented in the app, close to the amount of saved energy (Figure 4.3): if savings have been obtained and the comfort level has been maintained, the related amounts are presented in a green coloured textbox, otherwise a warning is shown in a red coloured textbox. Ideally, this screen shows that energy savings can be obtained with no or minimal decrease in comfort level, thus strengthening the intention to keep saving energy.

If this *self-reevaluation* process is successfully activated, opportunities for "tunnelling" would emerge. Namely, enCompass would start to guide users through a process of experiences, which provides them with opportunities to be persuaded to change along the way. App users would thus enter the *preparation* stage, in which additional enCompass features would be at work to support change.

[2] Use of the enCompass app requires that for each household at least one full year of electricity consumption data is available, disaggregated at the monthly level, that can be automatically accessed by the enCompass app back-end system. In the case of the enCompass project, this was easy to obtain, since the local utility company was among the project partners and ensured access to such data.

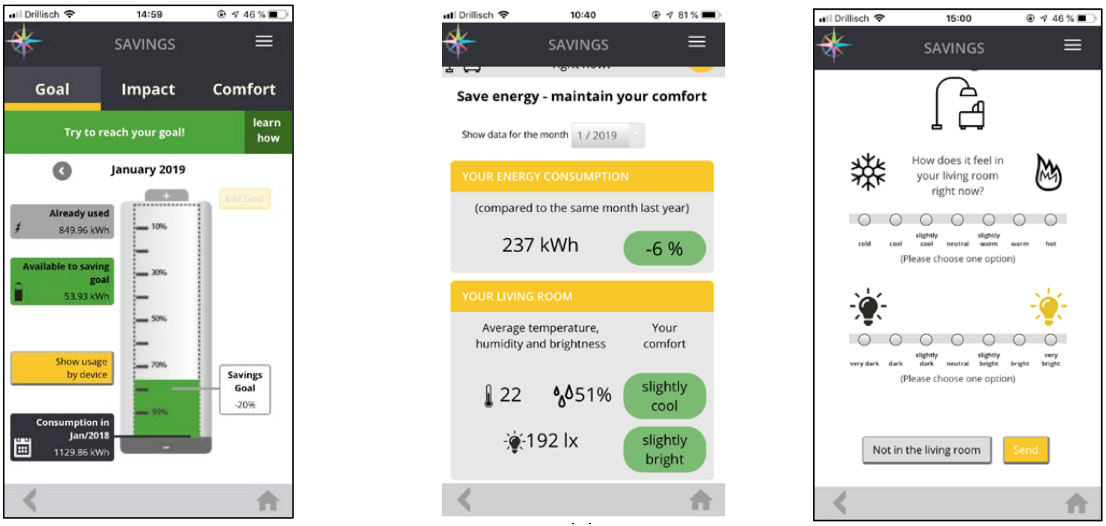

Figure 4.3: enCompass pages providing feedback on the energy saving goal achievement via the battery metaphor (left), on the amount of saved energy (center) and on the enjoyed comfort level (right).

4.1.3 Preparation stage

Preparation is the crucial stage during which individuals develop plans for action, with the intention to put them into practice in the very near future. At this stage, a *self-liberation* process occurs, namely individuals start to believe they can actually change, and commit to act coherently with such a belief. To support such a process, enCompass provides goal setting features: households are invited to move from the default energy saving target and to set their own target value, again to be compared with consumption values recorded in the same month of the baseline period (the year before). The choice of such a target indicates the amount of change they would like to achieve and is therefore related to the amount of effort they plan to invest in trying to reduce their energy consumption. This choice is on purpose fully left to the users (*personalisation*), to avoid any one-size-fits-all, patronising or super-imposed solution. Energy saving practices at home are in fact constrained by contextual factors, such as for instance family needs (number and age of household members, number of members spending time inside/outside home, type of activities required to be performed at home but also outside the home, ecc.), or infrastructural characteristics of the building (energy retrofits, sun exposure, ecc.), which affect practical feasibility of achieving ambitious energy saving targets: in some cases even small targets for change might imply a significant reorganisation of one's individual and family routines and habits.

Both the goal and the target value can be changed over time, thus allowing users to start with relatively easy targets and to later increase difficulty (*reduction* of complexity), or to simplify their target, in case they started with a too challenging one. Once they have started to take action, users are then free to progress at their own pace and in their own direction, while being stimulated by enCompass to achieve their personal goal for change.

4.1.4 Action and Maintenance stages

In the action stage individuals need to consciously act to change their daily consumption patterns. Further, such actions need to be sustained over time. Once households have committed themselves to their own goal, in order to enter the *action* stage they need practical support and guidance on how to perform the change. Namely, a *counter-conditioning* process needs to be activated, that allows household members to learn more sustainable behaviours with respect to their current one.

To help individuals to achieve their goal step-by-step (*reduction*), a “Learn how” button is shown in the goal page, which leads to the app section showing a selection of non-customised energy-saving tips, which are the same for all the households, and of customised energy-saving recommendations, which depend on the household profiling. Thanks to the information collected by smart meters, temperature, humidity and luminance sensors, activity logs of the in-app interactions, characteristics of the building and of its heating system, as well as socio-demographic information collected via in-app questionnaires, the enCompass app is in fact able to profile its users in order to dynamically provide them with customised suggestions to save energy. Provision of tips and recommendations allows to reduce the complexity of energy saving processes: by identifying simple actions to be frequently repeated, it supports the creation of new habits. Both tips’ and recommendations’ contents span a wide range of energy consuming practices that are performed into the household, such as washing (clothes or dishes), cleaning, using information technology appliances, showering, and heating/cooling rooms. Their full list is available on the project website, at `https://www.encompass-project.eu/project-materials/other-material/` (last accessed on January, 27 2023).

Furthermore, enCompass rewards households for the effort they have performed in trying to put the new behaviour into practice (*contingency management* process). For each received tip/recommendation, households are in fact invited to commit to perform the suggested action, via the button “Ok, will do”. The more the commitment to take action, the more the households are rewarded by points. The latter are used within a gamified incentive mechanics, which shows the households’ position on a weekly leaderboard exploiting social comparison and competition persuasive principles. Households in the top positions are rewarded with virtual badges in the app and with physical voucher prizes in real-life as well, which provides a feeling of recognition of one’s achievement.

Points are also attributed whenever household members perform actions in the app; furthermore, depending on the type of action performed on the app, users are also attributed unexpected badges, which either reward the performances they have achieved (such as for instance reaching one’s saving target or being in the top leaderboard positions) or reward for having responded to the profiling questions requested by the app in order to tailor recommendations. Together with the impact feedback features showing the saved money and CO_2 emissions, all these gamification elements contribute, either directly or indirectly, to rekindle the users’ commitment to implement their action plan towards their goal. Therefore, they support maintenance of households' new energy consuming practices.

Finally, enCompass provides general support by means of a notification and reminder system aimed at activating an *helping relationships* process. To retain the households' interest and attention over time, enCompass regularly sends them push notifications (from two to five per week, depending on the user preferences), that either provide tips/recommendations, notify if they are close to missing their energy saving goal, announce attribution of a badge/voucher, or suggest a reactivation of app usage, if the system automatically monitors low levels of in-app activity. Besides acting as a reminder, such notification systems congratulating for achieved results can substitute for an in-person counsellor, enhancing the perception of social support.

As long as individuals practice with the implementation of the new behaviour, they enter the *maintenance* stage, during which the need for external support decreases, they are less tempted by relapse and are more confident that change can be maintained over time. To avoid relapse, which is always possible and would lead individuals back to an earlier stage (in the worst case, to *pre-contemplation*), the same techniques as in the action stage are put into practice —however, they are needed less frequently. The behaviour change process concludes when individuals reach the *termination* stage, namely they have no temptation and the new behaviour is regularly put into practice for an indefinite period of time. In such a condition, the enCompass app would no longer be needed and users could confidently stop using it, having managed to change their domestic energy consumption practices.

4.2 Research design

I now introduce the overall methodology I follow in order to tackle the RQ1 - RQ3 research questions about the energy and CO_2 saving effectiveness of persuasive apps targeting households. I estimate the causal effect of use of the enCompass app on electricity consumption and CO_2 emissions, both in the short-term (namely, during app use, RQ1) and in the long-term (namely, a reasonably long period after its use, RQ2). Also, I verify if the magnitude of such a causal effect differs, on varying the observed characteristics of app users (heterogeneity analysis, RQ3).

4.2.1 Experimental research for evidence-based policy-making

The social sciences have long identified experimental research as the "gold standard" method for causal inference (Cartwright, 2007; Brancati, 2018). Typically borrowed from the natural and physical sciences, experimental research allows to identify the cause and effect relationship between two or more phenomena. It has however been seldom exploited in the social sciences, and particularly in sociology, due to practical and ethical reasons (Bruce and Yearley, 2006; Turner, 2006). Experiments in fact require that researchers manipulate one or more variables of interest (the supposed causes) under a strictly controlled procedure and measure the effects on an another dependent variable (the outcome), keeping constant all the other variables possibly affecting the outcome. Since such a manipulation is rarely possible for typical social phenomena, traditionally sociology tended to rely on naturally occurring phenomena, within observational studies, and to look for "robust associations" (or, following Goldthorpe, 2001, "robust dependencies") between variables, instead of causal inference.

Use of randomized controlled trials (RCTs), more generally known as experiments, has however started to emerge in social research in the last two decades (M. Jackson and Cox, 2013), particularly within evidence-based policy-making and impact evaluation studies (Cartwright, 2009; Gerber and Green, 2012), including energy-related research (Vine et al., 2014; Frederiks, Stenner, Hobman, and Fischle, 2016). Experiments are in fact particularly well-suited to assess the effectiveness of pilot policy interventions and to decide whether they are worth large investments to support their scaling up (Gertler et al., 2016), since they allow to account for the social dynamics characterising real-life phenomena (M. Jackson and Cox, 2013), which are crucial to evidence-based policy-making.

Experiments build on the counterfactual approach to causation proposed by the "Rubin causal model" (Holland, 1986), which allows to estimate the effect of an independent variable X (cause, such as a policy intervention targeting individuals) on another dependent variable Y (effect, such as the intervention outcome), through the differences on the dependent variable observed in two identical groups of individuals: the treated (treatment group) and the untreated (control group) ones. The difference between such groups in the average value of variable Y after the intervention would thus coincide with the average treatment effect on the treated (ATT), which produces a measure of the presence of a causal relation between X and Y and of its intensity. Statistical hypothesis testing would then allow to verify if the observed differences in Y between the control and the treatment group are likely to have been produced by chance or are actually caused by the intervention, namely by variable X (M. Jackson and Cox, 2013; Willer and H. A. Walker, 2007).

The key assumption under this approach is that the two groups of individuals are identical in all their observed and non observable characteristics and only differ for the manipulation consisting in the treatment. Such a comparability but for the treatment can be obtained by randomly attributing the treatment to individuals within a sample of units of analysis. Through the randomisation of treatment assignment, possible reasons for dependency between cause and effect are similarly distributed in the treatment and control group, since any covariates (including those that are not observable) have the same probability distribution between the two groups.

4.2.2 Choice of the quasi-experimental approach

Unfortunately, in the case of enCompass a true experiment cannot be performed. Basically, this is because, for households to be treated with the enCompass app, the following two requirements had to be met, which proved to be too restrictive for an experiment:

- availability of smart meters automatically providing the enCompass app with high granularity energy consumption data (consumption over 15 minutes periods), needed for the feedback and impact features of the app;
- willingness to install physical sensors in the house, in order to feed the enCompass app with the piece of information on indoor luminance, humidity and temperature needed for the assessment of comfort and the provision of customised energy saving recommendations.

In Switzerland electricity smart meters are not widespread yet, though the national Energy Act regulation compels all electricity providers to have at least 80% of smart meters by the end of 2027. Availability of districts fully covered by smart meters is thus still limited —and it was even more so at the start of the enCompass intervention. In this framework, the utility company responsible for setting and managing the Swiss enCompass intervention ("Società elettrica Sopracenerina SA" SES, the company providing electricity and water services to the region around the city of Locarno) identified the small village of Contone (about 800 inhabitants), hamlet of the municipality of Gambarogno (overall, about 5'150 inhabitants), as the proper area to locate the intervention. There, in fact, a wide roll-out of smart meters had recently been completed and therefore high granularity electricity consumption data were available (also retroactively, starting from the beginning of 2017) for all village buildings. Even though Contone is a very small village, where about 300 smart meters were installed, out of which 230 were associated with households, it was regarded as suitable for the enCompass intervention. Pilot activities envisioned by the EU enCompass project had in fact indicated a target of 100 households to be actively treated with the enCompass app. This means that a bit less than one out of two of the households of Contone were expected to be treated with the enCompass app.

Ideally, all the 230 households of Contone could have been included in an experiment, by randomly allocating half of them to the treatment group and half of them to the control group. This however was not possible, due to the second requirement, namely the need for physical installation of sensors into the treated households. Such a requirement was expected to largely reduce the number of households willing to accept the treatment. Furthermore, the SES utility company did not accept an opt-out strategy (namely, automatically considering all households as part of the experiment, though leaving the possibility to opt-out before random allocation of the treatment), and asked for interested households to actively opt-in and register for project participation. The only remaining possibility for an experiment would have been to consider the sample of registered households and to randomly allocate the treatment within them, thus generating two comparable treatment and control groups. The treatment and control sample sizes generated this way would however have been too small; furthermore, half of the households having already accepted the installation of sensors and willing to use the app would have been allocated to the control group, thus generating their dissatisfaction, to be best case, and/or creating contamination problems, to the worst case. To avoid these problems, the research team responsible for the enCompass project decided to perform a “quasi experiment” (Campbell and Stanley, 2015; Maciejewski, 2020). Even though quasi experiments are clearly less rigorous than true experiments in evaluating policy programmes, they are often the only viable solution for applied empirical research. Provided that they are developed under strict methodological conditions, authors such as Vine et al. (2014) list both experiments and quasi-experiments as “robust experimental designs” and acknowledge that, when practical considerations preclude use of true experiments, “it is not only appropriate but also necessary to use the best available quasi-experimental techniques to try to answer important policy questions using the best empirical evidence” (p. 628).

Under such a quasi-experimental design, the enCompass treatment group was made of all the self-selected households who, already equipped with a smart meter, decided to join the project and to install the related sensors in their premises, according to an opt-in framework. The control group was instead selected among other households located in the same region and equipped with smart meters, via a matching procedure aimed at ensuring as much comparability as possible in the observable characteristics of the two groups.

4.2.3 Choice of the panel data approach

In order to estimate the average treatment effect on the treated (ATT) of the enCompass quasi-experiment, the unit of analysis is the household, the independent variable X consists in a dichotomous variable taking value “1” if the household has been treated with the enCompass app, and “0” otherwise, and the outcome dependent variable Y is the amount of electricity consumed in a given period. Note that, through emission factors for the mix of consumed electricity available in scientific literature (for the case of Switzerland, estimated by Krebs and Frischknecht (2021) as equal to 128 g CO_2/kWh), once the amount of consumed electricity is available, an estimate of the outcome on CO_2 emissions can be automatically obtained as well.

Under these quasi-experimental conditions, households self-select themselves into treatment, which can cause a bias in the estimate of the treatment effect, if this is simply obtained by comparing mean outcomes in the treatment and control groups during the intervention. In order to remove such a bias, I opt for a before/after panel data (repeated measurement) approach. By comparing differences in outcome values within a single household before and after the intervention, and then comparing such differences between the treatment and control groups, panel data approaches allow to control for time-invariant systematic differences —both the observed and unobserved ones— between the members of control and treatment groups, thus removing potential biases in the estimates of the effects. Namely, panel data approaches allow to remove possible bias due to time-invariant unobserved heterogeneity among the self-selected households.

Quasi-experimental approaches require at least two periods of data to be compared, collected over the same treated and untreated units of analysis: before and after the treatment. For enCompass, during the EU project the choice was made to collect electricity consumption data over full year periods. Guidance for evidence-based policy-making for energy consumption-related interventions, in fact, suggest to run the treatment for at least a full year, in order to account for natural variations in energy consumption data due to seasonal effects (Sergici and Faruqui, 2011). This is particularly true in the case of enCompass, which addresses both heating consumption (provided by heat pumps, electric heating systems, or boilers for hot water) and non-heating consumptions for electric appliances and lighting. The enCompass treatment period was in fact set over an entire year, from June, 1 2018 to May, 31 2019. To guarantee that comparisons are not affected by seasonal effects and are made on a similar seasonal basis, pre-intervention data (baseline) were thus collected for the same period as the intervention, though one year before (June, 1 2017 - May, 31 2018[3]).

Under such a panel data approach, two methods can be used to estimate the ATT of the enCompass intervention. One is use of a "Difference-in-Differences" estimator (Wooldridge, 2015): mean differences in the outcome variable before and after treatment are computed within the treatment and control groups; then, such means are differenced between the two groups. Alternatively, a panel regression model can be used (Wooldridge, 2010). I computed them both, also in order to check accuracy of the results I obtained with both methods. However, also following Sergici and Faruqui (2011), here I only report results of the panel regression model, which is more versatile since it allows to account for more than two measurement periods.

Panel data regressions are in fact particularly well-suited to the enCompass case, for which I aim at estimating the ATT not only in the short-term, but also in the long-term. Electricity consumption data were already collected in the enCompass project for two yearly periods, from June 2017 to May 2019. For this dissertation I collected two additional years of data for the same panel of users, by accessing them through the SES utility company: besides electricity consumption during the baseline and intervention period, for my analyses I also considered electricity consumption over the June 2019 - May 2020 period (one year after the enCompass intervention) and June 2020 - May 2021 period (two years after the enCompass intervention). This was possible since the SES utility company had kept the pseudonymisation table that identifies correspondence between the identification number of the households included in the enCompass intervention and their POD number, namely the code that identifies each electricity end-user throughout Switzerland.

Overall, for each household of the treatment and control group four full years of electricity consumption data were available (as summarised in Table 4.2), that I profitably fed into a panel regression model to address my research questions.

[3]Incidentally, note that this added a further constraint on the requirement for households to be included in the quasi experiment: they had to be living in the same home in Contone since at least one full year before the start of the enCompass intervention —namely, their electricity consumption had to be available since at least June, 1 2017.

Table 4.2: The time periods I considered to assess effectiveness of the enCompass intervention.

Period	Year	Dates	Type of period
Baseline	Year Zero	June 2017 - May 2018	Pre-treatment
Intervention	Year One	June 2018 - May 2019	Treatment
One year after the intervention	Year Two	June 2019 - May 2020	Post-treatment
Two years after the intervention	Year Three	June 2020 - May 2021	Post-treatment

4.3 Identification of the treatment group

Having clarified the general methodology I used to estimate the short- and long-term ATT of the enCompass intervention, I now introduce how the treatment and control groups have been identified. As above mentioned, households of the treatment group were identified through an opt-in voluntary procedure, while households of the control group were identified via a matching procedure I performed on a later stage.

For small customers such as households, the electricity market liberalisation has not been activated yet in Switzerland. This implies that in each regional district only one electricity provider is active and that all households have to necessarily exploit the electricity services it provides. This was a great advantage for a policy intervention such as enCompass, since the SES local electricity provider actually supplied all households of Contone and could therefore easily send them all invitations to join the project. A partnership with the local municipality also allowed to activate a stronger recruitment campaign: first a press release was issued, followed by articles in the local media, and then a customised printed letter was sent to each household in the hamlet of Contone, with the invitation to join the project. No disguise was used and the energy and carbon saving impacts expected to stem from app use were clearly indicated in all recruitment materials.

As an additional recruitment strategy, SES offered a few rewards: at project sign up, each household received a 100 CHF electricity bill discount, plus an additional energy-saving gadget, approximately valued 10 CHF. Further, three “super-prizes” of the value of 700 - 1’000 CHF each were advertised to be offered in a final raffle open to all the households that would have remained active until the end of the enCompass intervention — and in June 2019 such prizes were indeed attributed to three participating households.

Overall, the recruitment concluded with a lower number of registered households than the 100 target value that was indicated in the enCompass project proposal: n= 75 households signed up to try use of the enCompass app and installation of the related sensors. This number was however regarded as satisfactory and sufficient for the purpose of the project, and therefore the intervention could start as planned.

The number of valid households for my analysis however is lower than n=75, for the following reasons:

- n= 10 households never logged in the enCompass app (case of non-compliance): indeed, they did not receive the enCompass treatment and for this reason I do not consider them among the treatment group;
- in the four-year observation period, n=6 households installed a photovoltaic (PV) plant for self-production and -consumption of electricity. This is critical for the assessment of the overall amount of their electricity consumption: the share of electricity that they directly self-consume via the PV plant is in fact not measured by the electricity smart meter. After installation of the PV plant, therefore, smart meters register a decrease in the household's electricity consumption, though this does not correspond to a real decrease in consumption: in addition to the meter's reading, the household is simply consuming part of the electricity directly produced by the new PV plant. For this reason, I do not consider these households in the treatment group;
- during the four-year observation period, n= 4 households changed the technical equipments of their house, by moving from an oil-based heating system to a heat pump heating system. Since the different heating system largely affects the amount of electricity consumption (presence of the heat pump implies much higher consumption for heating purposes, which before installation of the heat pump was not recorded by the smart meters, since heating was provided by oil), comparing electricity consumptions over four years is not possible. Therefore, I do not consider them in the treatment group either.

Ultimately, n= 55 households were available for the analytical sample I considered as the enCompass treatment group. Specifically, for each household of the treatment group, the following piece of information was available for my analyses, either provided me by the SES utility company or by the enCompass consortium:

- *Purpose of electricity use* (and related technical equipment): households that use electricity for heating purposes were classified based on their technical equipment: *boiler* (if they used electricity for hot water heating), *electric heating system* (if they used an electricity direct heating system for room heating), *heat pump* (if they used a heat pump for home heating). The remaining households only used electricity for lighting and appliances (from large appliances such as the dishwasher and washing machine, to all information technology-related appliances, such as computers, laptops, smartphones, or televisions, to small domestic appliances such as the hairdryer or the mixer). These households were classified as *appliances*. Indeed, for the treatment group final analytical sample, no cases of households were found in the category of electric heating systems: among the 55 treated households, only appliances, boiler, and heat pump categories were available. Please, note all categories include electricity use for appliances, in addition to which they respectively also account for production of hot water (boiler category) or of space heating (electric heating system and heat pump categories). This information was constant across the considered four-year time period;

- *Electricity consumption*: smart meters register consumption data in kilowatthours [kWh] on a 15 minute interval. Since for my specific research questions such a high granularity information was not necessary, SES provided me with monthly electricity consumption data for each household, from June 2017 to May 2021 included[4]. In order to account for the above-mentioned seasonal variability characterising household electricity consumption (the consumption of each month of the year tends to be inherently different from the other months, due to natural variability in outdoor temperature, luminance, and weather, which indirectly also influence the time spent at home and hence the related consumption), I aggregated the related data over four single years. Doing so, I obtained a set of four yearly electricity consumption data (baseline: year 0; treatment: year 1; one year after enCompass: year 2; two years after enCompass: year 3) for each household, which can be directly compared with each other. Besides the minor data cleaning operations mentioned in the footnote, in fact, no additional data pre-processing operation was needed. Starting from such electricity consumption data, I could automatically estimate the corresponding amount of CO_2 emissions, by means of the emission factor [gCO_2/kWh] provided by the literature. Therefore, I performed all the analyses by referring to electricity consumption and then at the end I also provided a final estimate of the ATT on carbon emissions;
- *Level of app use*: for each household and each month during the enCompass intervention, the managers of the enCompass back-end at Politecnico di Milano collected the monthly number of logins and interactions in the enCompass app. As shown in Section 4.8, I use such a piece of information for an additional in-depth analysis on the enCompass effect, on varying its level of use.

In the framework of the enCompass project, further rich data-sets were collected for each participating household, through two questionnaires delivered immediately before the start of field activities, in May 2018, and immediately after their end, in June 2019. The questionnaires collected socio-demographic information regarding households' composition and members, the type of building and energy system equipments, as well as behavioural and attitudinal constructs (perceived behavioural control, behavioural intention, and energy knowledge), and (only in the second questionnaire) the households' evaluation of the whole enCompass system. In this dissertation I however do not consider such data-sets, since they are not available for the control group I rely upon, and therefore they cannot be used in the panel data analyses I perform.

[4]More precisely, SES provided me with monthly readings of energy consumption. Since such readings are cumulative, I obtained the actual monthly consumption by subtracting readings between two consecutive months. For a limited number of months, due to smart meter data transmission problems, the monthly consumption data were missing in the data-set provided by SES; however the actual consumption during the related period was not lost, since it was simply accounted for in the reading of the following month. In such cases, I estimated the average monthly consumption over the two-month period and manually inputed the same value to both months. Such an operation did not introduce any biases in my analysis, since no missing data occurred for the months corresponding to the end of one year and the start of the following one.

4.4 Identification of the control group

For the identification of the control group, I adopted a largely different procedure: households of the control group were in fact identified while the enCompass intervention was already ongoing, with the aim of matching the characteristics of its components as much as possible with the characteristics of the treatment group. As I mentioned in Section 1.5.1, during the enCompass EU project a different control group was considered, with respect to the one I consider for this dissertation. In that case, documented in project deliverable 7.4 “Final overall validation and impact report’ (at the time of writing of this dissertation still under embargo to allow for scientific publications) and in Koroleva et al. (2019), a sample of self-selected households living in nearby villages (other hamlets of the municipality of Gambarogno) was considered. Such a sample was identified among households already equipped with smart meters, who accepted to answer two questionnaires asking the same questions as the above questionnaires for the treatment group members (apart for the questions on app evaluation). The questionnaires were advertised together with the electricity invoices and offered small prizes through a random draw open to all respondents, again supplied by the SES utility company. The aim was to gather a sufficiently large number of control households, to be compared with the treatment group not only in terms of their electricity consumption, but also in their behavioural attitudes and determinants, always in a Difference-in-Differences fashion. However, the number of obtained responses was very small: only n= 25 households answered both questionnaires and could therefore enter the control group. Research teams involved in the enCompass project found a 4.48% electricity saving average treatment effect on the treated, though, as expected due to the very low sample sizes, it was not statistically significant.

I therefore opted for a different strategy for the identification of the control group, which has the only drawback of ignoring the additional information collected through the questionnaire. For this purpose, I asked the SES utility company to provide me with the monthly electricity consumption data for all the Contone households connected to smart meters that had not registered for the enCompass intervention, in order to select from such a sample a control group of households that was as large as possible and at the same time comparable with the households of the treatment group, by means of matching techniques.

Overall, the number of Contone households equipped with smart meters since 2017 and not included in the enCompass treatment group is equal to n= 174. A first element of comparability with the treatment group is provided by geographic closeness: households of both groups are located in the very same hamlet of Contone, therefore they essentially share the same topographical conditions and enjoyed a comparable exposure to the sunlight and other meteorological factors that can drive their electricity consumption. Then, such households can be compared with the treatment ones based on the available information that the SES utility company owns for all their customers, namely the possible presence of photovoltaics power plants and purpose of electricity use: appliances only, appliances and boiler, appliances and direct electric heating, and appliances and heat pump.

Out of such 174 households, I removed those equipped with a photovoltaics plant, for the same reasons as the treatment group. I also removed those who changed their heating system during the four-year analysis period (from June 2017 to May 2021), by replacing their old oil heating plant with a modern electricity-fed heat pump. By removing such cases, a total of n =163 households remains, potentially available to be used for the control group. From now onwards I thus refer to this group as the "potential control group".

4.4.1 Check of imbalance

To verify comparability between the treatment and the potential control group, I perform two types of check of imbalances, as suggested by Sergici and Faruqui (2011) and Gerber and Green (2012) for experimental research. First, I check the percentage distribution of the available covariate, the purpose of electricity use, in both groups. Unfortunately, the check can only be performed by considering such a single covariate, since this is the only one that I can observe in both groups.

This single check is however sufficient to suggest the presence of critical imbalances (selection bias) between the two groups, as reported in Table 4.3: in the control group, households with boilers are present in larger percentages than the treatment group, in which to the opposite, the share of households with heat pump is definitely larger. Also, in the control group there are a few households equipped with electric heating, which instead is not present in the treatment group. Such differences can lead to different electricity saving potentials between the two groups, as households with heat pump, boilers, or electric heating can implement energy saving measures both for heating and for non-heating electricity consumption, while "only appliances" households can only implement savings in non-heating electricity consumption.

Table 4.3: Characteristics of the "potential control group" (n = 163) compared with characteristics of the treatment group (n = 55).

	Appliances		Appliances + Boiler		Appliances + Electric heating		Appliances + Heat pump		Total
	Num	%	Num	%	Num	%	Num	%	Num
Treatment	22	40	11	20	—	—	22	40	55
Control	76	47	51	31	6	3.7	30	18	163

Furthermore, presence of heat pumps and electric heating systems can be regarded as a proxy indicator of the age of the building: recent (or recently retrofitted) buildings tend to install heat pumps, due to their higher energy efficiency and thus energy and CO_2 saving capacity; instead, historical buildings tend to rely on direct electric heating systems, which were easier to install in a pre-existing building heated via wood chimneys, with no need to build an hydraulic circuit. Since recent or recently retrofitted buildings tend also to be equipped with double glazings and thermal insulating coatings, their energy consumptions and CO_2 emissions tend to be already largely optimised, and thus the energy saving potential of householders' behaviour and practices is lower in percentage terms. Furthermore, environmental awareness in households equipped with heat pumps might be higher than in households equipped with direct electric heating systems, if installation of the heat pump is an explicit choice made by household members. Again, this suggests that the group of heat pump households may have lower potential energy and CO_2 savings than the group of direct electric heating households, as their environmental awareness might have also driven their daily behaviour and practices at home. Therefore, the treatment group might have overall lower energy and CO_2 saving potentials than the control group, thus leading to underestimate the enCompass' app effectiveness.

To confirm presence of imbalances, as specifically suggested by Gerber and Green (2012) in the framework of experimental research, I verify if, for the observed characteristics, imbalances are larger than one would expect from chance alone. For this purpose, I perform a linear probability model and regress the assigned treatment on the available covariates. By computing the aggregate F-statistics, I can then test the null hypothesis that the covariates predict the treatment membership no better than one would expect by chance. If p-value is higher than 0.05, at the 5% significance level the covariates predict the treatment membership no better than would be expected by chance, therefore the two groups are balanced. Otherwise, they are not. For this purpose, as shown by equation (4.1), I regress a dichotomous variable taking value "1" if the household is treated, and "0" otherwise, on three dichotomous variables respectively indicating with "1" whether the household is equipped with boiler, electric heating or heat pump (in order to avoid collinearity, I do not consider the dichotomous variable representing households using electricity only for appliances). This provides me with the probability of households being assigned to treatment ($P(treat_i) = 1$), based on the available covariates.

$$treat_i = \beta_0 + \beta_1 boiler_i + \beta_2 heat\ pump_i + \beta_3 only\ appliances_i + u_i \tag{4.1}$$

The F-statistic's p-value resulting from regression (4.1), equal to 0.005945, indicates the regression's parameter estimates of standardised coefficients are jointly statistically different from 0 at the 0.01 level, which confirms that the treatment group and the potential control group of n=163 households are not balanced with respect to the covariates that I can observe about their characteristics.

Additionally, as suggested by Sergici and Faruqui (2011), I also look for differences between the pre-treatment mean baseline electricity consumption values of the treatment and potential control groups. Presence of large and statistically significant differences in the mean baseline values can imply the two groups have different electricity saving potentials (for instance, one group might have lower baselines since its members already have already implemented energy sufficient consumption behaviours, or adopted efficient plants and appliances, while the other group has not), which might preclude comparability between their outcomes.

Table 4.4: Comparison of baseline electricity consumption between treatment (n = 55) and potential control group (n= 163).

Baseline period (Year 0) Consumption [kWh]	Treatment group (n=55)	Potential control group (n=163)
Mean	7'168	5'968
Standard Deviation	4'572	4'923

In this case, mean yearly baseline values of the two groups, reported in Table 4.4, are statistically different at the 0.05 significance level, as indicated by a two-tailed Mann Whitney U Test (p-value= 0.02249), chosen instead of a t-test since Shapiro-Wilk tests indicate distributions of the electricity consumption variable in the two groups is not normal. The boxplot respresented in Figure 4.4 also provides a visual indication of the difference between baseline electricity consumption values in the two groups.

This additional check therefore confirms that treatment and potential control groups are not comparable regarding the amount of electricity they used during the baseline period. For this reason, I opt for not using the potential control group (n = 163) and instead exploit a matching approach, with the aim of identifying a control group with higher comparability to the treatment group.

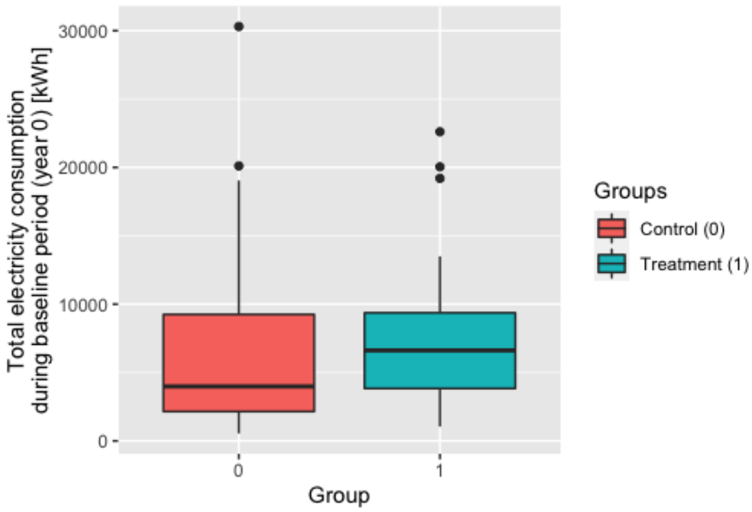

Figure 4.4: Comparison between yearly baseline electricity consumption values of potential control group (0) and treatment group (1).

4.4.2 Matching

Matching techniques are widely used in quasi-experimental research to build control groups from observed characteristics (Zhao et al., 2021), frequently coupled with regression adjustment techniques, as for instance suggested by Stuart (2010) for "double robustness" analyses. The aim of the matching I perform is to select, out of the potential control group of 163 households, a control group whose percentage composition in terms of electricity use purpose (the only covariate I can observe), is as close as possible as the treatment group's one: presence of systematic differences in covariates related to outcomes between treatment and control groups would in fact bias results.

Perfect matching requires each household in the treatment group to be matched with at least an household in the control group, which is identical on all relevant observable characteristics. When the number of covariates available for matching is very limited, stratification/complete matching techniques work satisfactorily. When instead more covariates are available, propensity score matching (PSM) procedures are advisable, since they manage to optimally deal with the different information provided by each covariate. Since in the enCompass case I can only observe the single covariate on the purpose of use of electricity, I opt for a complete matching technique, that I perform manually. Specifically, I adopt a stratification procedure that mimics the experimental designs' stratified (block) random assignment procedures as much as possible. My goal is in fact to favour as much variability as possible in unobserved characteristics of the households, such as for instance the size of the house, the number of household members, the energy efficiency of the building, the attitudes of the household members, or their constraints to daily practices.

To find the "matched control group", I adopt the following procedure:

- I stratify the treatment and potential control groups based on the covariate related to the purpose of use of electricity, thus creating four strata (blocks): appliances and lighting only; appliances and lighting + hot water heating; appliances and lighting + room heating via direct electric heating; appliances and lighting + room heating via heat pumps;
- then I select households of the potential control group that match the covariate value in the treatment group, with the aim of respecting the same stratum size percentage as in the treatment group;
- when more households are available to match households of the treatment group, I randomly select among them, with the aim of favouring as much variability as possible in unobserved covariates.

To increase statistical power of the analyses, I aim at keeping the size of the control group as large as possible, while respecting the above-mentioned size percentages. As indicated in Table 4.3, the percentage of households with heat pump in the treatment group is 40%. Since in the potential control group only 30 households with heat pump are available, this acts as a limiting factor for the size of the matched control group: such 30 households need to be about the 40% of the size of the matched control group. By opting for a matched control group size equal to 1.5 times the size of the treatment group (namely, n = 55*1.5 = 82), the percentage of heat pumps in the matched control group would be 36.6%, which is reasonably close to the 40% percentage of the treatment group.

I thus include in the matched control group all households with heat pump that are included in the potential control group. Then, since in the treatment group households with boiler are 20% of the total, I randomly select 16 households within the "boiler" stratum of the potential control group (19.4% of the n = 82 size of the matched control group). Finally, I randomly select 36 households within the "only appliances" stratum of the potential control group (44% of the n= 82 size of the matched control group). No households of the treatment group are included in the "direct electric heating" stratum, therefore no household of such a stratum has to be included in the matched control group. Overall, the composition of the resulting matched control group is represented in Table 4.5 and Figure 4.5, again in comparison with the treatment group.

Table 4.5: Characteristics of the "matched control group" (n = 82) compared with characteristics of the treatment group (n = 55).

	Appliances		Appliances + Boiler		Appliances + Electric heating		Appliances + Heat pump		Total
	Num	%	Num	%	Num	%	Num	%	Num
Treatment	22	40	11	20	—	—	22	40	55
Control	36	44	16	19	—	—	30	37	82

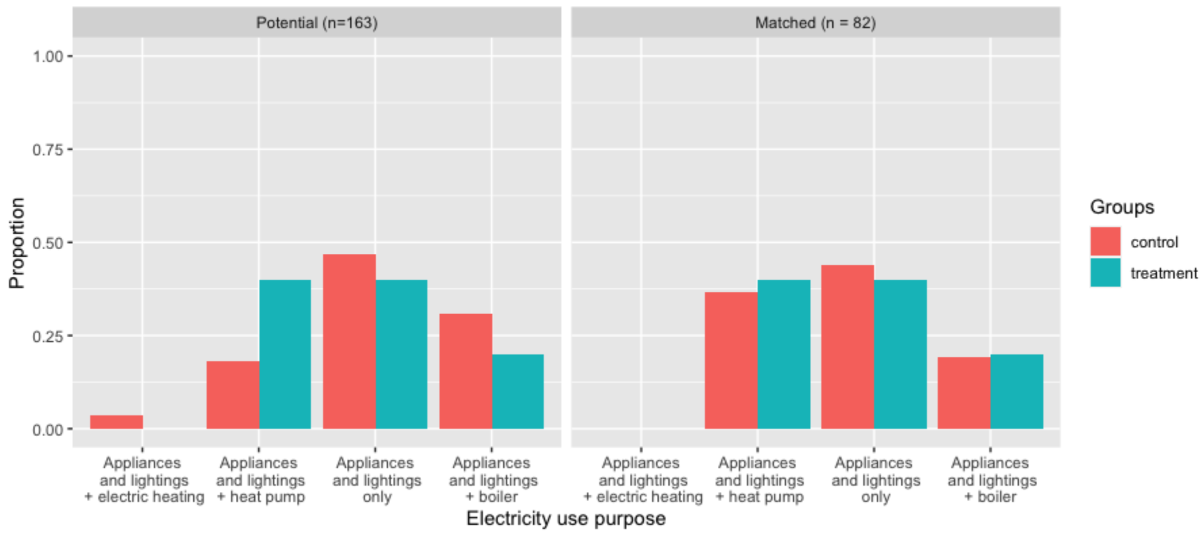

Figure 4.5: Comparison between observed characteristics of the treatment group and: potential control group (left), matched control group (right).

By design, the treatment and matched control groups are now comparable in their observed characteristic about their purpose of use of electricity. I anyway repeat the check of imbalance via the linear probability model of equation (4.1), and in fact obtain a confirmation that the two groups are now balanced. The p-value of the F-Statistics in fact in this case is equal to 0.6049, which means that the regression coefficients are not jointly statistically significant at the 10% level: the covariates predict membership to the treatment group no better than it would have been expected by chance.

Finally, I also check the values of the baseline electricity consumption (before the en-Compass intervention, year 0). Mean yearly values are represented in Table 4.6 and in Figure 4.6. Though differences in the mean baseline values between the two groups still appear, in this case a two-tailed Mann Whitney U Test (again chosen because the two distributions are not normal according to Shapiro-Wilk tests) indicates such differences are not statistically significant even at the 10% level (p-value = 0.4017).

Table 4.6: Comparison of baseline electricity consumption between treatment (n = 55) and matched control group (n= 82).

Baseline period (Year 0) Consumption [kWh]	Treatment group (n=55)	Matched control group (n=82)
Mean	7'168	6'678
Standard Deviation	4'572	4'869

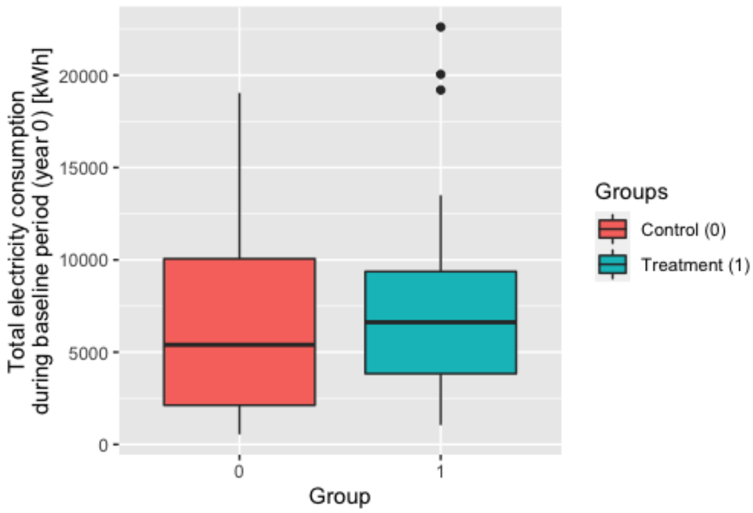

Figure 4.6: Comparison between baseline electricity consumption values of matched control group (0) and treatment group (1).

These additional checks therefore indicate that treatment group and matched control group are also comparable regarding the amount of electricity they used during the baseline period. From now onwards, therefore, this is the analytical control group sample I use in my analyses for the estimate of the average treatment effect. For the sake of simplicity, I will simply refer to it as to the control group.

4.5 Fixed effects regression modelling

To provide an estimate of the ATT and answer my research questions, following Wooldridge (2015) I 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 met[5],

[5]They 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.

Fixed Effects allow correlation between the error term and the independent variable. They in fact 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 powerful process allows to get rid of the unobserved heterogeneity between households that is constant over time. This characteristics is particularly relevant for the enCompass case, in which participating households are self-selected, and therefore it is likely that a correlation exists between the independent variable of my 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. I suspect, in particular, that households that decided to join the enCompass project 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 earnings), which makes them more inclined to actively join a research project.

The Fixed Effects estimator is therefore well-suited to the enCompass case —much more than the Random Effects, the other widely used estimator for panel data regressions, whose use is instead advised (and definitely more efficient) when one thinks the unobserved effect is uncorrelated with any of the explanatory variables. The only drawback of Fixed Effects estimators is that, during the differencing process, besides the unobserved covariates, also other independent variables that are constant over time are removed. This implies that Fixed Effects estimators do not allow to estimate the effect of observed covariates that remain constant over time, such as for instance the variable indicating the "purpose of use of electricity" that I observe in enCompass. Nevertheless, it is still possible to estimate the impact of such time constant covariates by interacting them with a time-variant variable indicating the post-treatment period.

4.5.1 The Two-Way Fixed effects estimator

Among the family of Fixed Effects estimators, I specifically opt for the Two Ways Fixed Effect (TWFE) estimator (Wooldridge, 2021), which 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 to 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 (e.g. the evolution of outdoor temperatures that drives the need for home heating of cooling). The general shape of TWFE regression models is as follows:

$$\begin{gathered} y_{it} = \boldsymbol{x}_{it}\boldsymbol{\beta} + \boldsymbol{c}_i + f_t + u_{it} \\ for\ t = 1, ...T\ and\ i = 1, ...N \end{gathered} \tag{4.2}$$

where:

- y_{it} is the observed dependent variable (here, the yearly electricity consumption, which varies between the households and over time);
- $\boldsymbol{x}_{it}\boldsymbol{\beta}$ is the vector of observed independent explanatory variables (here, a dichotomous variable indicating if the household i received the enCompass treatment at time t and the variable indicating the purpose of use of electricity, which varies among the households but is time-invariant);
- $\boldsymbol{c}_i$ 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, the "Heating degree days" or "Cooling degree days" variables that I introduce below);
- 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 4;
- 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 groups (n_T = 55 and n_C= 82, thus n= 137).

Through a panel data regression model and a TWFE estimator I can therefore estimate the causal effect of the enCompass intervention, by accounting for:

- the observed characteristics of the households (type of use of electricity);
- the unobserved variables affecting the observed behaviour (electricity 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.

4.5.2 Heating and cooling degree days

As said, besides information specifically related with the enCompass households, through the model I can also account for weather external factors that affect their electricity demand. Particularly, I consider the "Heating degree-days" and "Cooling degree-days" indicators, which are traditionally used in order to account for the different heating or cooling demands that characterise different periods, when one wants to compare them.

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 is directly proportional to the difference between a reference indoor temperature (for Switzerland set to 20 °C) and the outdoor temperature. The "heating degree-days" are 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 however the average temperature is above 12°, the number of heating degree-days for that day is 0. Heating degree-days 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, 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 region of the enCompass intervention, long historical series of daily averages can be directly downloaded from the "Osservatorio ambientale della Svizzera italiana OASI" (Environmental Observatory of Southern Switzerland) (`https://www.oasi.ti.ch/web/dati/meteo.html`, last accessed on January, 27 2023). In order to include heating and cooling degree-days into the model, I considered the station of "Magadino", which is located in the municipality of Gambarogno (the same as Contone), and computed monthly heating degree-days (HDD) and cooling degree-days (CDD) according to the SIA 381/3 standard for each month included in the enCompass analysis. Then, I aggregated each such values over the four years, thus obtaining four HDD and four CDD yearly values to be directly included in the panel regression model.

4.6 Estimate of the treatment effect

I have now completely introduced the methodology I rely upon in order to assess the effect of the enCompass intervention and can therefore present the specific analyses I performed and the related results. For all the analyses I use the R Statistical Software (v4.1.3, R Core Team, 2022) and the RStudio open source development environment (RStudio Team, 2022). Specifically, for panel regressions I use package "plm" (Croissant and Millo, 2008), while for figures I use package "ggplot2" (Wickham, 2016).

I start by research questions RQ1 and RQ2, by considering the treatment and control groups as a whole. For the enCompass case, I adopt the following specific formulations of RQ1 and RQ2, and state the related null hypotheses (H_0) for hypothesis testing as follows:

- Research question RQ1:
  - What is the average treatment effect on the treated (ATT) during the one-year period of use of the enCompass app (Treatment period, Year One), in terms of savings of electricity and CO_2 emissions?
  - H_0: the ATT in Year One is equal to zero;
- Research question RQ2:
  - What is the average treatment effect on the treated (ATT) over the two one-year periods after use of the the enCompass app (Post-treatment periods, Years Two and Three), in terms of savings of electricity and CO_2 emissions?
  - H_0: the ATT in Year Two is equal to zero;

- H_0: the ATT in Year Three is equal to zero.

In both cases, for hypothesis testing I use two-tailed tests. To estimate the ATT, I compute the following TWFE panel regression model:

$$\begin{aligned} ln\_kWh_{it} = {} & \alpha_1 Treat_i + \\ & \beta_1 Year\_One_t + \beta_2 Year\_Two_t + \beta_3 Year\_Three_t + \\ & \gamma_1 TreatxYear\_One_{it} + \gamma_2 TreatxYear\_Two_{it} + \gamma_3 TreatxYear\_Three_{it} + \\ & \delta_1 HDD_t + \delta_2 CDD_t + \\ & \boldsymbol{c_i} + u_{it} \\ & for\ t = 1, ...4\ and\ i = 1, ...137 \end{aligned} \tag{4.3}$$

where:

- ln_kWh_{it} is the observed dependent variable, namely the natural logarithm of the yearly electricity consumption. I use the logarithmic form in order to directly obtain the ATT, which is represented by γ_i coefficients, in terms of percentage electricity savings; note that, since I estimate CO_2 emissions via a multiplication factor starting from electricity consumption, the electricity saving percentage provided by the model is directly also a percentage of saved CO_2 emissions;
- $Treat_i$ is the observed independent variable characterising each household, namely a dichotomous variable indicating if household i received the enCompass treatment (the household is member of the treatment group) or not (the household is member of the control group);
- $Year\_One_t$, $Year\_Two_t$, $Year\_Three_t$ are three dichotomous variables respectively indicating the period. Note that an additional dichotomous variable $Year\_Zero_t$ is also defined, though not explicitly included in the model to avoid collinearity; indeed, it acts as a reference variable;
- $TreatxYear\_One_{it}$, $TreatxYear\_Two_{it}$ and $TreatxYear\_Three_{it}$ are three dichotomous interaction terms indicating if household i is treated and if its electricity consumption respectively refers to each of the years t. Namely, they take on value "1" if the household is treated and its electricity consumption is respectively related to Year One, Year Two, or Year Three; otherwise, they take on value "0";
- HDD_t and CDD_t respectively measure the amount of the Heating degree-days and Cooling degree-days recorded in each year;
- $\boldsymbol{c_i}$ is the vector of unobserved time-invariant household-specific effects;
- u_{it} is the unobserved idiosyncratic error term.

Overall, through model (4.3) I can directly estimate the average treatment effect on the treated, for each of the three years I am interested into: the design of the model equation is such that consumption of the control group in year 0 is taken as a reference and that the ATT values respectively over Year One, Year Two and Year Three directly correspond to the value of the γ_i coefficients, which represent the percentage electricity saving over the related year (on a 0-1 scale).

Note, I do not include interaction terms for the two variables heating and cooling degree-days (HDD and CDD). I have in fact included them in earlier versions of the model (via $TreatxYear\_OnexHDD_{it}$, $TreatxYear\_TwoxHDD_{it}$, $TreatxYear\_ThreexHDD_{it}$ and corresponding interaction terms for CDD), though they were automatically removed due to collinearity reasons. Therefore, for the sake of simplicity I do not include them in the model I report here. Since values of the HDD and CDD variables are constant at each time stage t for each household i, this implies that the TWFE model does not produce estimates of their coefficients.

The outcome of panel model regression, that I refer to as "Model I" in order to facilitate comparison with other models I present in the next sections, is reported in Table 4.7, which in brackets shows heteroskedasticity-robust clustered standard errors computed according to the Arellano method (Millo, 2017), and in equation (4.4), which shows in bold format the parameter estimates that are statistically significant at the 0.05 level.

$$\begin{aligned} \widehat{ln\_kWh} = &\mathbf{-0.0508}\ TreatxYear\_One + \\ &0.0016\ TreatxYear\_Two + 0.0046\ TreatxYear\_Three \end{aligned} \tag{4.4}$$

Table 4.7: Output of Model (I) panel regression.

	Model (I)	
	ln(kWh)	p value
$TreatxYear\_One$	-0.0508 ** [0.0253]	0.04662
$TreatxYear\_Two$	0.0016 [0.0498]	0.97360
$TreatxYear\_Three$	0.0046 [0.0468]	0.92130
Observations	n=137, balanced panel. T=4, N=548.	—
Total Sum of Squares	13.455	—
Residual Sum of Squares	13.386	—
Adjusted R-Squared	-0.34364	—
R-Squared	0.0051641	—
F-statistic	0.70077 on 3 and 405 degrees of freedom.	0.55202

R "plm" package; model= "within", effect= "twoways".
Heteroskedasticity-robust, clustered standard errors (Arellano method) in parenthesis.
Significance levels: * 0.1, ** 0.05, *** 0.01

The parameter estimate of coefficient γ_1 is significant at the 0.05 significance level. This means I can reject at the 0.05 significance level the null hypothesis H_0 related with RQ1.

I can thus conclude that the enCompass ATT in Year One is actually different from zero. Following Wooldridge (2015), the precise coefficient estimate can be computed as:

$$ATT\ Year\_One = 100(exp(\gamma_1) - 1) = 100(exp(-0.0508) - 1) = -4.95\% \quad (4.5)$$

In the short term (Year One), use of the enCompass app produced on average 4.95% electricity and CO_2 emission savings compared to the previous year. In order to understand the amount of such an effect, I computed the effect size, adopting the Cohen's d approach (Cohen, 2013). The result is a small effect size, equal to -0.35.

Instead, the parameter estimates of coefficients γ_2 and γ_3 are not significant even at the 0.1 significance level. This means that I cannot reject the two null hypotheses H_0 related with RQ2, stating that the enCompass ATT in Year Two and in Year Three is different from zero. The precise estimates of the coefficient, as well as the related effect size, are reported in Table 4.8, which summarizes results by Model I dealing with RQ1 and RQ2.

Table 4.8: Summary of Model (I) outcome addressing RQ1 and RQ2 for enCompass.

Model (I)	Year One	Year Two	Year Three
ATT [% change]	**- 4.95%****	+ 0.16%	+ 0.46%
ATT's Standard error	[0.0253]	[0.0498]	[0.0468]
ATT's 95% Confidence Interval	[-9.901, +0.009]	[-9.601, +9.921]	[-8.732, +9.652]
Effect size (Cohen's d)	- 0.35 (small)	- 0.09 (negligible)	- 0.05 (negligible)

Significance levels: * 0.1, ** 0.05, *** 0.01

The above evidence shows that during the intervention period (in Year One), the households treated with the enCompass app decreased their electricity consumption and CO_2 emissions by 4.95%, compared with the previous year and by accounting for a comparable control group.

The effect of the intervention was however not maintained over time: based on data collected during follow-up Year Two and Year Three, in fact, no statistically significant conclusions can be drawn about the effect of the enCompass treatment in the long term. By considering the practical significance of the outcomes of the panel regression model, however, it appears that treated households, always compared with their baseline before the enCompass intervention and by accounting for a comparable group, even slightly increased their consumption, though negligibly.

4.7 Heterogeneity on varying electricity use purpose

The above results indicate that use of the enCompass app produces an energy and CO_2 saving effect, tough only in the short term. Tackling research question RQ3, I now aim at analysing the heterogeneity of such effects, in order to understand whether effects of the same intensity are produced in all types of households, or if instead they change on varying the household's characteristics I can observe.

For such an heterogeneity analysis, two possibilities are available. The first one is to consider variable "purpose of electricity use" and to create three sub-groups of households, based on the category they fit in: households using electricity only for lighting and appliances ("only appliances" households, $n_T = 22$, $n_C = 36$), households using electricity also for hot water purposes ("boiler" households, $n_T = 11$, $n_C = 16$), and households using electricity also for room heating purposes ("heat pump" households, $n_T = 22$, $n_C = 30$).

An alternative possibility would be to create sub-groups of households based on their baseline level of electricity consumption. One might in fact assume that, depending on their initial amount of electricity consumption, households have different room for change and reduction of their consumptions. Categorisation in groups is not as immediate as the previous case, since electricity consumption is a continuous variable. I however performed an exploratory clustering of the enCompass households based on their level of baseline electricity consumption, by means of a k-means clustering algorithm (see the next Section). The outcome of such a clustering showed strong overlapping with the classification of households based on variable "purpose of electricity use". Indeed, there is very high correlation between the use of electricity for hot water and heating purposes and the amounts of electricity consumed. By means of such an exploratory analysis, therefore, I conclude that clustering based on the baseline level of electricity consumption cannot provide relevant additional insights to the simple categorisation based on the purpose of electricity use. Therefore, I focus my heterogeneity analyses on the three sub-groups of "only appliances", "boiler", and "heat pump" households.

For these analyses, the specific research question RQ3 can be formulated as: for each of the one-year periods related with the enCompass app (Intervention period - Year One, and Post-treatment periods - Years Two and Three), does the average treatment effect on the treated (ATT), measured in terms of electricity and CO_2 emissions savings, differ between the three sub-groups? Namely, to tackle RQ3 I cross-compare the ATTs' amount, across the above three sub-groups.

For this purpose, for each of such sub-groups I run another Two-Way Fixed Effects panel regression model, with the aim of obtaining the ATT within the sub-group, again for the three years under analysis. The null hypotheses H_0 I consider for each of such models are, again, that the resulting ATT computed for Year One, Year Two, and Year Three are equal to zero. And, again, to test such hypotheses I use two-tailed tests. This is however not sufficient to state if the ATT is actually different between the considered sub-groups (Assmann et al., 2000). An additional testing stage is in fact needed, which consists in *post-hoc* tests for statistical interaction that perform pairwise comparisons of the three sub-groups. For this purpose, I follow Christensen et al. (2021) and manually perform the tests, based on the ATTs computed for each sub-group and the corresponding standard errors.

4.7.1 Estimate of the treatment effect on varying electricity use purpose

The panel regression models I use coincide with TWFE model of equation (4.3). Simply, they are computed over different sub-samples of households. To distinguish their outputs, I refer to them as to Model (II). The results are reported in Table 4.9 and in equations (4.6), (4.7), and (4.8).

$$\begin{aligned} \widehat{ln\_kWh\_Only\_appliances} = & \mathbf{-0.156}\ TreatxYear\_One + \\ & - 0.032\ TreatxYear\_Two - 0.040\ TreatxYear\_Three \end{aligned} \tag{4.6}$$

$$\begin{aligned} \widehat{ln\_kWh\_Boiler} = & 0.062\ TreatxYear\_One + \\ & 0.081\ TreatxYear\_Two + 0.094\ TreatxYear\_Three \end{aligned} \tag{4.7}$$

$$\begin{aligned} \widehat{ln\_kWh\_Heat\_pump} = & 0.003\ TreatxYear\_One + \\ & - 0.014\ TreatxYear\_Two - 0.011\ TreatxYear\_Three \end{aligned} \tag{4.8}$$

Table 4.9: Output of Model (II) panel regression.

	Model (II)					
	ln(kWh)					
	Only appliances households	p value	Households with boiler	p value	Households with heat pump	p value
$TreatxYear\_One$	-0.156247 *** [0.045377]	0.00073	0.061708 [0.047339]	0.1964	0.0028829 [0.0280729]	0.9183
$TreatxYear\_Two$	-0.031600 [0.097782]	0.74697	0.081498 [0.073611]	0.2718	-0.0142901 [0.0507043]	0.7785
$TreatxYear\_Three$	-0.039765 [0.083437]	0.63427	0.093790 [0.093724]	0.3202	-0.0111437 [0.0481575]	0.8173
Observations	n=58, balanced panel. T=4, N=232.	—	n=27, balanced panel. T=4, N=108.	—	n=52, balanced panel. T=4, N=208.	—
Total Sum of Squares	9.4342	—	1.7104	—	1.6443	—
Residual Sum of Squares	9.2425	—	1.6765	—	1.6416	—
Adjusted R-Squared	-0.34706	—	-0.39837	—	-0.37777	—
R-Squared	0.020323	—	0.019835	—	0.0016176	—
F-statistic	1.16171 on 3 and 168 degrees of freedom.	0.32603	0.505918 on 3 and 75 degrees of freedom.	0.67938	0.0810128 on 3 and 150 degrees of freedom.	0.97025

R “plm” package; model= “within”, effect= “twoways”.
Heteroskedasticity-robust, clustered standard errors (Arellano method) in parenthesis.
Significance levels: * 0.1, ** 0.05, *** 0.01

Table 4.10 summarizes the results, by computing the precise ATTs based on formula of Equation (4.5), reporting the 95% Confidence Intervals and estimating the effect sizes using Cohen's d estimator. Table 4.11 then complements such results, by presenting the outcomes of the *post-hoc* interaction tests performing pairwise comparisons between the three considered sub-groups of households.

Table 4.10: Summary of Model (II) outcome addressing RQ3 for enCompass: sub-group analysis based on variable "purpose of use of electricity".

Model (II)		Year One	Year Two	Year Three
A. Only appliances households ($n_T = 22$, $n_C = 36$)	ATT [% change]	**- 14.46%*******	- 3.11%	-3.90%
	95% Confidence Interval	[-23.557, -5.362]	[-22.714, +16.444]	[-20.628, +12.828]
	Effect size (Cohen's d)	- 0.91 (large)	- 0.30 (small)	- 0.22 (small)
B. Households with boiler ($n_T = 11$, $n_C = 16$)	ATT [% change]	+ 6.36%	+ 8.49%	+ 9.83%
	95% Confidence Interval	[-3.433, +16.153]	[-6.734, +23.718]	[-9.558, +29.218]
	Effect size (Cohen's d)	+ 0.35 (small)	+ 0.23 (small)	+ 0.25 (small)
C. Households with heat pump ($n_T = 22$, $n_C = 30$)	ATT [% change]	+ 0.29%	- 1.42%	- 1.11%
	95% Confidence Interval	[-5.354, +5.934]	[-11.615, +8.766]	[-10.793, +8.573]
	Effect size (Cohen's d)	+ 0.07 (negligible)	+ 0.03 (negligible)	- 0.07 (negligible)

Significance levels: * 0.1, ** 0.05, *** 0.01

Table 4.11: Results of *post-hoc* interaction tests on Model (II) outcomes: p-values for the comparison of effects between sub-groups of households based on variable "purpose of use of electricity".

Model (II)	Year One	Year Two	Year Three
A. Only appliances households vs B. Households with boiler	**0.001498*****	0.343246	0.273879
A. Only appliances households vs C. Households with heat pump	**0.005704*****	0.878057	0.772116
B. Households with boiler vs C. Households with heat pump	0.270073	0.267560	0.299166

Significance levels: * 0.1, ** 0.05, *** 0.01

Overall, for Year One and for the subgroup of households using electricity for appliances and lighting only, an electricity saving and CO_2 saving effect emerges and thus I can reject the null hypothesis H_0. The interaction tests confirm that the effect on this sub-group is actually different from the effects on the other two sub-groups (households with boiler and households with heat pump), at the 0.01 significance level. Also, this effect is very relevant, in terms of both its practical and statistical significance. The amount of savings is in fact equal to 14.46%, compared with the baseline period Year Zero and accounting for the comparable control group of non-treated households using electricity only for lighting and appliances. The effect size, measured through the Cohen's d, is very large, equal to -0.91. Statistical significance is also very high, since the parameter estimates of coefficients γ_1 measuring the ATT is significant at the 1% significance level, and the same holds for p-values resulting from the interaction tests.

For Years Two and Three, instead, no statistically significant results are found, even at the 0.1 significance level: the related null hypotheses H_0 cannot be rejected. Also, the interaction tests indicate that the differences between the three groups are not significant at the 0.1 level. Finally, Cohen's d effect sizes are always small, or even negligible. According to this heterogeneity analysis, therefore, in years Two and Three, no differences in the effects on sub-groups of treated households emerge. Overall, I can conclude that in Years Two and Three the enCompass intervention had no statistically significant effects, not only at the level of the whole sample of treated households, but also within sub-groups of them, that use electricity for different purposes.

4.8 Heterogeneity on varying app use

As an additional analysis on the enCompass case, I consider the information provided by the available data on the level of app use, which was provided by the enCompass platform managers at Politecnico di Milano. As mentioned in Section 4.3, for each household of the treatment group I in fact know the monthly number of logins and interactions in the enCompass app. Through detailed app-analytics systems, platform managers could compute the amount of access to the following app features:

- the goal setting and energy saving feedback page (left-hand side of Figure 4.1);
- the energy consumption plots (right-hand side of Figure 4.1);
- the impact and comfort feedback page (Figure 4.2);
- the comfort page (Figure 4.3);
- the tips page;
- the achievements page, showing all the obtained badges;
- the leaderboard page.

The number of monthly app interactions represents the total number of any of such activities. Accounting for both logins and app interactions allows to account for the frequency of interactions with the app over the one-year intervention, as well as for their intensity, represented by the amount of exploited features. By relying on this additional dataset, I can thus verify if the electricity and CO_2 saving effect changes on varying the level of app use. Specifically, I aim at verifying whether differences in the ATT emerge, between sub-groups of very active and very inactive app users, which I suppose to represent exceptions in app use patterns, and the remaining sub-group(s) of intermediate app users, which I suppose to represent the most common app use pattern.

For this purpose, I perform a clustering analysis on the treated households based on the two observed variables "number of yearly app logins" and "number of yearly in-app interactions", which I obtain by aggregating (summing) the twelve monthly observations available for both pieces of information (app logins and app interactions). Once the clusters are available, I run again the same panel regression model as above, though this time on each of the sub-groups of household that are included in each cluster. This allows me to explore the heterogeneity of the effects of the enCompass intervention from another perspective, therefore offering an additional contribution to tackle research question RQ3.

For Year One, I expect to find negligible ATT for the very low app user households, while for intermediate user households I expect a large ATT. For the very high app user households, I do not exclude a saturation effect might occur: while I expect a large ATT as well, I suppose it to be similar as the one found in intermediate user households. For the following years Two and Three, I expect instead to find a statistically different from zero electricity and CO_2 saving ATT only for very active households, whose energy behaviour and practices might have been permanently affected by an intense use of the app's features and the underlying behaviour change processes. Based on the results for the whole group of treated households, I instead expect such an ATT not to be present among the very low and intermediate activity households.

4.8.1 Clustering

Cluster analysis aims at grouping subjects of analysis into homogeneous groups (clusters), so that those allocated to the same cluster are as similar as possible between each other according to a given measure of distance (internal homogeneity), and as different as possible from subjects allocated to other clusters, according to the same measure of distance (external separation). Patterns of subjects in the same cluster are similar to each other, while patterns of subjects attributed to different clusters are not.

Many clustering techniques and related algorithms are available, which are usually classified into the main categories of "hierarchical" and "partitional" clustering techniques (Xu and Wunsch, 2005). For my analysis I opt for the *k-means* technique, which is a "partitional" technique and one of the most used ones due to its simplicity. The k-means clustering technique divides the subjects into a "k" pre-defined number of distinct, non-overlapping clusters, by relocating each subject from one cluster to another via an iterative process, that starts from an initial partitioning (Hartigan and Wong, 1979).

Since there is no optimal prior number of clusters, for my analysis I consider a number of clusters from two to six, and explore the results produced by the k-means algorithm, implemented through the "cluster" package in R (Maechler et al., 2021). The results are reported in Table 4.12 and in Figure 4.7 (plots produced with "factoextra" R package by Kassambara and Mundt, 2020, in addition to "ggplot2" package).

Table 4.12: Summary of clustering output by the k-means algorithm.

Number of clusters	Within cluster sum of squares (SS)	Between SS/Total SS [%]
2	28.75, 14.25	65.3
3	28.75, 2.40, 2.57	72.8
4	28.75, 0.42, 0.58, 1.98	74.4
5	1.98, 0.03, 0.12, 0.45, 28.75	74.7
6	1.98, 0.03, 0.06, 0.11, 28.75, 0.11	75.0

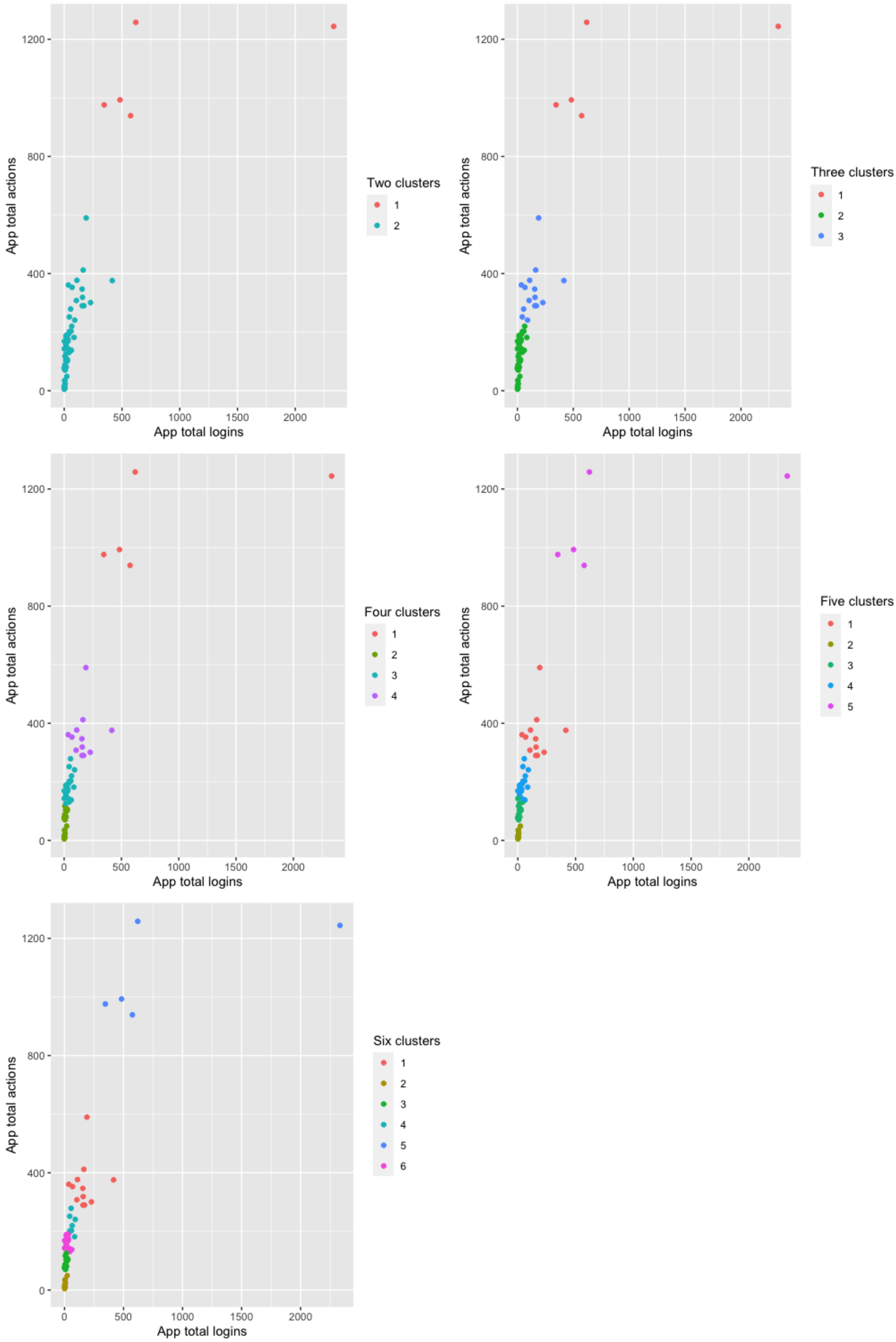

Figure 4.7: Clustering of the n= 55 treated households produced by the k-means algorithm, respectively with 2, 3, 4, 5, and 6 clusters.

Among the five clustering options, the classification in five clusters (Figure 4.8) allows to clearly identify the very low (cluster 2) and very high app users (cluster 5). The three clusters lying in between (clusters 1, 3, and 4) can be regarded as the "intermediate level of activity" users, which I suppose represent the most common type of app users. By re-aggregating households of cluster 1, 3, and 4 into a single sub-group, I can thus obtain three sub-groups of households, respectively characterised by negligible (cluster 2), very high (cluster 5), and intermediate level of app use (clusters 1, 3, and 4). I believe this threefold classification starting from a five-cluster classification suits better my goal of breaking-down app users based on their level of in-app activity, compared with the direct three-cluster classification reported in Figure 4.7, since the former offers higher granularity for the identification of the extreme cases. The latter instead does not manage to properly isolate the negligible activity households, which are put in the same cluster as low activity households.

I thus start from the five-cluster classification and then re-aggregate the intermediate clusters, obtaining three sub-groups of households based on their in-app level of activity: n= 9 households (cluster 2) are classified as "negligible app use", n= 41 households (clusters 1, 3, and 4) as "intermediate app use", and n= 5 households (cluster 5) as "very high app use".

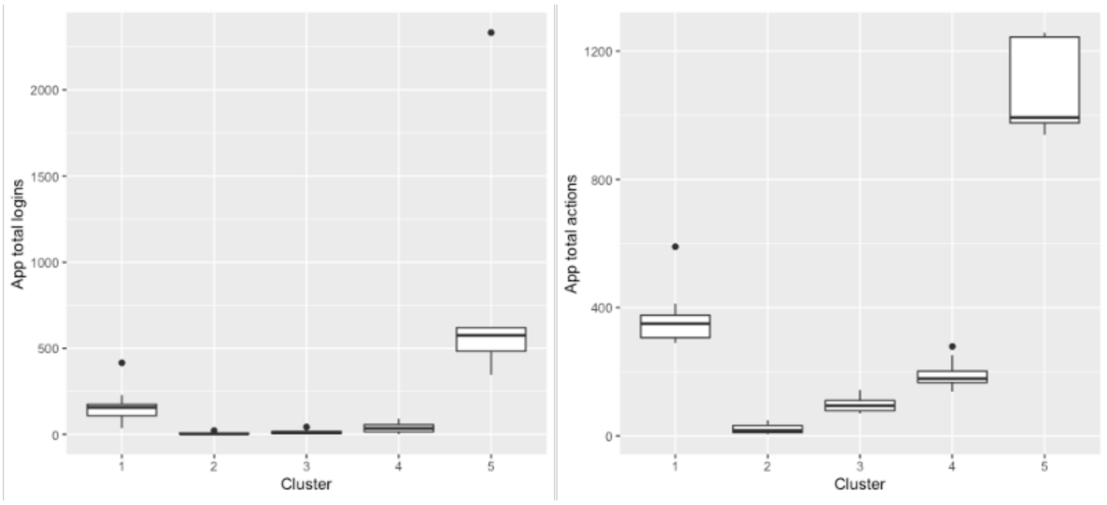

Figure 4.8: Characteristics of clusters of treated households under the five-cluster classification based on the number of app logins and of app actions.

4.8.2 Estimate of the treatment effect on varying app use

Having identified three sub-groups of treated households, I now feed them into three regression models, with the aim of estimating the enCompass effect in each sub-group. For this purpose I use the Two-Way Fixed Effects model of Equation (4.3) to perform three regressions on each of the household sub-groups. As a comparison, in each case I consider all the n= 82 households of the control group.

For two-tailed testing of the regression outcome, I consider the same null hypotheses H_0 as the previous cases: in each of the considered sub-groups of the treated households, the

ATT respectively computed for Year One, Year Two, and Year Three is equal to zero. Like the previous analyses of heterogeneity, also in this case I use outcomes of the regression to feed *post-hoc* interaction tests, to assess whether the obtained differences in the ATTs between sub-groups are statistically significant.

Outcomes of the set of TWFE regressions, which I refer to as to Model (III), are reported in Table 4.13 and in Equations (4.9), (4.10), and (4.11) (with estimates of the coefficients that are statistically significant at the 0.05 level of significance reported in bold).

$$\begin{aligned} \widehat{ln\_kWh\_Very\_Low} = & -0.079\ TreatxYear\_One + \\ & +0.006\ TreatxYear\_Two + 0.023\ TreatxYear\_Three \end{aligned} \tag{4.9}$$

$$\begin{aligned} \widehat{ln\_kWh\_Intermediate} = & \mathbf{-0.035}\ \ TreatxYear\_One + \\ & 0.002\ TreatxYear\_Two - 0.003\ TreatxYear\_Three \end{aligned} \tag{4.10}$$

$$\begin{aligned} \widehat{ln\_kWh\_Very\_High} = & -0.129\ TreatxYear\_One + \\ & -0.010\ TreatxYear\_Two + 0.033\ TreatxYear\_Three \end{aligned} \tag{4.11}$$

Table 4.13: Output of Model (III) panel regression.

	Model (III)					
	ln(kWh)					
	Negligible app use households	p value	Intermediate app use households	p value	Very high app use households	p value
$TreatxYear\_One$	-0.0795266 [0.0533563]	0.1373	-0.0349526 [0.0274025]	0.2029	-0.129182** [0.059168]	0.0299
$TreatxYear\_Two$	0.0056848 [0.0523273]	0.9136	0.0022382 [0.0539968]	0.9670	-0.010391 [0.117195]	0.9294
$TreatxYear\_Three$	0.0233888 [0.0870087]	0.7883	-0.0029135 [0.0496039]	0.9532	0.032815 [0.111052]	0.7679
Observations	n=91, balanced panel. T=4, N=364.	—	n=123, balanced panel. T=4, N=492.	—	n=87, balanced panel. T=4, N=348.	—
Total Sum of Squares	11.157	—	12.668	—	10.97	—
Residual Sum of Squares	11.106	—	12.643	—	10.879	—
Adjusted R-Squared	-0.35336	—	-0.34994	—	-0.3520	—
R-Squared	0.00456	—	0.00198	—	0.0065	—
F-statistic	0.407 on 3 and 267 degrees of freedom.	0.7479	0.240 on 3 and 363 degrees of freedom.	0.8683	0.553 on 3 and 255 degrees of freedom.	0.6464

R “plm” package; model= “within”, effect= “twoways”. Heteroskedasticity-robust, clustered standard errors (Arellano method) in parenthesis. Significance levels: * 0.1, ** 0.05, *** 0.01

Table 4.14 summarizes the results, by computing the precise ATTs based on formula of Equation (4.5), reporting the 95% Confidence Intervals, and providing estimates of the effect sizes using Cohen's d estimator. Table 4.15 then complements such results, by presenting the outcome of the *post-hoc* interaction tests (pairwise comparisons between the three considered sub-groups of households).

Results are indeed different that what I expected: despite for Year One the ATT on the (very small indeed) sub-group of very high app users is statically different from zero at the 0.05 level, and is also characterised by a very large effect size (-0.87), the interaction tests show that there are no statistically significant differences in the ATT between the three subgroups of households, even in Year One. Though the lack of statistical significance might also be due to the very low size of the sub-groups (the very active and negligibly active households are respectively only n= 5 and n= 9), these results indicate that, no matter the level of intensity of the interaction with the enCompass app, the same effect is produced on electricity and CO_2 savings. Furthermore, in any case savings only occur in Year One. The effect is not maintained and, as time goes by, consumption (and hence emissions) reverts to the baseline value.

Table 4.14: Summary of Model (III) outcome addressing RQ3 for enCompass: sub-group analysis based on level of app use.

Model (III)		Year One	Year Two	Year Three
D. Negligible app use households ($n_T = 9$, $n_C = 82$)	ATT [% change]	- 7.65%	+ 0.57%	+ 2.37%
	95% Confidence Interval	[-18.255, +2.955]	[-9.831, +10.971]	[-14.923, +19.664]
	Effect size (Cohen's d)	- 0.53 (medium)	- 0.14 (small)	0.02 (negligible)
E. Intermediate app use households ($n_T = 41$, $n_C = 82$)	ATT [% change]	- 3.43%	+ 0.22%	- 0.29%
	95% Confidence Interval	[-8.856, 1.996]	[-10.472, +10.912]	[-10.112, +9.532]
	Effect size (Cohen's d)	- 0.25 (small)	- 0.07 (negligible)	- 0.08 (negligible)
F. Very high app use households ($n_T = 5$, $n_C = 82$)	ATT [% change]	**-12.12%****	- 1.03%	+ 3.36%
	95% Confidence Interval	[-23.888, -0.352]	[-24.340, +22.280]	[-18.728, +25.370]
	Effect size (Cohen's d)	- 0.87 (large)	- 0.10 (small)	+ 0.05 (negligible)

Significance levels: * 0.1, ** 0.05, *** 0.01

Table 4.15: Results of *post-hoc* interaction tests on Model (III) outcomes: p-values for the comparison of effects between household sub-groups based on their level of app use.

Model (III)	Year One	Year Two	Year Three
D. Negligible vs E. Intermediate app use households	0.4817136	0.9163245	0.7905566
D. Negligible vs F. Very high app use households	0.5747668	0.9007907	0.9440553
F. Intermediate vs F. Very high app use households	0.1826267	0.922828	0.7641023

Significance levels: * 0.1, ** 0.05, *** 0.01

4.9 Conclusions

In this Chapter I presented the persuasive and gamified enCompass app, by framing its features according to a stage model of behaviour change and principles for persuasive systems design. I also introduced the enCompass policy intervention, which took place in the Swiss small village of Contone between June 2018 and May 2019. Finally, I presented the novel data collection procedure and analyses I performed in order to assess the effect of the enCompass intervention, together with the results I obtained.

By adopting a quasi-experimental panel data approach, I identified a comparable control group of households to the self-selected ones that were treated with the enCompass app, and estimated the average treatment effect on the treated (ATT) of the enCompass intervention, both during the intervention itself and in the following two years. For this purpose, I collected four full years of electricity consumption data by the households of the treatment and control groups: the consumption over the year before the intervention (Baseline), the consumption during the enCompass intervention (Year One), and the consumption for the two years after the end of the intervention (Year Two and Year Three). With the collected data I fed three Two Way Fixed Effects panel regression models, which allowed me to answer my three research questions about the effects of the enCompass intervention in the short and long term, as well as to look for differences in the effects on varying the observed characteristics of sub-groups of intervention households (heterogeneity analysis).

Table 4.16 provides a summary visualisation of the results of the three models I developed, by reporting the ATTs and the effect sizes for each model and sub-group of households considered for the heterogeneity analyses.

According to the results by Model (I), a significant ATT of the enCompass intervention, at the 0.05 significance level, is only found during the intervention itself (Year One); then, over time it disappears. On average, in Year One the enCompass treated households decreased their electricity consumption and $CO2_2$ emissions by 4.95%, compared with the previous year (Baseline) and comparable untreated households represented by the control group. These saving percentages are similar to those by early smart meter feedback studies reviewed by Darby et al. (2006), Fischer (2008) or Delmas et al. (2013). More specifically, these results are similar to those that the latter authors associate with weaker studies from the methodological point of view —even though in this case I devoted particular care to ensure that the whole assessment procedure respects strict methodological criteria.

By considering the observable heterogeneous characteristics of the households included in the enCompass intervention, namely the purpose of use of electricity, a larger electricity and CO_2 saving effect (14.5% decrease, compared with the Baseline year and comparable untreated households) is found on households that use electricity only for appliances and lighting (Model II). The effect on this household sub-group is statistically different from the effects on the other two sub-groups of households (those using electricity respectively

Table 4.16: Summary of the three model outcomes estimating the average treatment effect on the treated (ATT) of the enCompass intervention.

		Year One (intervention)		Year Two (One year after)		Year Three (Two years after)	
		ATT (Electricity consumption) [% change/year]	Effect size (Cohen's d)	ATT (Electricity consumption) [% change/year]	Effect size (Cohen's d)	ATT (Electricity consumption) [% change/year]	Effect size (Cohen's d)
Model (I)	All households ($n_T = 55$, $n_C = 82$)	**-4.95****	-0.35 small	+0.16	-0.06 negligible	+0.46	-0.05 negligible
Model (II)	A. Only appliances households ($n_T = 22$, $n_C = 36$)	**-14.46*****	-0.91 large	-3.11	-0.30 small	-3.90	-0.22 small
	B. Households with boiler ($n_T = 11$, $n_C = 16$)	+6.36	+0.35 small	+8.49	+0.23 small	+9.83	+0.25 small
	C. Households with heat pump ($n_T = 22$, $n_C = 30$)	+0.29	+0.07 negligible	-1.42	+0.03 negligible	-1.11	-0.07 negligible
	Post-hoc interaction tests between sub-groups A, B, C	**A vs B** *** **A vs C** ***	—	No signif.	—	No signif.	—
Model (III)	D. Negligible app use households ($n_T = 9$, $n_C = 82$)	-7.65	-0.53 medium	+0.57	-0.14 small	+2.37	+10.02 negligible
	E. Intermediate app use households ($n_T = 41$, $n_C = 82$)	-3.43	-0.25 small	+0.22	-0.07 negligible	-0.29	-0.08 negligible
	F. Very high app use households ($n_T = 5$, $n_C = 82$)	**-12.12****	-0.87 large	-1.03	-0.10 small	+3.36	+0.05 negligible
	Post-hoc interaction tests between sub-groups D, E, F	No signif.	—	No signif.	—	No signif.	—

Significance levels: * 0.1, ** 0.05, *** 0.01

also for water and for space heating). Again, however, the effect is only found in Year One and no statistically significant differences are found between the sub-groups in Years Two and Three, even at the 0.1 significance level.

By considering all treated households, the effect size found for the ATT is small (Cohen's d equal to 0.35); however, by accounting for heterogeneity within the sub-groups, for the sub-group in which statistical significance of the effect is found ("only appliances households" in Year One, significant at the 0.05 level), the effect size is large. This shows that, in the households in which the enCompass intervention managed to produce an effect, such an effect was tangible and relevant —though transient over time.

Finally, by considering heterogeneous effects on varying the level of in-app activity by treated households as done in Model (III), for the very active households a statistically significant effect is found (12.1% decrease, compared with the Baseline year and all the untreated households). However, *post-hoc* interaction tests show this effect is not different from the effects on the households with very low or intermediate levels of in-app activity, even at the 0.1 significance level. Furthermore, also Model (III) shows no significantly different effects in Years Two and Three between the app-use sub-groups. I thus conclude that the intensity of use of the enCompass app is not a driver for different magnitude electricity and CO_2 saving effects: the amount of electricity and CO_2 savings is not influenced by the intensity of app use.

These specific results suggest that the saving effect of the enCompass app might not be due to its specific features, but simply to the existence of the app itself and its availability for the household's use. Framing the behaviour change process from the perspective of the theory of Planned Behaviour, on which the app's features were grounded, this suggests that the very fact of being part of the enCompass intervention and potentially receiving the support by the app, might have temporarily increased either their perceived behavioural control over electricity saving activities or the households' subjective norms about energy saving, thus leading to tangible saving results. Once the app use was discontinued, these feelings progressively disappeared and households relapsed to their previous behaviours. From the perspective of stage models of behaviour change, which are the other theoretical reference of the enCompass app, maintenance of the new behaviour was not effective, and relapse took place, gradually leading app user households to go back to their previous energy consuming behaviour and routines.

By focusing on Year One, for which a statistically significant effect was found, the heterogeneity analyses based on the purpose of electricity use suggest that saving electricity by changing the way appliances and lighting are used is easier than saving electricity that is used for heating or hot water purposes. I cannot even exclude that, despite the goal of enCompass was to promote energy sufficient consumption behaviour, the observed savings were obtained by replacing high electricity consumption appliances, such as for instance the fridge or the washing machine, exactly during the year of the enCompass intervention, maybe also due to the increase in energy and environmental awareness stimulated by the intervention itself.

Overall, the enCompass findings suggest that effectively intervening on energy consumption to satisfy hot water and heating demand may require different types of interventions, that are more deeply entrenched with daily practices about personal hygiene and home living, as well as with conventions and social norms about cleanliness and comfort, as for instance suggested by Social Practice Theories. Future research would therefore benefit by interventions that, while targeting a larger sample of households equipped with heat pumps or boilers, in order to verify whether the lack of statistical significance in enCompass is mostly due to the very small sample sizes of each sub-group, also aim at specifically addressing heating-related social norms and conventions.

Particularly, I expect that addressing such conventions could also favour the consolidation of new behaviour over time, thus addressing the primary limitation emerging from the enCompass case, and contributing to the long-term maintenance of the beneficial effects that the use of the app produced in Year One.

5 Case two: Social Power

> *Nobody realizes that some people expend tremendous energy merely to be normal.*
>
> — **Albert Camus**
> Writer

The Social Power project[1] was performed during years 2015-2018 by an inter-disciplinary research team involving two Swiss universities of applied sciences (SUPSI, the University of Applied Sciences of Southern Switzerland, and ZHAW, the Zurich University of Applied Sciences), two software and data analysis private companies, and the two utility companies "Azienda elettrica di Massagno" (AEM) and "Stadtwerk Winterthur" (SWW), respectively operating in the Italian- and in the German-speaking part of Switzerland. I was part of the SUPSI research team that was responsible for the quasi experimental assessment of the effects of the behaviour change intervention, as well for activities on the field performed in the Swiss-Italian region.

The project aimed at developing and field testing a persuasive, gamified app targeting the reduction in electricity consumption and the related CO_2 emissions in households. The peculiarity of the app was that, besides providing energy consumption feedback and other persuasive features at the individual level, it also explicitly tackled the social dimension of energy consumption practices in households, by leveraging social norms through collaborative and competitive motivational features, respectively offered by two different app versions. It is therefore an interesting case study to explore, especially combined with the enCompass one, which, to the opposite, strictly relied on individual, as customised as possible, persuasive features.

Specifically, two partially different versions of the Social Power app were created and field tested, which differed on the "gamified structure" they relied upon. One version leveraged *collaborative-based* types of social interactions between its users, thus relying on a collaborative gamified structure, while the other version leveraged *competitive-based* types of interactions between its users, thus relying on a competitive gamified structure.

The field intervention aimed at assessing the Social Power app's effectiveness took place in Spring 2016 for thirteen weeks, by involving a sample of voluntary self-selected households living in the municipalities of Massagno (about 8'500 inhabitants, entirely embedded in the urban agglomeration of Lugano, which, with about 65'000 inhabitants, is the largest municipality of Canton Ticino, the Italian speaking-part of Switzerland)

[1]Funded by Gebert Rüf Foundation under the BREF program in the field of Social Innovation, Grant agreement No GRS-065/14, `https://www.grstiftung.ch/de/media/portfolio~grs-065-14~.html` (last accessed on January, 27 2023).

and Winterthur (about 105'000 inhabitants, located in Canton Zurich, in the German-speaking part of Switzerland). The treated households, that were randomly allocated to use either the collaborative or the competitive gamified structure, were recruited with the support of the local AEM and SW utility companies. Such utilities also guaranteed the automatic collection of electricity smart meter consumption data and their feeding into the Social Power app, supported the identification of comparable control groups of households (control group) and provided baseline electricity consumption data for both the treatment and the control groups households.

Besides analyses on such data to quasi-experimentally estimate the average treatment effect on the treated (ATT), the Social Power research team also performed a three-wave questionnaire (before, after the intervention, and one year further after) targeting the treated household, as well as a few in-person, one-hour interviews with some of the treated households, to collect additional elements for a broader evaluation of the effectiveness of the approach. The outcomes of analyses by the Social Power research team are reported in Wemyss, Castri, et al. (2018) and Wemyss, Cellina, Lobsiger-Kägi, et al. (2019), as well as in a few conference papers and presentations available on the project website (`http://www.socialpower.ch/index.php/publications/`, last accessed on January, 27 2023). Earlier analyses by the project team did not look for possible heterogeneity of the effects between the involved households, which is a missed opportunity, since a few characteristics were observed for both the treatment and the control group households, and are thus available for specific analysis. In this chapter I therefore fill this gap and, focusing on electricity consumption data only, explore in more details the short and long term effects of the Social Power app.

I open the chapter by introducing the characteristics of the Social Power app, in its two gamified structures, then present the quasi experimental methodology I used to assess the effects of the intervention's average treatment effect on the treated (ATT), thus dealing with the overarching research questions RQ1- RQ3 I introduced in Section 1.4. I then present the obtained results, focusing on the outcomes of the heterogeneity analysis between the households, and conclude the chapter by summarising the main findings. For their discussion, I refer the reader to Chapter 7, wherein I offer a general discussion also considering the other two case studies.

With respect to the previous analyses performed by the Social Power team, my specific contribution can be summarised as follows:

- the theoretical clarification of the app's features, by fitting them within stage models of behaviour change and principles for persuasive systems design;
- the use of panel regression models to estimate the short- and long-term effects of app use;
- analyses of heterogeneity of the effects of the intervention on varying the observed characteristics of the households, namely the composition of the household, the type of house, and the city where they live.

5.1 The Social Power persuasive app(s)

Social Power adopts a gamification approach and exploits a number of persuasive features supporting electricity saving. Central to Social Power is the idea of leveraging "social support" persuasive features, as they are referred to by the Framework for persuasive systems design (Oinas-Kukkonen and Harjumaa, 2008). In Social Power, in fact, households are not regarded as isolated agents for change; rather, they are seen as socially embedded actors, that share values and goals, and together are part of broader collective dynamics that they all contribute to shape —and are shaped by.

Each household using the app is automatically put into a team, invited to "play" the Social Power game for three months, together with the team members, and provided with an electricity saving goal at the team level. Depending on the app's gamified structure, the team level goal is different. The collaborative gamified structure invites the team to globally achieve a 10% electricity saving goal, while the competitive gamified structure invites the team to save more electricity than a rival team, which is automatically created by the app as well. The collaborative structure thus leverages intra-group collaboration within members of the same team, while the competitive structure leverages inter-group competition between members of different teams, to support the same energy saving behaviour. Note however that, ultimately, both gamified structures build on the cooperation between team members, within their team: also in the competitive gamified structure, in fact, team members have to cooperate with their team in order to beat the rival team. Further, the app design is such, that a household is either associated with the collaborative gamified structure or with the competitive one, and has no possibility to move from the collaborative to the competitive structure, and vice versa. Apart for the social support features directly associated with the collaborative and competitive gamified structures, however, Social Power is characterised by a number of other features, which do not differ between the gamified structures.

Regarding the specific type of energy consumption targeted by Social Power, only electricity smart meters are available in the two regions of AEM and SWW, the two utility companies supporting project activities. Therefore, Social Power focuses on electricity saving only. However, differently than the enCompass case, in which advanced algorithms and computational competences were available among the app development team, Social Power is not able to deal with electricity consumption for the purposes of heating rooms or producing hot water. The research team in fact decided to on purpose only focus on electricity consumption for lighting and electric appliances, excluding the consumption of electricity to satisfy the demand for hot water or space heating. This design choice thus brought about some limitations in the requirements for app use (households with boilers or heat pump are excluded by app use, at least during the intervention activities) and also resulted in the lack of app's features aimed at favouring the decrease in households' consumption for hot heating or hot water production purposes.

Table 5.1 shows an overview of all the Social Power app's features, summarising them from the perspective of both their theoretical background and the persuasive principles and techniques they exploit. As for the enCompass case, I refer to the Transtheoretical model of behaviour change (Prochaska and Velicer, 1997), to the techniques listed in the taxonomy for behaviour change interventions by Abraham and Michie (2008) and to the principles by the framework for Persuasive Systems Design (PSD) by Oinas-Kukkonen and Harjumaa (2009).

5.1.1 Pre-contemplation stage

In this stage, households have no motivation for change and do not intend to take action. Similarly to enCompass, to trigger their motivation, Social Power activates a *consciousness raising* process. For this purpose, it offers daily feedback on the household's electricity consumption, allowing to monitor its hourly evolution throughout the day. The "Energy diary" section (Figure 5.1) in fact reports the automatically collected smart-meter data, by aggregating it at the hourly level, starting from the 15-minute data provided by utility companies[2].

Like in enCompass, consumption is represented through a barplot and the user can freely move along the horizontal axis representing time, in order to explore the evolution of the household's consumption data and start to glimpse its daily and weekly patterns. Differently than enCompass, however, the barplot is more simplified and provides less technical details. For instance, the specific amount of consumed kilowatthours is only reported in the lower part of the screen and it is not shown in the plot, which aims at intuitively representing the amount of consumption, and its evolution over time, through the height of the bar.

Furthermore, every week app users also receive feedback on the impact of their behaviour in terms of saved electricity, compared to their own historical baseline (namely, their average weekly electricity consumption collected over a comparable period —the same considered for the evaluation of the Social Power intervention, that I present in details in Section 5.2.1). Inspired by (Schultz et al., 2007), such feedback is accompanied by an injunctive norm feedback, which is provided through an unhappy, neutral, or smiling face graphic (emoticon), for whether the household used more, the same, or less electricity, respectively, as compared to the personal baseline (Figure 5.1, right-hand side). Through such feedback, which at the start of app use is mostly an unhappy face, since no electricity saving has been performed yet, users are pushed towards the contemplation stage.

5.1.2 Contemplation stage

Once households start contemplating they might change, they need to be triggered to soon start changing, instead of indefinitely postponing action. For this purpose, the Transtheoretical model suggests to activate a *self-reevaluation* process, aimed at

[2]For technical reasons, consumption data is shown in the Energy diary on the day after the consumption has occurred and not in real time.

Table 5.1: Features of the Social Power app targeting households.

Stage	Process	Feature	Technique	PSD framework
Pre-contemplation	*Consciousness raising* Increasing awareness for causes, consequences and cues about a behaviour	Individual electricity consumption feedback Individual electricity saving feedback	2. Provide information on consequences 3. Provide information about others' approval 12. Prompt self-monitoring of behaviour	Self-monitoring Normative influence
Contemplation	*Self-reevaluation* Cognitively and affectively assessing one's self-image, with and without a particularly unhealthy habit	Individual-to-group (collaborative structure) or Group-to-group (competitive structure) electricity saving feedback	13. Provide feedback on performance 19. Provide opportunities for social comparison	Self-monitoring Social learning Social facilitation Social comparison
Preparation	*Self-liberation* Believing that one can change and committing to act on such a belief	Step-by-step weekly challenges	4. Prompt intention formation 7. Set graded tasks 8. Provide instruction	Reduction Personalization
Action and Maintenance	*Counterconditioning* Learning of more sustainable behaviours that can substitute the less sustainable ones	Tips and external blog Step-by-step weekly challenges	7. Set graded tasks 8. Provide instruction 9. Model or demonstrate the behaviour	Suggestion Tunnelling
	Contingency management Providing consequences (rewards) for taking steps in a particular direction	Individual and team-level points Congratulation push notifications Monthly quizzes with tangible rewards	14. Provide contingent rewards	Praise Rewards Recognition
		Social bonus points Individual-to-group (collaborative structure) or Group-to-group (competitive structure) electricity saving feedback	19. Provide opportunities for social comparison	Social comparison Cooperation Competition
	Helping relationship Providing social support (caring, trust, general support) for new behaviour	Facebook page Monthly quizzes	6. Provide general encouragement 17. Prompt practice	Social role

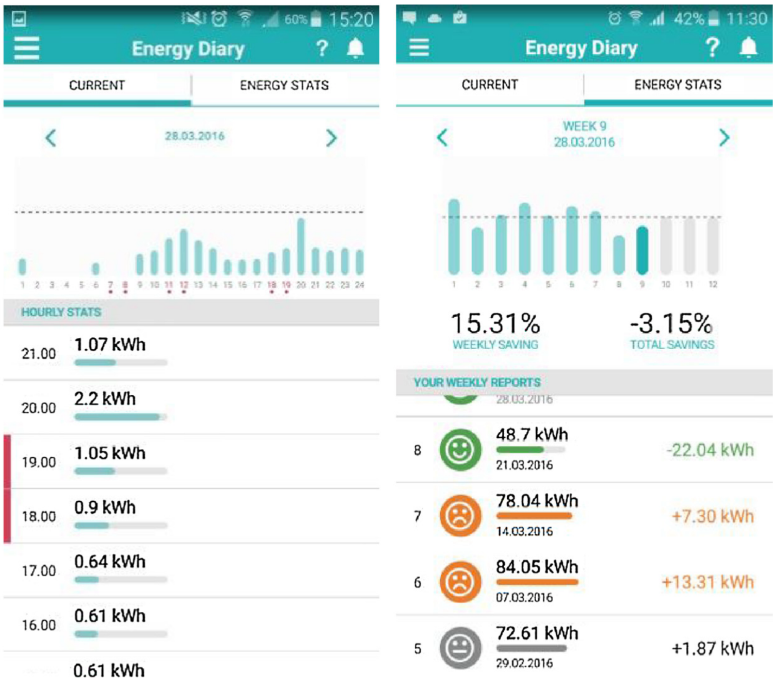

Figure 5.1: Social Power pages providing feedback on electricity consumption at the household level. These pages are offered in both the collaborative and competitive app versions.

cognitively and affectively assessing one's self-image, with and without a particularly unhealthy habit. In Social Power, this is triggered by the social features offering feedback on the electricity saving impact by other households. Such a "social feedback" is different, on varying the gamified structure.

In the collaborative gamified structure (Figure 5.2, left), the individual household's electricity saving is directly compared with the average saving by the team, via an individual-to-group comparison. This is represented by means of a coloured plot showing the average weekly savings from the beginning of the intervention, in absolute kWh and percentage values, of both the single household and the team it belongs to. If the household members perceive the other households of their team are saving energy, they are potentially led to think of themselves as "electricity savers" as well, via a process of identification with the other team's households, and thus are triggered to activate the desired energy-saving behaviour. The provided feedback can favour the household's identification with other households that actually managed to achieve an energy saving impact.

In the competitive gamified structure (Figure 5.2, right), a group-to-group comparison is provided, by means of two plots, which respectively represent the average weekly saving by the team that the household belongs to, and the saving of the rival team. In this case, a process of identification might be activated as well, however it is expected to be weaker than in the collaborative case. The household in fact is only shown average feedback about the whole team's performances and has no direct and specific feedback on its own performances compared with those by the other team members.

Figure 5.2: Social Power pages providing individual-to-group feedback (collaborative gamified structure, left) and group-to-group feedback (competitive gamified structure, right).

5.1.3 Preparation stage

In the *Preparation* stage, households concretely develop plans for actions, to be implemented in the very next future. Thanks to progress through the previous stages, they start to believe they can actually change, and therefore commit to act coherently with such a belief (*self-liberation* process). To support such a process, Social Power invites its users to join weekly challenges, namely to perform electricity-saving activities within their household routine.

Challenges refer for instance to use the dishwasher only if it is full, to laundry at low temperatures, to use the oven only to prepare two dishes at a time, thus halving energy consumption for cooking, or to perform freezer maintenance operations, in order to keep its consumption efficient (Figure 5.3). Every week a set of new challenges dealing with a specific topic is released[3] and households are provided with a hands-on, step-by-step

[3] The full list of the twelve weekly challenges offered by Social Power is as follows: Energy check-up (get acquainted with the overall amount of household consumption according to the electricity bill), Fridge and freezer, Washing machine and tumbler, Dishwasher, Oven, Cooking, Lighting, Home office and electronics, Kitchen appliances, Green power, Cleaning utensils, Bathroom appliances.

process that guides them from setting their commitment towards the challenge to its completion, with practical guidance at each step. Overall, Social Power is built around twelve weekly topics, which are launched every week apart for the first one, which acts as an onboarding period.

This process supports the *reduction* of complexity of the new behaviour, by splitting it into simple tasks that the householders can easily perform. Households are free to join challenges or to ignore them. If they join a challenge, they can choose when and at which speed to complete it, in a *personalised* fashion. Namely, the completion of challenges sometime during the week is self-regulated and households can perform them at the times that best fit their lifestyle and weekly schedule and is compatible with the constraints affecting them.

5.1.4 Action and Maintenance stages

When finally households start to take action, Social Power provides further support, by guiding them in the practical implementation of electricity saving actions through the four-step challenges. To support activation of a *counter-conditioning* process, for each "topic of the week" that the challenges refer to, both versions of Social Power also provide non customised tips, that suggest actions to be performed to complete the challenge (Figure 5.3). Most of the tips also offer additional information on an online weblog created on purpose[4], providing more suggestions on the topic of the week. Doing so, Social Power *tunnels* households through new experiences, with the aim of persuading them to keep implementing the new behaviour, as long as they experience it.

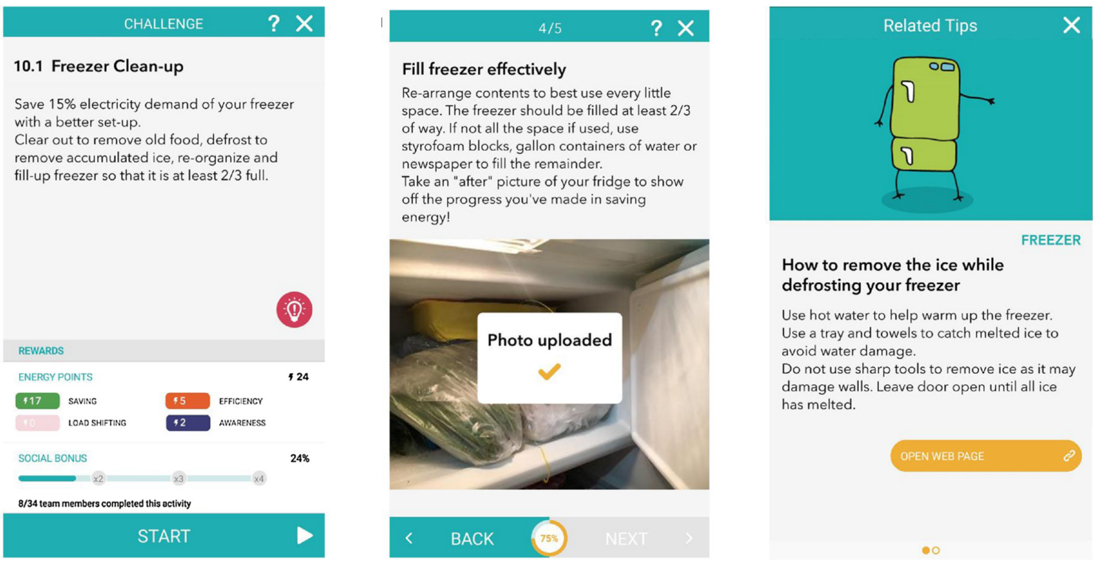

Figure 5.3: Social Power pages reporting a challenge step, the challenge completion photo upload, and the supporting tip. Examples refer to the "Freezer clean-up" challenge.

[4]Unfortunately, the weblog was only active during the intervention period and it is no longer available for documentation.

At challenge completion, households are requested to upload a picture in the app demonstrating their effort and the results they obtained, such as for instance a picture of their oven with a cake and a baking pan of lasagna being cooked at the same time. At picture upload, they receive points, which act as a *reward* for their achievement and as a stimulus for future behaviour repetition. Points are earned at both the individual household level and at the team level, and are attributed exactly in the same way in the collaborative and competitive gamified structures. Additionally, "social bonus" points are attributed at the household and team level on a weekly basis, if on average the whole team manages to save electricity compared to its weekly baseline consumption. If so, all team households also receive push notification messages congratulating for their results. Further, to stimulate the feeling of belonging to a team, motivate households to remain engaged, and activate a feeling of social pressure within the team members, whenever a team attains a high percentage of members completing a specific challenge, extra social bonus points are attributed to the team. These mechanics are also expected to reinforce perception of descriptive social norms, since they remark that other households have already implemented the target behaviours.

Though challenges, tips, and points are the same in both gamified structures, the type of team-level achievement feedback differs between the gamified structures: the collaborative one provides individual-to-group feedback (number of completed challenges, number of points, and percentage of electricity saving obtained by the individual household and by the whole team), while the competitive one provides a group-to-group feedback (number of completed challenges, number of points, and percentage of electricity saving obtained by each team).

No direct communication options are made available by Social Power to its users, not even for chat exchanges within the team. However, during the intervention period an external, publicly accessible Facebook page was created (now discontinued). Allowing potential exchanges between the app users, this external feature is supposed to support exchange of experiences and of electricity saving suggestions, thus enhancing the activation of an *helping relationship* process.

Finally, every four weeks, Social Power offers quizzes with real prizes (about 100 CHF overall per quiz session), focusing on the weekly topics presented in the previous weeks. Despite the low amount of offered prizes, these are expected to act as reminders that revive the interest by the households and favour their repetition of energy saving actions during the *maintenance* stage.

To conclude the presentation of the Social Power app features, Table 5.2 summarises the app's key components and indicates similarities and differences between the collaborative and competitive gamified structures and the related app versions.

Table 5.2: Features of the Social Power app's gamified structures and the related app versions.

	Collaborative	Competitive
Individual electricity consumption feedback	✓	✓
Team level electricity saving feedback	Individual-to-group	Group-to-group
Weekly challenges	✓	✓
Weekly tips and blog page	✓	✓
Points (individual and at team level, including social bonuses)	✓	✓
Team level point feedback	Individual-to-group	Group-to-group
Congratulation notifications	✓	✓
Facebook page	✓	✓
Monthly quizzes	✓	✓

5.2 Research design

I now introduce the methodology I follow in order to tackle my RQ1 - RQ3 research questions about the energy and CO_2 saving effectiveness of persuasive apps targeting households. Applying these questions to the Social Power case requires to estimate the causal effect of use of the Social Power app on electricity consumption and CO_2 emissions, both in the short-term (namely, during app use, RQ1) and in the long-term (namely, a reasonably long period after its use, RQ2), and to verify if the magnitude of such a causal effect differs, on varying the heterogeneity of observed characteristics of app users (sub-group analysis, RQ3). To tackle these questions, I refer to the research design I developed together with the Social Power team during the Social Power project and perform novel analyses developed on purpose for this dissertation.

As for enCompass, in Social Power we adopted a quasi-experimental approach, characterised by the voluntary, self-selection of the treated households and by the posterior identification of comparable control households. The design of the Social Power intervention is however more complex than the enCompass case and offers further opportunities for novel analyses on the very same datasets we collected during the project itself.

First of all, in Social Power we tested the effect of two partially different types of treatment: the Social Power app adopting the collaborative gamified structure and the Social Power app adopting the competitive gamified structure. Indeed, the original goal of the project was exactly to assess the effects of the collaborative and competitive app versions and to verify if and to what extent they produced different effects.

Then, we specifically designed the project with the aim to assess the long-term effect of the app-based treatment, by considering three monitoring periods: a pre-intervention period to collect baseline electricity consumption data, the intervention period itself, and a post-intervention period, which was set exactly one year after the intervention took place, to verify presence and magnitude of the effect in a reasonably long-term.

Further, we performed the same intervention in two cities (Massagno and Winterthur), exactly in the same periods. Besides the different geographical location, these cities are also characterised by different languages (Italian and German) and by overall different cultural contexts (the Mediterranean and the German one), which might also favour different levels of environmental awareness and social support for pro-environmental behaviour. Even though at the time of the project we only marginally explored this line of enquiry, the location of the intervention might be a good predictor of its effects, as it may act as a proxy of a richer set of variables related to the cultural context: performing an analysis of the effects by location sub-groups might therefore provide useful insights for policy-making.

And, finally, in the case of Social Power utility companies hold more information on the characteristics of their customers, compared to the enCompass case. Available data in both Massagno's (AEM) and Winterthur's (SWW) utilities allows in fact to categorise any of their household customers according to the two following variables:

- the type of household: "single adults" or "families" (if at least one son or daughter lives in the household, independently on their age);
- the type of building: "apartment" or "house" (namely, anything other than apartment).

Availability of such pieces of information offers interesting opportunities for analysis, that at the time of the Social Power project we did not consider: it allows to perform novel analyses about possible different effects in terms of heterogeneity on the location, the type of household, and the type of building. Understanding whether an app-based intervention as Social Power is *more* effective, say, for families, for households living in apartments, or for households surrounded by a German cultural context, can provide policy-makers with relevant practical information for future app-based interventions to effectively support the energy and climate transition. Furthermore, understanding for which target categories the intervention is *less* effective can suggest venues for future research aimed at understanding the reasons for the lower effectiveness and possibly at identifying corrective and improvement measures.

Overall, to deal with my research questions I refer to the research design we had already adopted in the Social Power project, by integrating it with heterogeneity analyses that were not originally envisioned. I exploit the same data we collected during the project for the three monitoring periods, both for the treated and the untreated households. To address RQ1 and RQ2, however, instead of the Difference-in-Differences approach that we adopted during the project, I perform novel analyses based on a panel data regression approach. Also, I address RQ3 with novel regression analyses aimed at estimating differences —if any— in the ATT on household sub-groups based on the location, the household type, and the building type.

In the next sections I first present the timeline of the Social Power intervention and the related data collection periods. Then, I report how, during the Social Power project, we identified the treatment and control groups, and summarise the observed characteristics of the analytical sample I consider for the novel analyses I address here. Finally, I present the panel regression models I developed to address my research questions.

5.2.1 Timeline of the intervention

For Social Power we adopted a three-step repeated measurement design: electricity consumption collected in treatment and control groups before the intervention (baseline) was compared with consumption collected during the intervention, in order to asses its short-term ATT, and with long-term consumption collected one year after, in order to assess its long-term ATT. Each monitoring period lasted for about three months, as shown in Table 5.3. For the baseline, we considered the 2015 October-December months (eleven weeks), excluding the last weeks of December, in which electricity consumptions are highly influenced by Christmas feast days. The Social Power intervention was run for thirteen weeks (one "onboarding" week with no challenges and tips and twelve weeks, each one devoted to a specific "topic of the week") in the 2016 February-May months. The long-term assessment of the effects was performed through a follow-up monitoring period, consisting of thirteen weeks between February and May 2017.

Table 5.3: The time periods I considered to assess effectiveness of the Social Power intervention.

Period	Period name	Dates	Type of period
Baseline	Period Zero	October, 1 2015 - December, 15 2015	Pre-treatment
Intervention	Period One	February, 1 2016 - May, 1 2016	Treatment
Follow-up	Period Two	February, 1 2017 - May, 1 2017	Post-treatment

For baseline and intervention periods we chose two relatively comparable periods in terms of daylight hours, which were expected to be similar in terms of households' electricity demand for appliances and lighting. The seasonal difference in outdoor temperatures between these periods was not regarded as critical, since it mainly affects heating demand, which was on purpose excluded from the intervention. By design, in fact, only households whose electricity consumption was due to electric appliances and lighting only were eligible to join the intervention. The effect of outdoor temperatures, which might influence the time spent indoor and thus affect electricity consumption, was instead considered negligible, and therefore not considered in the analysis —in any case, presence of control groups allows to account for seasonal temperature evolution. In addition, by opting for two consecutive periods, we expected limited changes in households' composition (no students would leave home to study, for example), thus reducing chances of large spurious influences on electricity consumption other than the use of the Social Power app.

As the length of the monitoring periods was slightly different in terms of number of weeks, we measured the outcome of the intervention in terms of "average weekly electricity consumption" over each monitoring period. For the baseline period, utility companies only provided us with total consumption over the whole period, for each household. From this data, we computed weekly averages by dividing it by the number of weeks (eleven).

For the intervention and follow-up periods, instead, for each household utility companies provided us with the weekly consumption data. Therefore, we first aggregated weekly consumptions over the respective periods, and then divided them by the number of weeks (thirteen) in order to obtain the weekly averages we were interested into. Overall, for each household three average weekly electricity consumption data are available to estimate the treatment effect.

5.2.2 Identification of treatment groups

The goal of the Social Power project was to recruit at least 100 households to receive treatment in each city. In that case, no sensor installation was needed for the households to join app use and to be included in the intervention. The only eligibility criteria to be respected were as follows:

- households had to be already covered by a smart meter device automatically monitoring their electricity consumption; such a meter should have been active since October, 2015, when the baseline monitoring period started;
- the meter had not to measure electricity consumption for heating or hot water purposes. Namely, if heating demand for the household was satisfied by electric boilers, direct electric heating systems, or heat pumps, they had not to be associated with the smart meter also measuring their electricity consumption for lighting and other appliances. Operatively, this implied that households equipped with these heating systems were not eligible to use the Social Power app and join project activities.

In the case of AEM (Massagno), these requirements were not too critical to potentially recruit a sufficiently large number of interested households via a public call and then randomly assign the Social Power treatment(s) to only half of them. In the case of SWW (Winterthur), instead, the total number of households respecting both the eligibility criteria was very low, equal to about 350. With the colleagues of the research team we therefore *a priori* excluded the idea of recruiting a large number of interested households and then randomly assigning them to the treatment or control group. Despite the support by the utility companies to recruit interested households, we considered this goal to be too ambitious, as it would have required to recruit 200 households out of the 350 available ones. Therefore, we opted for a quasi-experimental design: treatment groups were identified via open recruitment strategies customised to each location and utility company. Namely, they were self-selected volunteers. Comparable control groups were instead identified on a posterior basis, with the aim of matching them to the characteristics of the treated households. To guarantee comparability of the outcomes, we followed the same quasi-experimental approach both for Winterthur and for Massagno.

In both locations, in Fall 2015 an open call was performed, via a press release and related posts on social networks, by both the utility companies and the universities involved in the project. In Massagno we also organised school classes, to recruit households starting from the pupils, as well as flyers that were distributed across the city. In Winterthur, instead, printed and personalised letters were sent to each of the 350 households meeting the above requirements. In both cases no disguise was used: recruitment materials transparently indicated the energy and carbon saving goals of the Social Power intervention. Monetary incentives were also offered: in each location, all households who remained active until the end of the intervention were included in a final draw, offering three vouchers of approximately 700 euro value sponsored by the local utility company.

Overall, n= 108 households joined the call —thus, well below our initial goal of recruiting 100 households per city— exactly half in each city (n=54). By considering their available characteristics (type of household and type of building), we stratified the group and then randomly assigned households in each stratum to either the collaborative or the competitive group, in equal proportions. At the end of the procedure, in each city n= 27 households were assigned to each group, as summarised by Table 5.4. The n_{collab} = 27 + 27 = 54 households of the collaborative group were assigned to treatment with the app version adopting the collaborative gamified structure, while the other n_{compet} = 27 + 27 = 54 households of the competitive group were assigned to treatment with the app version adopting the competitive gamified structure.

Table 5.4: Organisation and size of treatment groups recruited for the Social Power intervention.

Group	App version	Goal	City	n
Collaborative	Collaborative gamified structure	With your team members, reach a 10% electricity saving goal	Massagno	27
			Winterthur	27
Competitive	Competitive gamified structure	With your team members, save more electricity than the rival team	Massagno	27
			Winterthur	27

The groups thus obtained in each city were not further divided in teams: all the 27 households of the competitive group in Massagno were put in a single team and invited to compete against the team of the 27 households of the competitive group in Winterthur. Similarly, all the 27 households of the collaborative group in Massagno were put in the same collaborative team, and the same happened for the 27 households of the collaborative group in Winterthur. Doing so, we exploited the idea of "natural teams" related to the geographical location ("Massagno against Winterthur"). Mimicking sport championships and leveraging the feeling of belonging to their region, we expected such a team composition strategy could increase households' engagement and maintain it over time. Its downside was a quite large team size (27 households per team). Anyway, we regarded this was acceptable, also because no direct communication channels between team members were made available through the app.

Analyses of the imbalance of the characteristics of the obtained treatment groups (the competitive and the collaborative ones) were performed, by considering the available information about the type of household and the type of house. I do not report them here, however, since the groups I can consider for impact analysis are largely different from the originally recruited ones, due to non-compliance. Between the period of household recruitment and the start of the field intervention (February, 1 2016), in fact, a large number of non-compliance cases occurred in both cities. A few households explicitly asked us to be removed from field activities; others, simply never logon into the Social Power app throughout the intervention period, and therefore did not receive the treatment they were assigned to.

I therefore removed both these types of non-compliant households from the analyses of the treatment effect. Ultimately, a total of n= 46 households constitute the analytical sample available for the analysis —half attributed to the collaborative treatment and half to the competitive treatment. Luckily, proportions in group sizes spontaneously remained constant. The characteristics of the collaborative and competitive treatment groups of the analytical sample are reported in Table 5.5, which also offers a comparison with the households identified for the control group. The non-compliance rate is definitely very high (out of the 54 self-selected households initially recruited in each city, 31 of them in each city did not receive the treatment, which corresponds to a 57.4% non compliance rate). Unfortunately, however, enforcing compliance was impossible for our research team as well as for the utility companies involved in project activities, and we could nothing but observe and report its value.

Table 5.5: The composition of Social Power treatment (collaborative and competitive) and control groups, according to their observed characteristics.

			Type of household				Type of building				Total
			Single adult		Family		Apartment		House		
			Num	%	Num	%	Num	%	Num	%	Num
Treatment	Collaborative	Massagno	4	40	6	60	8	80	2	20	10
		Winterthur	4	31	9	69	8	62	5	38	13
	Competitive	Massagno	5	38	8	62	10	77	3	23	13
		Winterthur	3	30	7	70	7	70	3	30	10
Control		Massagno	9	61	14	39	18	78	5	22	23
		Winterthur	10	43	13	57	15	65	8	35	23

5.2.3 Identification of control groups

In each city, comparable control groups of households were identified *ex post*, at the end of the field intervention, in order to match characteristics of the treated households that complied with the intervention. They were selected among households that respected the two requirements regarding availability of a smart meter connection and lack of use of electricity-based systems for room heating or hot water production. They were not aware that their consumption data was used in the quasi-experiment, in order to avoid any Hawthorne-like effect in the following data collection period.

The sample size was set, for each city, to correspond with the size of both treated groups, namely the collaborative and competitive one, by considering treatment-compliant households: in each city, 23 control households were identified, for a total of 46 control households. Matching to the observable characteristics of the treatment groups (type of household and type of building) was performed through a stratified sampling approach, starting from the samples of household customers of each utility company that met the above requirements. Within each stratum, a number of households was randomly picked in order to create a similar proportion of household (adults only vs. families) and building types (apartments vs. houses) as the overall intervention groups, and no further criteria were used. This operation was performed by the utility companies AEM and SWW, by applying a protocol we provided them, since we were not allowed to access their customer data-base, not even a pseudonymised extract of the customers meeting the selection criteria.

In each city, the overall composition of the control groups was therefore comparable to the composition of the competitive and collaborative groups. Table 5.5 reports the group composition based on the observed characteristics. No additional information associated with control group households is available, apart for the two observed covariates and their electricity consumption. Utility companies in fact provided us with pseudonymised data only, which does not allow for interaction possibilities with control group members.

5.2.4 Check of imbalance

To verify that the two treated groups resulting from such operations, and subjected to non-compliance, were balanced compared with each other and with the control group, we performed an imbalance check based on the observed information on the household and building type, as well as on the electricity consumption during the baseline period.

Figure 5.4 shows the composition of the three groups based on the two observed variables "house type" and "household type". From a visual inspection, the groups appear to be well-balanced: the random allocation between the two types of treatment and the later matching of the control group via a stratified random sampling procedure seem to have worked well.

As for the enCompass case, I follow Gerber and Green (2012) and verify if, for the observed characteristics, imbalances among the three groups are larger than one would expect from chance alone. For this purpose, I regress the assigned treatment on the available covariates household and building type, and then compute the aggregate F-statistics. Specifically, I regress a variable indicating the treatment type (collaborative, competitive, or control) on two dichotomous variables respectively taking on value "1" if the household is a family and lives in a house, and "0" otherwise (equation 5.1). In order to avoid collinearity, I do not consider the corresponding dichotomous variables representing single adult households and apartment building types.

$$treat\_type_i = \beta_0 + \beta_1 family_i + \beta_2 house_i + u_i \tag{5.1}$$

This allows me to test the null hypothesis that the covariates predict the treatment membership no better than would be expected by chance. The F-statistic's p-value

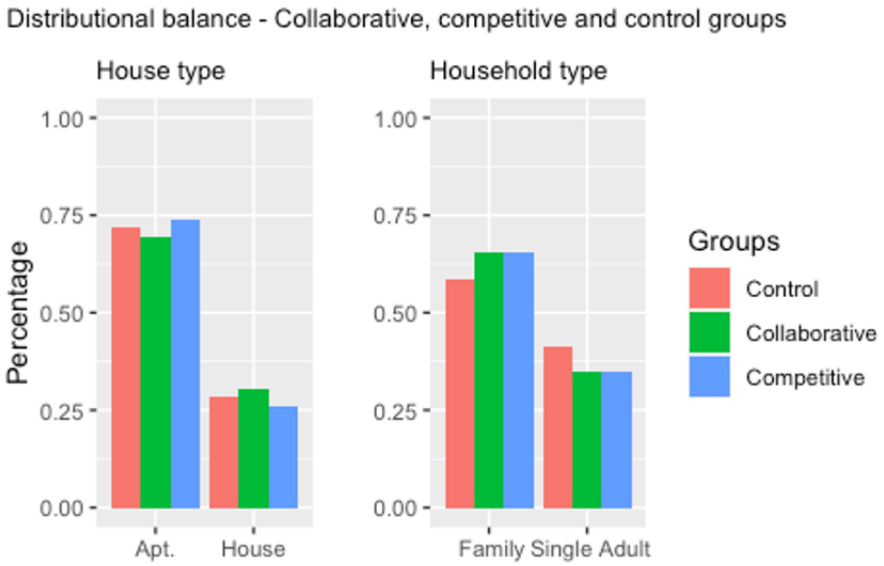

Figure 5.4: Visual comparison of the observed characteristics of the Social Power treatment (collaborative and competitive) and control groups.

resulting from regression (5.1) is equal to 0.8275. This indicates I can reject the null hypothesis that the regression's parameter estimates of standardized coefficients are jointly equal to 0, even at the 0.1 significance level, and confirms that the three groups are balanced with respect to the two observed household characteristics.

Finally, as suggested by Sergici and Faruqui (2011), I also look for differences between the pre-treatment mean baseline electricity consumption values of the three groups, which are reported in Table 5.6. In this case, mean yearly baseline values of the three groups are not statistically different, even at the 0.1 significance level, as indicated by a two-tailed Kruskal-Wallis test (p-value= 0.7376), chosen instead of an ANOVA since Shapiro-Wilk tests indicates distributions of the electricity consumption variable in the three groups is not normal. The boxplot reported in Figure 5.5 also provides a visual indication of the relatively small differences between baseline electricity consumption values in the three groups. This additional check therefore confirms that the two treatment (collaborative and competitive) groups are comparable between each other and with the control group also regarding the amount of electricity they consumed during the baseline period. The three groups can therefore be used to estimate the effects of the Social Power intervention.

Table 5.6: baseline electricity consumption of collaborative, competitive, and control groups.

Baseline electricity consumption [kWh/week]	Collaborative group (n_{collab}=23)	Competitive group (n_{compet}=23)	Control group ($n_{control}$=46)
Mean	67.11	63.02	75.74
Standard Deviation	48.09	52.41	84.36

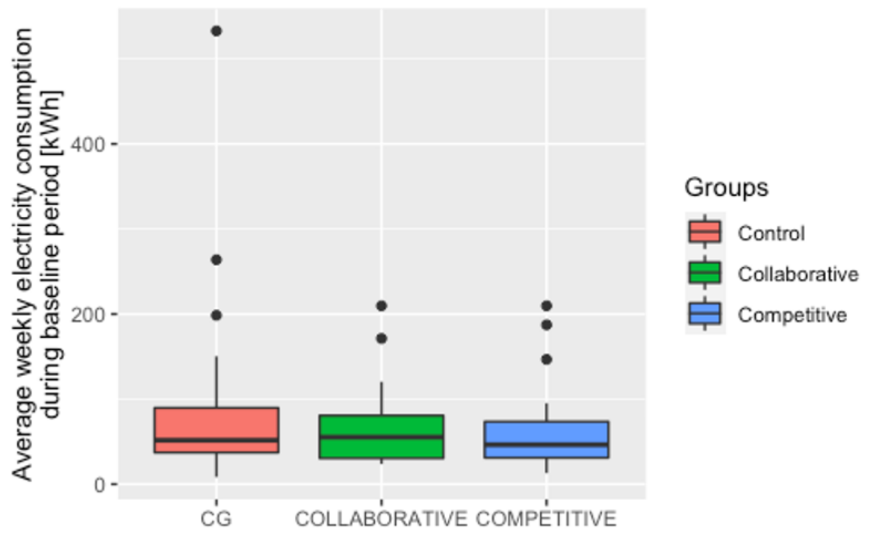

Figure 5.5: Yearly baseline electricity consumption values of the three Social Power groups. Label "CG" indicates the control group.

5.2.5 Attrition

A final consideration is needed, before estimating the intervention effect. The composition of the groups reported in the previous section refer to the groups that complied with the Social Power intervention and to the control group identified to match them. From the end of the intervention in May, 2016 until the end of the follow-up period in May 2017, however, five households moved to another location. Their consumption data is therefore no longer available for the estimate of the effect in the long-term.

Tables 5.7 and 5.8 respectively indicate in which groups attrition occurred and show the composition of the available groups for assessment of the long-term impact. Due to the very limited size of the analytical samples after accounting for non-compliance, I opted for keeping these five households in the analytical sample used to assess the short-term effect. A check of imbalances performed via the model of Eq. (5.1) showed that the three groups resulting out of attrition in follow-up period are still not significantly different in their observed characteristics (p-value of the F-statistics equal to 0.9428), which reassured me about the choice to consider two slightly different analytical samples for the two periods.

Overall, for analyses of short-term effectiveness I consider the sample reported in Table 5.5. For analyses of the long-term effectiveness, instead, I consider the sample reported in Table 5.8. Such a choice is not critical for my results: simply, the panel regression model will be slightly unbalanced.

Overall, Figure 5.6 shows the evolution of average weekly electricity consumption for the three groups (treated with the collaborative Social Power app version, treated with the competitive Social Power app version, and control), over the three monitoring periods.

Table 5.7: Number of households that, between the end of the intervention and the end of the follow-up period, moved to another location.

Missing households due to attrition between intervention and follow-up periods	Massagno	Winterthur
Collaborative group	0	0
Competitive group	1	0
Control group	2	2

Table 5.8: The composition of Social Power treatment (collaborative and competitive) and control groups at follow-up. This is the analytical sample used for long-term analysis.

			Type of household				Type of building				Total
			Single adult		Family		Apartment		House		
			Num	%	Num	%	Num	%	Num	%	Num
Treatment	Collaborative	Massagno	4	40	6	60	8	80	2	20	10
		Winterthur	4	31	9	69	8	62	5	38	13
	Competitive	Massagno	5	42	7	58	9	75	3	25	12
		Winterthur	3	30	7	70	7	70	3	30	10
Control		Massagno	8	62	13	38	16	76	5	24	21
		Winterthur	10	48	11	52	15	71	6	29	21

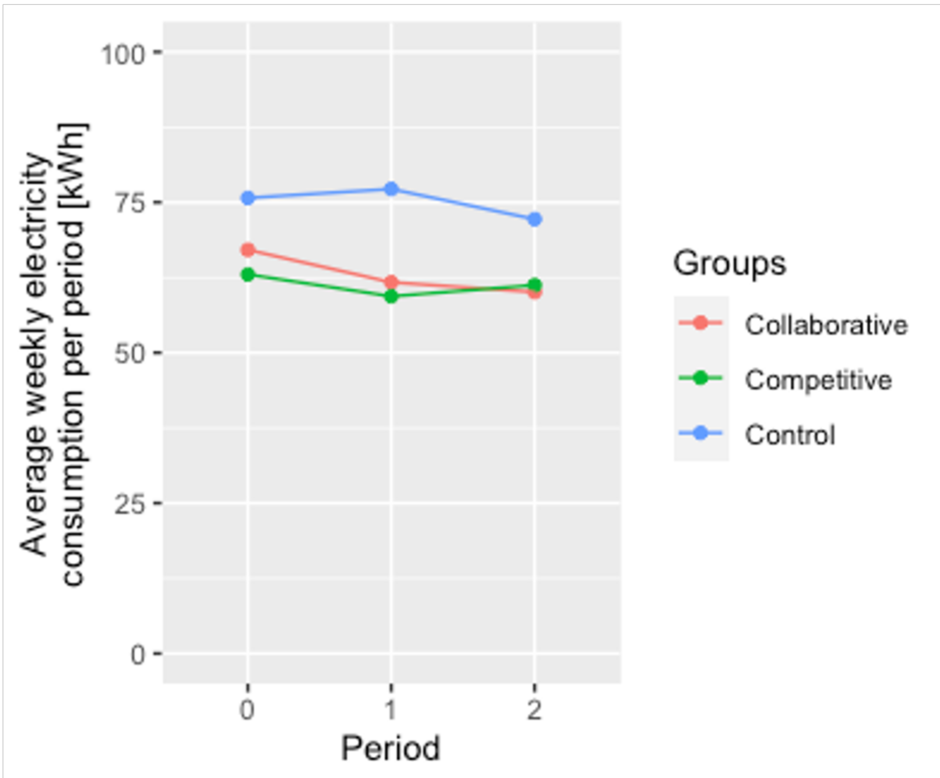

Figure 5.6: Evolution of the mean weekly electricity consumption in the collaborative, competitive and control groups, over the three monitoring periods.

5.3 Panel regression model

To estimate the average treatment effects (ATT) on the treated households, I use a Two Way Fixed Effects (TWFE) panel regression model, whose general structure is presented by equation (4.2). I opt for this approach for the same reasons as the enCompass case: households receiving the Social Power treatment were self-selected, therefore it is likely that their choice to join project activities is correlated with unobserved variables, such as for instance above average values of environmental attitudes or education. Namely, it is likely that the error term in the panel regression model is correlated with the model's independent variable indicating whether a household received the treatment or not.

The specific RQ1 and RQ2 research questions and null hypotheses (H_0) I formulate for the Social Power case are reported in Table 5.9.

Table 5.9: Research questions RQ1 and RQ2 and related null hypotheses.

ID	Research question	Null hypothesis H_0
RQ1_ Collab	What is the average treatment effect on the treated by the collaborative version of the Social Power app, during its period of use (Intervention period, Period One), in terms of savings of electricity and CO_2 emissions?	$H_0\_Collab\_1$: the ATT in Period One is equal to zero
RQ1_Compet	What is the average treatment effect on the treated by the competitive version of the Social Power app, during its period of use (Intervention period, Period One), in terms of savings of electricity and CO_2 emissions?	$H_0\_Compet\_1$: the ATT in Period One is equal to zero
RQ1_Diff	What is the difference between such ATTs for the collaborative and competitive treatments?	$H_0\_Diff\_1$: the difference in ATTs in Period One is equal to zero
RQ2_Collab	What is the average treatment effect on the treated by the collaborative version of the Social Power app, one year after its use (Follow-up period, Period Two), in terms of savings of electricity and CO_2 emissions?	$H_0\_Collab\_2$: the ATT in Period Two is equal to zero
RQ2_Compet	What is the average treatment effect on the treated by the competitive version of the Social Power app, one year after its use (Follow-up period, Period Two), in terms of savings of electricity and CO_2 emissions?	$H_0\_Compet\_2$: the ATT in Period Two is equal to zero
RQ2_Diff	What is the difference between such ATTs for the collaborative and competitive treatments?	$H_0\_Diff\_2$: the difference in ATTs in Period Two is equal to zero

For hypothesis testing I use two-tailed tests and to estimate model parameters I again rely on package "plm" offered by the R statistical software. The specific TWFE panel regression model I compute is the following one:

$$
\begin{aligned}
kWh\_week_{it} = {} & \alpha_1 Collab_i + \alpha_2 Compet_i + \\
& \beta_1 Period\_One_t + \beta_2 Period\_Two_t + \\
& \gamma_1 CollabxPeriod\_One_{it} + \gamma_2 CompetxPeriod\_One_{it} + \\
& \gamma_3 CollabxPeriod\_Two_{it} + \gamma_4 CompetxPeriod\_Two_{it} + \\
& \boldsymbol{c_i} + u_{it} \\
& for\ t = 1, ...3\ and\ i = 1, ...92
\end{aligned} \tag{5.2}
$$

where:

- kWh_week_{it} is the observed dependent variable, namely the average weekly electricity consumption collected over each monitoring period for each household;
- $Collab_i$ and $Compet_i$ are the observed time-invariant independent variables characterising each household. They are dichotomous variables respectively indicating if household i received the Social Power collaborative treatment ($Collab_i = 1$, $Compet_i = 0$), the competitive treatment ($Collab_i = 0$, $Compet_i = 1$), or was part of the control group ($Collab_i = 0$, $Compet_i = 0$). Note, that variable $Control_i$ is not included in the model to avoid collinearity; indeed, it acts as a reference variable;
- $Period\_One_t$ and $Period\_Two_t$ are three dichotomous variables respectively indicating the period (One and Two). Note, that variable $Period\_Zero_t$ is not included in the model to avoid collinearity; indeed, it acts as a reference variable;
- $CollabxPeriod\_One_{it}$, $CollabxPeriod\_Two_{it}$, $CompetxPeriod\_One_{it}$, and $CompetxPeriod\_Two_{it}$ are dichotomous interaction terms indicating if household i is collaboratively or competitively treated and if its electricity consumption respectively refers to each of the treatment periods t. For instance, $CollabxPeriod\_One_{it}$ takes on value "1" if the household is collaboratively treated and its electricity consumption is related to Period One; otherwise, it takes on value "0";
- $\boldsymbol{c_i}$ is the vector of unobserved household-specific effects, which are time-invariant (household fixed effects);
- u_{it} is the unobserved idiosyncratic error term;
- t is the subscript for the monitoring periods, varying from 1 to T, which here is equal to 3;
- i is the subscript for the household, varying from 1 to N, which here is equal to the sum of the households of the treatment and control groups ($n_{collab} = 23$, $n_{compet} = 23$, and $n_{control} = 46$, thus N= 92).

Overall, through model (5.2) I can directly estimate the average treatment effect on the treated (ATT) by the two types of Social Power treatments, for each of the periods I am interested into: the design of the model equation is such that consumption of the control group in year 0 is taken as a reference and that the ATTs by the collaborative and competitive app versions respectively over Period One and Period Two directly correspond to the model estimates of the γ_i coefficients. Since in this case I opted for the non-logarithmic model version, the ATT is expressed in terms of change in weekly electricity consumption [kWh/week]. The corresponding electricity saving percentage (which also corresponds to the CO_2 saving percentage, since I assume that CO_2 emissions are directly proportional to electricity consumption), can be obtained by comparison with the baseline average weekly consumption of each treated group.

Finally, note that, differently than the enCompass case, here I do not include time-specific fixed effects in the model, such as the "Heating degree days" or "Cooling degree days" that were instead relevant for the enCompass case. Social Power in fact does not account for electricity consumption for heating purposes —and in any case neither Winter heating nor Summer cooling months are included in the three monitoring periods.

5.4 Estimate of the treatment effect

The outcome of panel regression model 5.2, that I refer as to "Model I" in order to facilitate comparison with other models I present in the next sections, is reported in Table 5.10. In the table, heteroskedasticity-robust clustered standard errors computed according to the Arellano method (Millo, 2017) are reported in brackets. Model results are also reported in equation (5.3), which shows in bold format the parameter estimates that are statistically significant at the 0.05 level.

$$\begin{aligned} \widehat{kWh\_week} = &\mathbf{-6.88}\, CollabxPeriod\_One - \mathbf{5.12} CompetxPeriod\_One + \\ &- 1.60 CollabxPeriod\_Two + 4.36 CompetxPeriod\_Two \end{aligned} \tag{5.3}$$

Table 5.11 summarizes the model outcomes, also reporting 95% confidence intervals of the ATTs and their effect size, estimated via the Cohen's *d* statistics. The parameter estimates of coefficients γ_1 and γ_2 are significant at the 0.05 significance level. This means I can reject at the 0.05 significance level the $H_0\_Collab_1$ and $H_0\_Compet_1$ null hypotheses related with RQ1. I can thus conclude that in the short-term the ATTs by both collaborative and competitive Social Power app versions are actually different from zero. In percentage terms, compared with baseline average weekly consumptions of the respective groups, these ATTs correspond to 10.25% (collaborative group) and 8.11% (competitive group) savings in electricity consumption and in CO_2 emissions. To verify whether the observed differences in the collaborative and competitive ATTs are also statistically significant, I perform a *post-hoc* interaction test, with the same manual approach suggested by Christensen et al. (2021) that I already used for enCompass. Table 5.12 reports such results: differences between the collaborative and the competitive ATTs are not statistically significant, even at the 0.1 significance level.

The parameter estimates of coefficients γ_3 and γ_4 are instead not statistically significant, not even at the 0.1 significance level: I cannot reject the $H_0\_Collab_2$ and $H_0\_Compet_2$ null hypotheses related with RQ2. Therefore, I cannot conclude that in the long-term the ATTs by both collaborative and competitive Social Power app versions are actually different from zero.

Table 5.10: Output of Model (I) panel regression.

	Model (I)	
	Average weekly consumption [kWh/week]	p value
$CollabxPeriod\_One$	-6.8816** [2.8556]	0.01701
$CollabxPeriod\_Two$	-1.6026 [5.0061]	0.74926
$CompetxPeriod\_One$	-5.1192** [2.4946]	0.04167
$CompetxPeriod\_Two$	+4.3642 [6.1071]	0.47582
Observations	n=92, unbalanced panel. T=2/3, N=271.	—
Total Sum of Squares	30964	—
Residual Sum of Squares	30067	—
Adjusted R-Squared	-0.5155	—
R-Squared	0.02896	—
F-statistic	1.28978 on 4 and 173 degrees of freedom.	0.27591

R "plm" package; model= "within", effect= "twoways".
Heteroskedasticity-robust, clustered standard errors (Arellano method) in parentheses.
Significance levels: * 0.1, ** 0.05, *** 0.01

Table 5.11: Summary of Model (I) outcomes addressing RQ1 and RQ2 for Social Power.

Model (I)	Treatment	Period One	Period Two
ATT [kWh/week]	Collaborative	**- 6.88**** [2.8556]	- 1.60 [5.0061]
	Competitive	**- 5.12**** [2.4946]	+ 4.36 [6.1071]
ATT's 95% Confidence Interval	Collaborative	[-12.55741, -1.205787]	[-11.55276, 8.347564]
	Competitive	[-10.07749, -0.1609131]	[-7.774321, 16.50272]
Effect size (Cohen's d)	Collaborative	- 0.52 (medium)	0.07 (negligible)
	Competitive	- 0.41 (small)	0.21 (small)

Heteroskedasticity-robust, clustered standard errors (Arellano method) in parentheses.
Significance levels: * 0.1, ** 0.05, *** 0.01

Table 5.12: Results of *post-hoc* interaction test on Model (I) outcomes: p-value for the comparison of effects between the collaboratively and competitively treated households.

Model (I) p-values of interaction test	Period One	Period Two
Collaboratively treated households vs Competitively treated households	0.642077	0.4498844

Significance levels: * 0.1, ** 0.05, *** 0.01

To conclude, analysis of the electricity consumption data collected during the Social Power project shows that in the short-term, namely during the intervention itself, both versions of the Social Power app were effective in reducing electricity consumption and CO_2 emissions. Savings produced by the collaborative app version were equal to 6.88 kWh/week, corresponding to 10.25% with respect to the baseline of the collaboratively treated group. Savings produced by the competitive app version were equal to 5.12 kWh/week, corresponding to 8.11% with respect to the baseline of the competitively treated group. The effect sizes of ATTs are quite similar: respectively, -0.52 and -0.41, which can be regarded as an intermediate and relatively small size. Further, an interaction test showed that the two ATTs are not statistically different from each other. This means therefore that the two versions of the Social Power app had the same effect.

However, such an effect is only found for the short-term: in the long-term, one year after the Social Power intervention, the observed ATTs are no longer statistically significant, even at the 0.1 significance level. Similarly to the enCompass case, the beneficial effect produced by the Social Power app did not persist over time.

5.5 Heterogeneity analyses

Since there are no statistically significant differences between the effects of use of the two gamified structures of the Social Power app, I perform the RQ3 heterogeneity analyses by considering the Social Power treated groups as a whole ($n_{treated}$= 46). Doing so, I can increase statistical power and, despite the low sample size, increase chances that the outcomes of sub-group analyses are statistically significant.

Characteristics of the entire group of households treated with one of the two versions of the Social Power app are reported in Table 5.13. An imbalance test performed through model (5.4)

$$treat_i = \beta_0 + \beta_1 family_i + \beta_2 house_i + u_i \tag{5.4}$$

confirms that, with respect to the observed characteristics “type of household” and “type of building”, imbalances between the entire treatment group and the control group are not larger than one would expect from chance alone (p-value of the F-statistics equal to 0.8143).

Table 5.13: The composition of Social Power treatment and control groups, according to their observed characteristics.

			Type of household				Type of building				Total
			Single adult		Family		Apartment		House		
			Num	%	Num	%	Num	%	Num	%	Num
Treatment	16	35	30	65	33	72	13	28	46		
Control	19	41	27	59	33	72	13	28	46		

Average weekly electricity consumptions measured in each group during the baseline period are reported in Table 5.14 and represented in Figure 5.7. A Mann Whitney U

Test, chosen because in both groups the variable distribution is not normal, confirms that differences in baselines are not statistically significant (p-value= 0.6662). The two groups are therefore comparable in both their observed characteristics and their average weekly electricity consumption during baseline period.

Table 5.14: Comparison of baseline electricity consumption between treated and control groups.

Baseline electricity consumption [kWh/week]	Treatment group (n=46)	Control group (n=46)
Mean	65.06	75.74
Standard Deviation	49.78	84.36

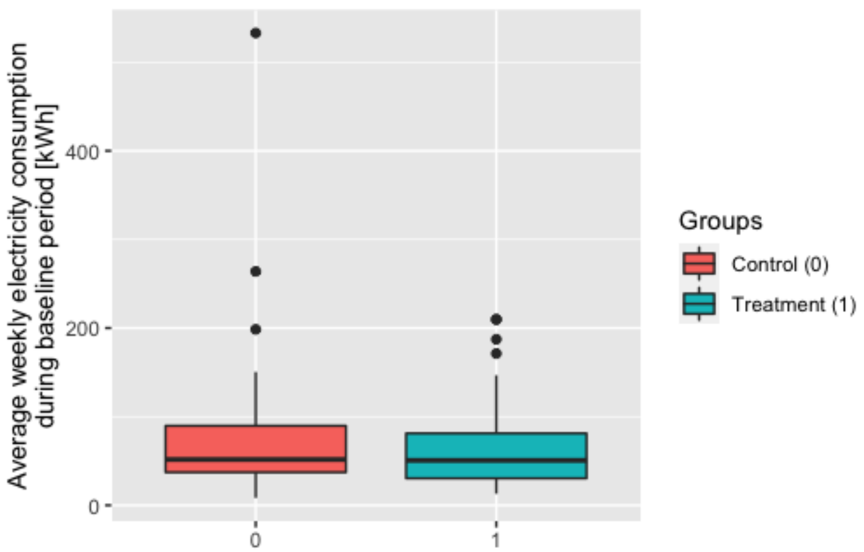

Figure 5.7: Average weekly baseline electricity consumption of the two Social Power groups.

The TWFE regression Model (II) that I thus use as a basis for heterogeneity analysis is represented in equation (5.5):

$$\begin{aligned} kWh\_week_{it} = {} & \alpha_1 Treat_i + \\ & \beta_1 Period\_One_t + \beta_2 Period\_Two_t + \\ & \gamma_1 TreatxPeriod\_One_{it} + \gamma_2 TreatxPeriod\_Two_{it} + \\ & \boldsymbol{c_i} + u_{it} \\ & for\ t = 1, ...3\ and\ i = 1, ...92 \end{aligned} \tag{5.5}$$

where $Treat_i$ is a dichotomous variable taking on value "1" if the household was treated with the Social Power app, "0" otherwise. Outcomes of Model (II), reported in Table 5.15, indicate that the average treatment effect on the treated households is a short-term 6.00 kWh/week decrease in electricity consumption, corresponding to a 9.22% decrease in both electricity consumption and CO_2 emissions, with respect to baseline consumption of the treatment group. Such a decrease is statistically significant at the 0.05 level and the related effect size, computed with Cohen's *d* statistics, is equal to 0.51. Namely, the overall short-term effect of the Social Power intervention is intermediate. In the long-term, Model (II) cannot but confirm the lack of statistical significance of the effect —which is by the way characterised by a low effect size (Cohen's *d* equal to 0.07).

Starting from this model, I thus perform heterogeneity analyses tackling RQ3, to verify if, and to to what extent, the average treatment effect on households treated with the Social Power app is different on sub-groups of households on varying the location (Massagno or Winterthur), the type of household (single adult or family), and the type of building (apartment or house). In all the three cases, I consider the null hypothesis H_0 that the average treatment effect ATT is the same between each couple of sub-groups.

Table 5.15: Output of Model (II) panel regression.

	Model (II)	
	Average weekly consumption [kWh/week]	p value
$TreatxPeriod\_One$	-6.004**	0.01471
	[2.4353]	
$TreatxPeriod\_Two$	1.3243	0.77031
	[4.5287]	
Observations	n=92, unbalanced panel. T=2/3, N=271.	—
Total Sum of Squares	30964	—
Residual Sum of Squares	30279	—
Adjusted R-Squared	-0.50873	—
R-Squared	0.02212	—
F-statistic	1.97926 on 4 and 175 degrees of freedom.	0.14125

R "plm" package; model= "within", effect= "twoways". Heteroskedasticity-robust, clustered standard errors (Arellano method) in parentheses. Significance levels: * 0.1, ** 0.05, *** 0.01

5.5.1 Heterogeneity on varying the location

Finding differences in the treatment effect depending on the city where the intervention took place might indicate that the different cultural background characterising the two locations actually plays a role. The null hypothesis H_0 is that the ATT in Massagno is the same as the ATT in Winterthur, both in the short- and in the long-term. To verify whether differences exist between Massagno and Winterthur, I run the panel data regression model of equation (5.5) on the two subgroups of households of Massagno and Winterthur (Model III) and then perform an interaction test to verify if the treatment effects emerging from the regressions are statistically different from each other.

The outputs of Model (III) regression are reported in Table 5.16 and in equations (5.6) and (5.7). Table 5.20 summarizes the results, by reporting the ATT's 95% confidence intervals, the effect sizes (estimated via Cohen's d estimator), and the outcome of the interaction test comparing the two sub-groups of households.

$$\widehat{kWh\_week\_Massagno} = -3.1125\, TreatxPeriod\_One + \\ +5.8282\, TreatxPeriod\_Two \tag{5.6}$$

$$\widehat{kWh\_week\_Winterthur} = \mathbf{-8.8884}\ TreatxPeriod\_One + \\ -3.2151\, TreatxPeriod\_Two \tag{5.7}$$

The average treatment effect ATT appears to be statistically significant only in Winterthur and, as expected, in the short period only (-8.89 kWh/week, significant at the 0.01 level). In this case, besides statistical significance, the effect size is also large (Cohen's d equal to -0.82). However, the interaction tests based on Christensen et al. (2021), also reported in Table 5.20, indicate that the observed differences in the ATT between the two sub-groups Massagno and Winterthur are not statistically significant. Therefore, I cannot reject the null hypothesis H_0 and conclude that the average treatment effect is not different between the two cities: the cultural context did not play a tangible role in driving the effects of the Social Power intervention.

Table 5.16: Output of Model (III) panel regression - Heterogeneity effects on sub-groups based on the location.

	Model (III)			
	Average weekly consumption [kWh/week]			
	Massagno	p value	Winterthur	p value
$TreatxPeriod\_One$	-3.1125 [3.4538]	0.3700	-8.8884*** [3.1130]	0.00539
$TreatxPeriod\_Two$	5.8282 [7.9195]	0.4368	-3.2151 [4.1366]	0.43916
Observations	n=92, unbalanced panel. T=2-3, N=271.	—	n=46, unbalanced panel. T=2-3, N=136.	—
Total Sum of Squares	21452	—	8742.1	—
Residual Sum of Squares	21006	—	8276.8	—
Adjusted R-Squared	-0.56146	—	-0.48621	—
R-Squared	0.032825	—	0.053228	—
F-statistic	0.902963 on 2 and 85 degrees of freedom.	0.40922	2.41747 on 2 and 86 degrees of freedom.	0.09518

R "plm" package; model= "within", effect= "twoways".
Heteroskedasticity-robust, clustered standard errors (Arellano method) in parenthesis.
Significance levels: * 0.1, ** 0.05, *** 0.01

5.5.2 Heterogeneity on varying the building type

I now investigate if the average treatment effect produced by use of the Social Power app changes on varying the building type, namely "apartment" or all the other building types, which I classify as "house". The null hypothesis H_0 is that the ATT in the sub-group of "apartment households" is the same as the ATT in the sub-group of "house households", in the short- and in the long-term. However, I expect to find differences between the sub-groups. I suspect in fact that those living in "houses" have on average larger homes, which in turn may also mean they are equipped with more appliances, electronic devices, and lighting. This might not only imply a larger electricity consumption; it may also correspond to larger room for saving.

To verify whether differences exist between apartments and houses, I run the panel data regression model of equation (5.5) on the related sub-groups of households (Model IV) and then perform an interaction test to verify if the treatment effects emerging from the regressions are statistically different from each other.

The outputs of Model (IV) regression are reported in Table 5.17 and in equations (5.8) and (5.9). Table 5.20 summarizes the results, by reporting the ATT's 95% confidence intervals, the effect sizes (via Cohen's d estimator), and the outcome of the interaction test comparing the two sub-groups of households.

$$\begin{aligned} kWh\_\widehat{week}\_House = & -6.77\ TreatxPeriod\_One + \\ & +3.08\ TreatxPeriod\_Two \end{aligned} \tag{5.8}$$

$$\begin{aligned} kWh\_\widehat{week}\_Apartment = & \mathbf{-5.70}\ \ TreatxPeriod\_One + \\ & 0.67\ TreatxPeriod\_Two \end{aligned} \tag{5.9}$$

As expected, the average treatment effect (ATT) is statistically significant in the short period only. However, it is only significant in apartments (-5.70 kWh/week, significant at the 0.01 level), where it also has a large effect size (Cohen's *d* equal to 0.80). For households, the ATT is larger in absolute terms (-6.77 kWh/week), however it has a small effect size (Cohen's *d* equal to 0.35) and is not statistically significant, even at the 0.1 significance level.

This appears to be contrary to my expectations: the saving effect seems to be larger in apartments. To verify the statistical significance of such differences, I perform the interaction test based on the approach by Christensen et al. (2021) and compare the ATTs in the two sub-groups "apartment" and "house". The result indicates that the observed differences in the ATTs are not statistically significant, even at the 0.1 level ((p-values equal to 0.98873 and 0.8063, respectively for the short- and the long-term). Therefore, I cannot reject the null hypothesis H_0 and conclude that the average treatment effect does not differ between the two types of building: possible home size and use of more appliances or lighting did not play a tangible role in driving the effects of the Social Power intervention.

Table 5.17: Output of Model (IV) panel regression - Heterogeneity effects on sub-groups based on the building type.

	Model (IV)			
	Average weekly consumption [kWh/week]			
	House	p value	Apartment	p value
$TreatxPeriod\_One$	-6.7664 [7.3325]	0.3609	-5.69869*** [1.72868]	0.00127
$TreatxPeriod\_Two$	3.0810 [8.2398]	0.7102	0.66157 [5.43092]	0.90324
Observations	n=26, unbalanced panel. T=2-3, N=76.	—	n=66, unbalanced panel. T=2-3, N=195.	—
Total Sum of Squares	10698	—	20163	—
Residual Sum of Squares	10383	—	19766	—
Adjusted R-Squared	-0.43461	—	-0.54556	—
R-Squared	0.083444	—	0.020084	—
F-statistic	0.69727 on 2 and 46 degrees of freedom.	1.2563	1.4426 on 2 and 125 degrees of freedom.	0.2883

R "plm" package; model= "within", effect= "twoways".
Heteroskedasticity-robust, clustered standard errors (Arellano method) in parenthesis.
Significance levels: * 0.1, ** 0.05, *** 0.01

5.5.3 Heterogeneity on varying the household type

Finally, I investigate if the average treatment effect produced by use of the Social Power app changes on varying the household type, namely the composition of the household in terms of the characteristics of its members. The two available categories are "single adults" and "family", which includes households with at least one son or daughter, independently on their age. The null hypothesis H_0 is that the ATT in the sub-group of "single adult" households is the same as the ATT in the sub-group of "family" households, both in the short- and in the long-term. Again, I expect to find differences between these sub-groups: families may have a larger set of electricity needs to be satisfied, and therefore offer more room for saving. I thus expect I will be able to reject the null hypothesis.

To verify whether differences exist between these sub-groups, I run the panel data regression model of equation (5.5) on the related subgroups of households (Model V) and then perform an interaction test to verify if the treatment effects emerging from the regressions are statistically different from each other. The outputs of Model (V) regression are reported in Table 5.18 and in equations (5.10) and (5.11).

$$\widehat{kWh\_week\_Single\_adult} = -4.12\ TreatxPeriod\_One + \mathbf{12.36}\ TreatxPeriod\_Two \tag{5.10}$$

$$\widehat{kWh\_week\_Family} = \mathbf{-7.38}\ TreatxPeriod\_One + -5.04\ TreatxPeriod\_Two \tag{5.11}$$

Table 5.18: Output of Model (V) panel regression - Heterogeneity effects on sub-groups based on the household type.

	Model (V)			
	Average weekly consumption [kWh/week]			
	Single adult households	p value	Family households	p value
TreatxPeriod_One	-4.1240 [4.5566]	0.36877	-7.3802*** [2.5542]	0.00468
TreatxPeriod_Two	12.3605** [7.1353]	0.08796	-5.0425 [5.6342]	0.37282
Observations	n=35, unbalanced panel. T=2-3, N=104.	—	n=57, unbalanced panel. T=2-3, N=167.	—
Total Sum of Squares	15119	—	15209	—
Residual Sum of Squares	13871	—	14806	—
Adjusted R-Squared	-0.43461	—	-0.55175	—
R-Squared	0.12252	—	0.027818	—
F-statistic	2.92532 on 2 and 65 degrees of freedom.	0.06074	1.44265 on 2 and 106 degrees of freedom.	0.2409

R “plm” package; model= “within”, effect= “twoways”.
Heteroskedasticity-robust, clustered standard errors (Arellano method) in parenthesis.
Significance levels: * 0.1, ** 0.05, *** 0.01

Table 5.20 summarizes the results, by reporting the ATT’s 95% confidence intervals, the effect sizes (estimated via Cohen’s d estimator), and the outcome of the interaction test comparing the two sub-groups. In the short-term, model V outcomes seem to confirm my hypotheses on the differences between the ATTs: a higher (and statistically significant at the 0.01 level) ATT appears in the “family” sub-group: -7.38 kWh/week, effect size 0.76, compared with 4.12 kWh/week and 0.28 effect size in the single adult sub-group, for which no statistical significance appears. However, the interaction test, always performed via the Christensen et al. (2021) approach, indicates the observed differences between the two sub-groups are not statistically significant, even at the 0.1 level (p-value= 0.5530). Therefore, I cannot reject the null hypothesis H_0 and conclude that in the short-term the average treatment effect does not differ between the two types of household: in the short-term, the household composition did not play a tangible role in driving the effects of the Social Power intervention.

For the the long-term, instead, I find definitely unexpected results: the ATT in the “single adult” subgroup is statistically significant at the 0.05 level, and it shows an increase in consumption, by 12.36 kWh/week, with an estimated effect size equal to 0.55 (medium level of Cohen’s *d* estimator). Moreover, the interaction test performed between the sub-groups indicates that the difference between this ATT and the one for families (-5.04 kWh/week, 0.21 Cohen’s *d*) is statistically significant at the 0.1 level (p-value equal to 0.0056). This is not coherent with previous model outcomes, therefore I explore this result in more details.

Single adults sub-groups are respectively made of n = 16 (treatment) and n = 19 (control) households. Among the treatment sub-group, there are three outliers (Figure 5.8), one of which is particularly relevant in terms of its distance from the other sub-group members: for this household, a 250% increase in electricity consumption is observed between Period Zero (the baseline) and Period Two (follow-up). Considering the small sample size, such a large increase in electricity consumption has a leverage effect and drives an increase in the average consumption by the whole sub-group and in the related ATT, which is found to be statistically significant by the t-test routinely performed by the R software package estimating the coefficients of the panel regression model.

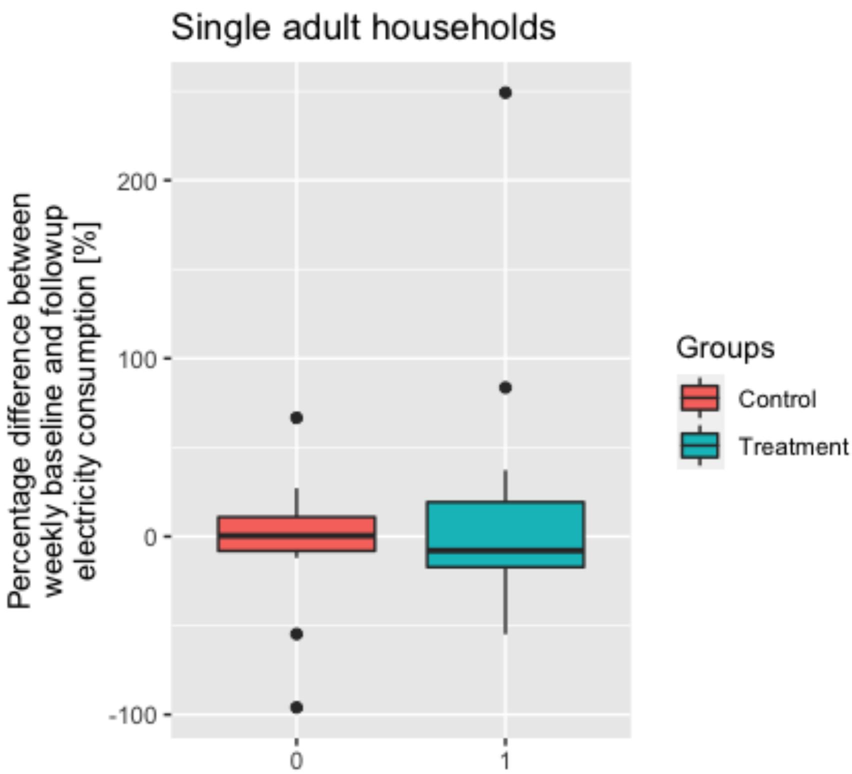

Figure 5.8: Percentage differences between baseline and follow-up average weekly consumption, in the "single adult" household sub-groups.

To get more insights on this unexpected result, I replicate the ATT estimate by adopting a Difference-in-Differences approach, which in fact produces the same result for the ATT. Analysis of the distributions of the variable «difference between electricity consumption in the baseline and followup period» for the treatment and control sub-groups of single adult households, however, shows they are not normally distributed (Figure 5.9), which is also confirmed by two Shapiro-Wilk normality tests, whose p-values are 0.0034 for the treatment sub-group and 0.0002137 for the control sub-group. Considering that this is a small-n case (the overall sample size is equal to 35, slightly above the n = 30 sample size which is usually suggested for use of t-test by relying on asymptotic normality conditions), I therefore repeat the test for statistical significance of the ATT by using a non-parametric Mann Whitney U Test. The p-value resulting from this test is equal to 0.7722, which indicates a lack of statistical significance, at the 0.1 level. Namely, by adopting a non-parametric test, statistical significance does not appear any longer.

Figure 5.9: Distribution of variable "Difference between baseline and followup average weekly consumption" for the treatment and control sub-groups of households of category "Single adult".

Since I think a non-parametric test is more appropriate for such a small-n sample characterised by non-normality of distributions, I rely on the latter estimate of the statistical significance of the ATT. By means of such a closer look, therefore, the very high ATT that I found appears to be due to chance alone (namely, it cannot be associated with statistical significance). Specifically, it appears to be driven by the high consumption of the above identified very large outlier.

I thus perform the analyses by removing the "extreme outlier" characterised by a 250% increase in consumption between baseline and followup periods. The results of this final model (Model VI) are reported in Table 5.19 and then also included in summary Table 5.20. Statistical significance of the long-term ATT in "Single adult" households now disappears, as well as the significance of the interaction test between the sub-groups of "Single adult" and "Family" households for the long-term (p-value= 0.1095124). Also, the Cohen's *d* effect size, decreases to -0.28 (short-term) and 0.43 (long-term).

By removing the extreme outlier in "Single households" sub-group, therefore, results are more coherent with results of the Difference-in-Differences estimator: no statistically significant difference appears in ATT among the sub-groups and the ATT in the "Single adults" sub-group is not significant either. Anyway, these results do not support my initial expectations, and I cannot reject the null hypotheses that the short- and long-term effects measured in the two sub-groups are the same.

Table 5.19: Output of Model (VI) panel regression - Heterogeneity effects on sub-group based on the household type, without the extreme outlier.

	Model (VI)			
	Average weekly consumption [kWh/week]			
	Single adult households (no outlier)	p value	Family households	p value
$TreatxPeriod\_One$	-4.2308 [4.5929]	0.3605	-7.3802*** [2.5542]	0.00468
$TreatxPeriod\_Two$	7.8708 [5.7760]	0.1778	-5.0425 [5.6342]	
Observations	n=34, unbalanced panel. T=2-3, N=101.	—	n=57, unbalanced panel. T=2-3, N=167.	—
Total Sum of Squares	11339	—	15209	—
Residual Sum of Squares	10720	—	14806	—
Adjusted R-Squared	-0.50066	—	-0.55175	—
R-Squared	0.054586	—	0.027818	—
F-statistic	1.81874 on 2 and 63 degrees of freedom.	0.17065	1.44265 on 2 and 106 degrees of freedom	0.2409

R "plm" package; model= "within", effect= "twoways".
Heteroskedasticity-robust, clustered standard errors (Arellano method) in parenthesis.
Significance levels: * 0.1, ** 0.05, *** 0.01

5.6 Conclusions

In this Chapter I have presented the persuasive, gamified Social Power app, by framing its features according to a stage model of behaviour change and principles for persuasive systems design. I also introduced the Social Power policy intervention, which took place in the two Swiss municipalities of Massagno and Winterthur in Spring 2016. Finally, I presented the novel panel regression data analyses I performed in order to assess the short and long term effects of the Social Power intervention and to identify possible heterogeneous effects among the sub-groups of treated households, thus respectively tackling my RQ1, RQ2, and RQ3 research questions. Outcomes of the analyses I developed are summarised in Table 5.20, which reports the estimated ATTs and effect sizes for each group and sub-group of households I could consider for heterogeneity analysis.

The Social Power intervention field tested two slightly different versions of the related app: despite most of the key persuasive elements were the same (individual electricity consumption feedback, challenges, tips, points, notification system, Facebook page, and monthly quizzes), one app version was characterised by a collaborative gamified structure, while the other one was characterised by a competitive gamified structure. To assess the intervention's ATT, I adopted a quasi-experimental approach, by comparing a group of voluntary, self-selected households engaged in the treatment groups (a collaborative and a competitive treatment group, each of n = 23 households, to which the self-selected households were randomly assigned), against a group of comparable control households (n = 46), selected on a posterior basis via matching techniques.

The ATT can be measured both in the short-term (during the Social Power intervention) and in the long-term (one year after the end of the intervention), thanks to the availability of weekly electricity consumption data for these periods and for a comparable baseline period, for all the involved households. With the available data I fed a Two Way Fixed Effects panel regression model, which showed that the intervention effect was actually not statistically different between the collaborative and competitive treatment groups (Model I). I therefore opted for considering the two collaborative and competitive groups as a whole. Due to the Social Power small sample sizes, in fact, a unified treatment group is more appropriate to address my research question on the heterogeneity of effects.

By comparing electricity consumption data between the unified treatment and the control group during the intervention (Model II, again via a TWFE panel regression), I found a statistically significant ATT, equal to a reduction of 6 kWh/week (0.05 significance level). This corresponds to 9.23 % electricity and CO_2 emission saving, with respect to the baseline average value of the treatment group. The savings obtained during the intervention (RQ1) have a medium effect size (Cohen's *d* equal to 0.51) and are for instance larger than the average ones obtained in the enCompass case, both in percentage terms and in effect size. However, if the sub-sample of enCompass households that do not use electricity for heating purposes is considered (i.e., if the ATT on the sub-sample of "only electricity households", which is comparable with the households treated with the Social Power app, is considered), enCompass' savings result larger than Social Power ones. This might for instance be due to the personalisation of recommendations offered by enCompass, as well as to the individual, customised goal setting features it offered, which may have resulted in a more engaging experience, favouring higher feelings of autonomy as predicted by the Self-Determination Theory.

In any case, the obtained savings are comparable with the upper end of the range by early smart meter feedback studies reviewed by Darby et al. (2006), Fischer (2008) or Delmas et al. (2013). With respect to the latter, results are closer to studies that the authors consider weaker from the methodological point of view. However, just like enCompass, also in this case I devoted particular care to ensure a rigorous evaluation procedure. Furthermore, again like enCompass, in the long-term the statistical significance disappears (RQ2).

Having obtained such results, I further explored the ATT, in order to identify possible —if any— differences between the sub-groups of treated households, both in the short- and in the long-term (RQ3). For this purpose, despite the low sample size and the related low statistical power, I performed heterogeneity analyses that consider sub-group differences due to available observed characteristics of the households, namely the city where they live (municipality of Massagno or of Winterthur), the type of building (apartment or house), and the type of household (single adults or families). For these analyses I adopted the same TWFE panel regression models, by running them on the specific sub-groups of target households (Model III, IV, and V). In all cases, the regression models indicate that in the short-term the ATT is statistically significant in one sub-group only (respectively, households living in the municipality of Winterthur, apartments, families), always at the 0.01 significance level. *Post-hoc* interaction tests conducted on pairwise comparisons of the sub-groups of households, however, have shown that those ATTs are not significantly different from each other, even at the 0.1 significance level. Furthermore, in the long-term, no statistically significant ATTs have emerged, even in the sub-group analysis.

Indeed, in one case (single adults vs families) the regression model has shown a statistically significant difference for the long-term, which is related to an apparently relevant increase in consumption (ATT for single adult households equal to +12.36 kWh/week, corresponding to a 23.30% increase compared with the baseline of treated households of that sub-group, 0.55 effect size). I have however shown that this is mostly due to the presence of an “extreme outlier” in the sub-group of “single adult” treated households. By removing this outlier (Model VI), statistical significance of the long-term ATT disappears, as well as the differences between the two sub-groups in the *post-hoc* pairwise comparison test. Further investigation would be needed to understand the reasons for presence of such an extreme outlier. Specifically, I cannot totally exclude that the observed increase in electricity consumption is due to a change in the composition of the household itself. For instance, the single adult might have become a couple, or a son or daughter might have been born. Unfortunately, households characteristics were provided by the utility companies at the start of the field intervention and not updated over time, therefore I could not check this hypothesis.

I thus conclude that location, type of household and type of building do not drive different magnitudes of the electricity and CO_2 saving effects produced by use of the Social Power app: the amount of savings is not directly influenced by these factors. The heterogeneity analyses I performed therefore suggest there are no specific household profiles that app-based electricity and CO_2 saving interventions should directly and actively target. Or, better: under the limited number of household characteristics I could observe, no relevant profiles to focus on emerge.

The lack of statistical differences on the effects on varying the location, in particular, provides a useful insight from the policy-making perspective: as there is no treatment effect heterogeneity, I conclude that the treatment effect can be generalised across the sites. This is a relevant finding, which suggests that, if app use were scaled-up to other regions and contexts, similar effects would likely be obtained.

Nevertheless, the problem of lack of persistence of the effects over time remains open: the persuasive, gamified elements focusing on social interactions and social norms that were introduced in Social Power did not manage to permanently produce a change in behaviour and practices. Or, better: the way such social interactions and norms were exploited in the Social Power app did not result to be effective. Two comments in the questionnaire performed at the end of the three-month intervention period (Period One), not analysed in details here, were indeed very explicit about the weaknesses of the way social interaction features are implemented in the app: “I wasn’t clear about the whole game: who is on my team and how to contact them?” and “ I found it very difficult to orient myself. What is where? What does what mean? Which team am I on? How do I get in touch with the team? What is social about the app?”.

Similarly to the case of enCompass, the relapse to the previous behaviour in the maintenance phase might be due to a decrease in Perceived Behavioural Control or in subjective norms (as suggested by the Theory of Planned Behaviour). However, these findings also suggest that effectively intervening on electricity consumption would benefit from a deeper and more direct and explicit engagement with conventions and social norms, based on richer social interaction possibilities. I expect that directly addressing such conventions, for instance by providing venues for open discussion and exchanges about them, could favour the consolidation of new behaviour and practices over time and contribute to the long-term maintenance of the electricity and CO_2 saving effects observed during direct app use.

The Social Power case provides insights into another relevant aspect: the target number of households to be involved in field activities (set to n = 100 in each city) was not reached: overall, only 108 households were identified at the end of the recruitment period. This can be regarded as an indication of limited interest by the population especially in the city of Massagno, were the eligibility requirements were less strict and a broad number of households was meeting them. The case of enCompass was not different (though in that case the population of eligible households was smaller): instead of the 100 target households, the intervention started with 75 recruited households. Furthermore, in Social Power the number of engaged households sharply decreased between the recruitment period and the intervention period. Despite participation to field activities was voluntary, and households were free to opt-in into project activities or simply ignore the invitation to do so, only 43% of the households that had indicated their willingness to join project activities then dowloaded the Social Power app and at least registered on it. For the enCompass case, non-compliance issues were smaller, since only 10 households out of the 75 voluntary recruited never logged into into the app (13% non-compliance rate).

Difficulties in recruiting households and the very high non-compliance rate observed in Social Power suggest that the interest in app-based types of interventions is still limited among the population. Particularly, even if such apps were found to be effective also in the long-term, based on the collected evidence I argue that specific additional motivational approaches would be needed in order to favour their spontaneous and voluntary usage and genuine interest in the population. These findings therefore indicate a clear obstacle to a possible large-scale deployment of app-based persuasive interventions and suggest to carefully evaluate their actual energy saving and climate change mitigation potential, beyond the average effect found on the small samples of households that were actually exposed to the treatment.

Table 5.20: Summary of the model outcomes estimating the ATTs of the Social Power intervention, including heterogeneity analyses.

		Period One (intervention)		Period Two (One year after)	
		Consumption ATT and C.I. [kWh/week]	Effect size (Cohen's d)	Consumption ATT and C.I. [kWh/week]	Effect size (Cohen's d)
Model (I)	Collaborative households ($n_T = 23$, $n_C = 46$)	**-6.88**** [-12.557, -1.206]	-0.52 medium	-1.60 [-11.553, 8.348]	0.07 negligible
	Competitive households ($n_T = 23$, $n_C = 46$)	**-5.12**** [-10.077, -0.161]	-0.41 small	4.36 [-7.774, 16.503]	0.07 negligible
Model (II)	All households ($n_T = 46$, $n_C = 46$)	**-6.004**** [-10.843, -1.165]	-0.51 medium	1.3243 [-7.674, 10.323]	0.07 negligible
Model (III)	A. Massagno households ($n_T = 23$, $n_C = 23$)	-3.11 [-10.078, 3.853]	-0.58 medium	5.83 [-10.143, 21.790]	0.21 small
	B. Winterthur households ($n_T = 23$, $n_C = 23$)	**-8.89***** [-15.166, -2.610]	-0.82 large	-3.21 [-11.557, 5.127]	-0.18 small
	Post-hoc interaction test between sub-groups A, B	p-value: 0.21416 No significant differences		p-value: 0.31147 No significant differences	
Model (IV)	C. House ($n_T = 16$, $n_C = 19$)	-6.77 [-21.935, 8.402]	-0.35 small	3.08 [-13.964, 20.126]	0.22 small
	D. Apartment ($n_T = 30$, $n_C = 27$)	**-5.70***** [-19.479, 5.955]	-0.80 large	0.66 [-11.006, 17.167]	-0.18 small
	Post-hoc interaction test between sub-groups C, D	p-value: 0.8873 No significant differences		p-value: 0.8063 No significant differences	
Model (V)	E. Single adult ($n_T = 16$, $n_C = 19$)	-4.12 [-15.250, 5.957]	-0.28 small	**12.36****[1] [-9.563, 18.741]	0.55 medium
	F. Family ($n_T = 30$, $n_C = 27$)	**-7.38***** [-12.501, -2.260]	-0.76 nearly large	-5.04 [-18.610, 6.253]	-0.21 small
	Post-hoc interaction test between sub-groups E, F	p-value: 0.5330 No significant differences		p-value: 0.0556 **E. vs F.** **	
Model (VI)	G. Single adult without extreme outlier ($n_T = 15$, $n_C = 19$)	-4.23 [-13.598, 5.136]	-0.28 small	7.87 [-9.563, 18.741]	0.43 small
	H. Family ($n_T = 30$, $n_C = 27$)	**-7.38***** [-12.501, -2.260]	-0.76 nearly large	-5.04 [-18.610, 6.253]	-0.21 small
	Post-hoc interaction test between sub-groups G, H	p-value: 0.5490 No significant differences		p-value: 0.1095 No significant differences	

[1] As indicated in the text, by adopting a more appropriate Mann Whitney U Test, statistical significance is not found, even at the 0.1 level.
Significance levels: * 0.1, ** 0.05, *** 0.01

6 Case three: Social Power Plus

> *Real dialogue isn't about talking to people who believe the same things as you.*
>
> — **Zygmunt Bauman**
> Sociologist and philosopher

In this chapter I introduce and analyse the case of Social Power Plus (from now onwards, SPP), which is an ongoing research project that I am involved into, developed in collaboration with researchers from the Zurich University of Applied Sciences (ZHAW, the same team as the Social Power project) and three utility companies operating in the Swiss-German part of Switzerland. The project builds on the experience of Social Power and, learning from its results, aims at favouring the diffusion of energy-sufficient routines in households by means of a novel combination of app-based motivational features, which again puts social interactions at its core.

SPP aims at creating a virtual community of households collectively engaged for the energy transition, by providing its users with novel spaces and opportunities for interaction around energy-related topics. Besides offering energy consumption feedback at the household level, in fact, the SPP app also addresses single households through challenges aimed at re-crafting eight specific energy related routines towards energy sufficiency (heating, showering, washing, cleaning, cooking, dishwashing, studying and working, recreation) and supports dialogue between app user households through an internal forum ("pinboard"), which is also backed-up by monthly online meetings. Such features draw on the “everyday behaviour” perspective conceptualised by Kaufman et al. (2021) (see Section 2.9) and on interventions inspired by Social Practice Theories within living lab contexts (Sahakian, Rau, et al., 2021; Matschoss et al., 2021). The rationale behind these features is that impact-focused challenges and interaction in the app pinboard might trigger peer-to-peer learning opportunities, resulting in the creation of novel social norms and competences around energy-sufficient household energy consumption.

An additional element characterising SPP is that it deals with energy consumption for both heating and non-heating purposes. Thus, similarly to enCompass, it accounts for the largest share of household's energy demand and carbon emissions, which is heating. To support households in reducing their heating demand, SPP also provides the estimated breakdown of energy consumption into heating (including hot-water production) and non-heating purposes (electric appliances, lighting, computers, etc.), which is produced by an algorithm developed on purpose, and addresses heating and showering routines by household members through dedicated challenges.

At the time of writing (Summer - Fall 2022), the SPP app is being tested in a field intervention involving about 200 voluntary households in the three German-speaking Swiss regions of Schaffhausen, Wil, and Winterthur. Details of the field intervention and the related research design are provided in Section 6.3. As explained there, within the time-frame of this dissertation, analysis of energy consumption data automatically measured by smart meters is not possible. For this case therefore I do not estimate the app's quantitative treatment effect on energy consumption and carbon emissions. The case is however well-suited to address my RQ4 research question ("Which app features can foster higher user engagement, thus providing greater support to the reduction of energy consumptions and CO_2 emissions?"), as it allows me to perform an in-depth analysis on the role and contribution by the different app's features to performing energy sufficient activities, from the perspective of the user experience and the evaluation by its users. Specifically, by means of two questionnaires administered before and half-way the SPP intervention and by analysing the amount and type of in-app interaction by its users, I collect insights on the user experience, suggesting which app features are most likely to affect household members' routines, even though maybe only on a short-term basis.

SPP is a collaborative research project, in which different research team members equally contribute to a common outcome. The design of the SPP app and the operationalisation of its features, as well as the design of the questionnaires to elicit household's preferences about such features, were performed by the whole research team, including myself. The analyses and results I present in this chapter, instead, are the outcome of my sole work.

6.1 The Social Power Plus persuasive app

The Social Power experience described in the previous chapter taught us that, despite the goals and ambition by the research team, app users could not really interact with other team members, since they did not previously know each other and the app was lacking an internal communication channel. Post-intervention surveys and interviews conducted with project participants, not reported in the previous chapter, in fact indicated that within-team social bonds were perceived as very low and that app users would have appreciated an internal chat, to strengthen feelings of team belonging and affiliation and establish social interactions.

More generally, the Social Power experience suggested us to tackle the call by Buchanan et al. (2015) outlining the need for innovative forms of feedback, that enable actual engagement by households. To increase chances that the app resulted appealing and stimulating for its users, we thus opted for designing the features of the novel SPP app in a participatory co-creation process (Sanders and Stappers, 2008), in order to tailor them to actual needs, constraints, and motivations by real-life households ("living lab").

6.1.1 Co-design in a living lab setting

Living labs are open-innovation processes aimed at co-creating and validating innovation within collaborative, real-world environments (Pallot et al., 2010; Bergvall-Kåreborn et al., 2010; Almirall et al., 2012; Dell'Era and Landoni, 2014). They aim at designing, testing and learning from innovative socio-technical practices (namely, new ways of doing something) in real-world conditions, by engaging a diversity of key stakeholders, including policy-making institutions. They are characterised by situated experimentation, diversity, evaluation and shared learning (Følstad, 2008; Hillgren, 2013; Leminen, 2013; Cellina, Castri, Diethart, et al., 2018). Processes inspired by the living labs approach open up to "participatory mindsets", where users become active partners of the value creation process (Sanders, 2002; Schuler and Namioka, 1993): beyond "designing for the users", living labs support "designing with the users".

Design involving users has been previously applied to energy transition research to improve smart meter-based behaviour change 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), or capture multi-faceted benefits including increasing comfort, energy savings, transparency and overall consumer awareness (Böhm and Szwec, 2013). The co-creation approach, and specifically co-design, involves users during the whole design process, through interviews, surveys, focus groups, or testbed activities (Sanders and Stappers, 2014). Doing so, the product is designed for its intended use and is argued to be ultimately more effective and efficient (Abras et al., 2004). Co-design processes are in fact expected to increase trust in the individuals called to perform a given behaviour and to favour the resulting policy interventions to better fit with the specific contexts in which they are implemented (Della Valle and Bertoldi, 2021).

The co-design process we performed for SPP is presented in details in a paper under review (Wemyss, Lobsiger-Kägi, et al., 2022). It took place by means of three workshops, through which we gathered inputs from around 50 voluntary households, which were identified through an open call, in the same Swiss regions where later on SPP was field-tested: Schaffhausen, Wil, and Winterthur. Besides suggestions for individual goal setting features and customised recommendations to improve the energy efficiency of the household's technical equipments, one major wish from workshop participants was to include a pinboard section into the app, where tips, recommendations, experiences and also fun facts could be shared among the group of app users, engaged together in exploring energy sufficient activities in their homes.

The features resulting from the co-design process are presented in details in the next sections and are summarised in Table 6.1. The Table shows an overview of the Social Power Plus app's features, summarising them from the perspective of both their theoretical background and the persuasive principles and techniques they exploit. As for the previous cases, I refer to the Transtheoretical model of behaviour change (TTM, Prochaska and Velicer, 1997), to the techniques listed in the taxonomy for behaviour change interventions by Abraham and Michie (2008), and to the principles for Persuasive Systems Design (PSD) by Oinas-Kukkonen and Harjumaa (2009).

Table 6.1: Features of the Social Power Plus app targeting households.

Stage	Process	Feature	Technique	PSD framework
Contemplation	*Self-reevaluation* Cognitively and affectively assessing one's self-image, with and without a particularly unhealthy habit	Individual energy consumption feedback Energy consumption change compared with individual baseline Feedback on energy consumption change by comparable households	3. Provide information about others' approval 12. Prompt self-monitoring of behaviour 13. Provide feedback on performance 19. Provide opportunities for social comparison	Self-monitoring Social comparison Normative influence
Preparation	*Self-liberation* Believing that one can change and commiting to act on such a belief	Goal setting Challenges	4. Prompt intention formation 10. Prompt specific goal setting 7. Set graded tasks	Reduction Personalization
Action and Maintenance	*Counterconditioning* Learning of more sustainable behaviours that can substitute the less sustainable ones	Challenges Energy saving tips Pinboard	8. Provide instruction	Reduction Suggestion
	Contingency management Providing consequences (rewards) for taking steps in a particular direction	Goal achievement congratulation messages	13. Provide feedback on performance 14. Provide contingent rewards	Praise
	Helping relationship Providing social support (caring, trust, general support) for new behaviour	Pinboard Regional energy saving competition in teams Online meetings	6. Provide general encouragement 19. Provide opportunities for social comparison 20. Plan social support or social change	Social learning Social comparison Normative influence Social facilitation Cooperation Competition Recognition
		Notification and reminder system to stimulate action maintenance	17. Prompt practice 18. Use follow-up prompts	Reminders

6.1.2 Contemplation stage

According to TTM, the behaviour change process starts from a pre-contemplation stage, in which households have no motivation for reducing their energy consumption and do not intend to take action to change. To support them towards change, Prochaska and Velicer (1997) suggest to activate a *consciousness raising* process, which increases awareness of causes, consequences, and cues about a behaviour. In SPP, however, we opted for not including features aimed at raising awareness of the need for behaviour change, which were instead delivered through the communication material accompanying the launch of the app, for the enrolment of its users. Voluntary households willing to use SPP are therefore expected to have a certain level of awareness of the need for change, and therefore the app aims at directly addressing households in the contemplation stage.

To persuade households to take action instead of "chronically contemplating" an abstract idea of change, SPP provides detailed feedback on the households' energy consumption. First, in "My Home" section, which is the landing page when the app is opened, it provides daily feedback. Thanks to a direct connection with smart meters automatically recording energy consumption, the household receives feedback on the evolution of its consumption, compared to its own historical baseline, measured over a comparable period. If consumption is higher than the baseline, the app background is red; if it is smaller, the app background is green (Figure 6.1). This provides immediate evaluation metrics, letting app users immediately understand if 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 be shown green backgrounds only.

Furthermore, "My Home" provides a comparison with other similar households, by reporting the average change in consumption by similar[1] households of the SPP community. In this case, SPP exploits a descriptive norm, showing the energy consumption performances by the other households. Finally, SPP shows a simple bar chart reporting the household's hourly consumption in the previous twenty-four hours, in the last week, or the weekly consumption since the beginning of app use. Through this piece of information, household members start self-discovering 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 –and vice-versa. Users can also visualise an estimate of their disaggregated consumption for heating and non-heating purposes, which is obtained by non-intrusive load monitoring algorithms (P. Huber et al., 2021).

The combined feedback on individual energy consumption and change compared to the baseline, as well as the comparison with similar households, is expected to activate a *self-reevaluation process* (Prochaska and Velicer, 1997): household members start to cognitively and affectively assess their own self-image, under both current and novel energy consumption behaviours, leading them to energy saving outcomes. If this process is properly activated, household members enter the preparation stage.

[1]Three household categories are considered: <65 year old adults, >65 year old adults, families with kids.

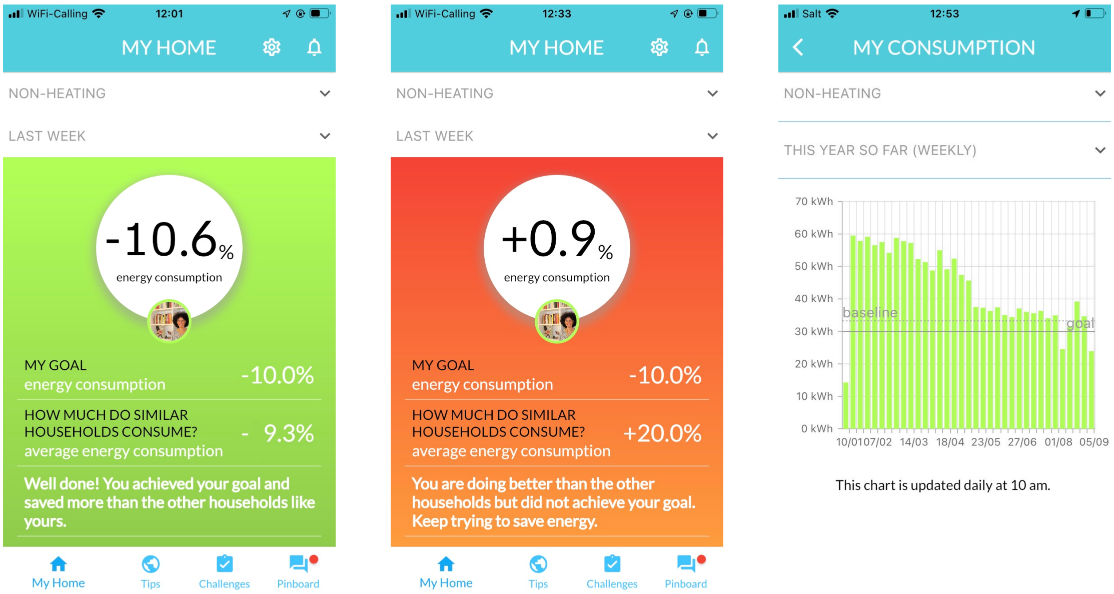

Figure 6.1: Social Power Plus pages showing individual energy consumption and saving feedback.

6.1.3 Preparation stage

Household members start developing concrete plans for action and commit to stick to them (*self-liberation* process, according to Prochaska and Velicer, 1997). Social Power Plus supports them by prompting goal setting: every household is invited to specify its own energy saving target, for heating and non-heating purposes (Figure 6.2).

In order to support household members in achieving their goal, Social Power Plus invites its users to join energy sufficiency challenges, which aim at modifying dominant household practices (heating, showering, washing, cleaning, cooking, dishwashing, studying and working, enjoying recreation time) and questioning current conventions on comfort, convenience, or cleanliness. Challenges also provide suggestions on how to re-craft household routines towards energy sufficiency. For instance, challenges aim at reducing room thermostat settings by a few degrees, at reducing the number of weekly laundry/washing machine cycles, or at enjoying digital-detox free-time at home, without using electronic devices.

Each challenge lasts for two weeks and every two weeks new challenges are released, dealing with a different household routine. All household members can potentially engage in each challenge, by contributing to practice re-crafting within the household. Households are free to ignore the challenges or to engage in more of them at the same time. The completion of challenges sometimes during the fortnight is thus self-regulated and personalised: 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 in fact start if households commit, via the app, to achieve a specific target they set for themselves, such as for instance reducing indoor temperature by three Celsius degrees, reducing the number of washing machine cycles to two per week, or at enjoying at least three "digital-free" nights per week.

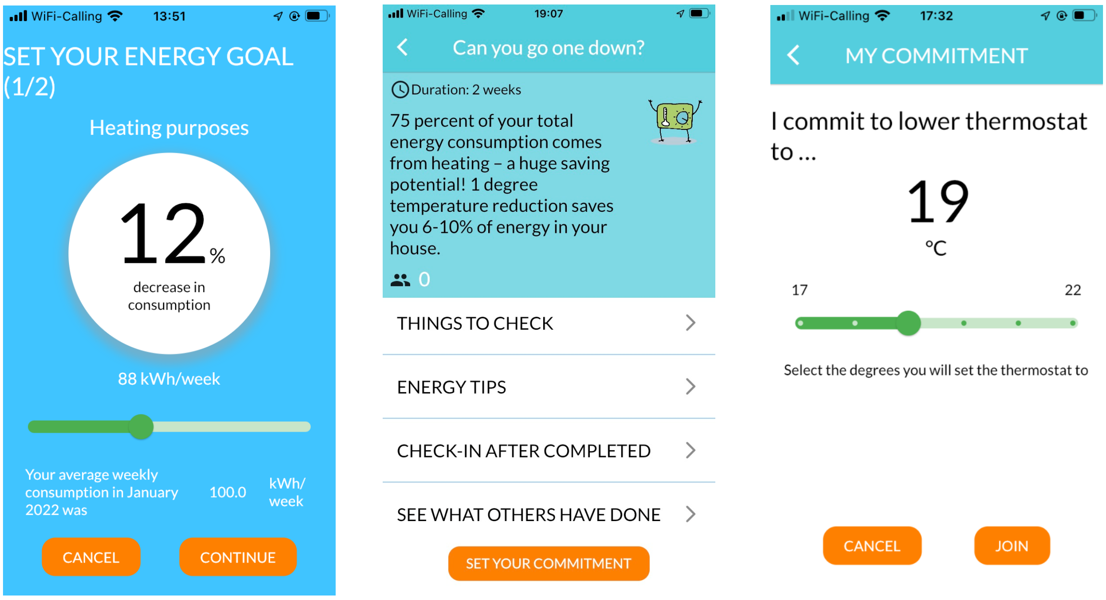

Figure 6.2: Social Power Plus pages providing individual goal setting opportunities and challenge introduction and commitment.

6.1.4 Action and maintenance stages

Once households have set their overall energy saving goal and have committed to a challenge target, they need to start taking action. To support them, Social Power Plus accompanies each challenge with a list of (non-customised) tips, which suggest simple energy-sufficiency actions or lower energy demand ways to perform the related household routines, inspired from a literature research on household practices (Figure 6.3). Every two weeks, when new challenges about a new practice are released, the related tips are released as well, and then they remain always accessible via the app. Following Prochaska and Velicer (1997), challenges and tips activate a *counterconditioning* process, namely they reduce complexity of the whole behaviour change process, and help households to learn of more sustainable behaviours, that can substitute the less sustainable ones.

Besides such a support by the project team via challenges and tips, SPP aims at leveraging social interactions and at activating a social learning process between the peer households engaged in app use. For this purpose, an app section is dedicated to a "pinboard" wall, similar to a virtual forum, where app-users are invited to post their comments, suggestions, as well as success or failure experiences about the challenge achievement process (Figure 6.3). Everything posted in the pinboard is visible to any app users, and whenever a new message is posted, a visual notification is shown in the app (though no push notifications are sent through the smartphone notification system, to avoid annoying app users). App users can also "like" messages they appreciate or respond to them, by activating an asynchronous dialogue with their peers. The pinboard therefore allows the creation of a feeling of *helping relationships* for new behaviours, which from the Transtheoretical model perspective is essential for action maintenance over time.

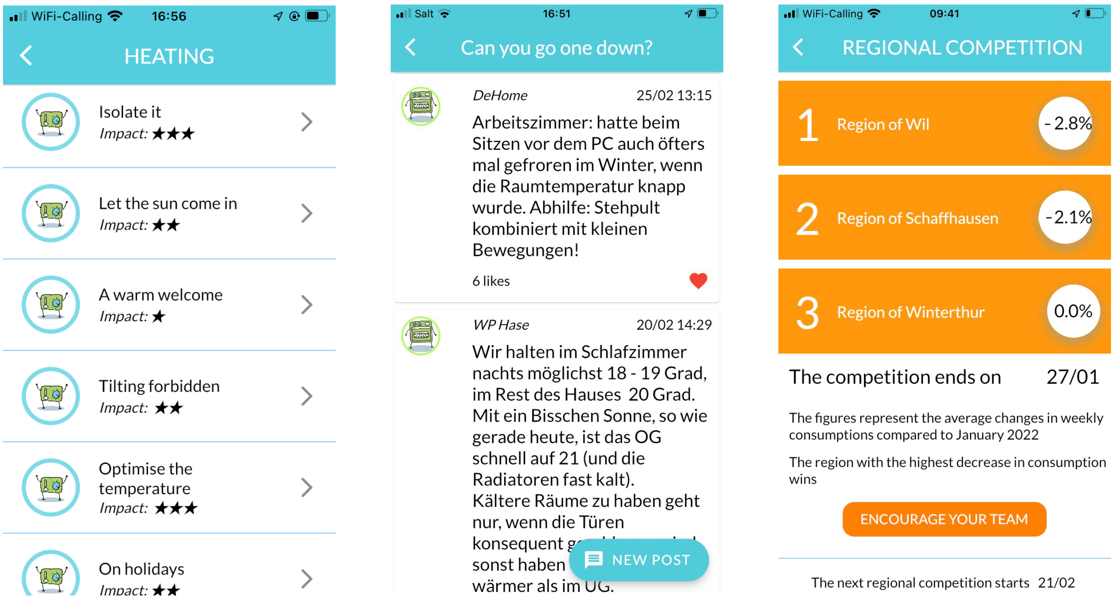

Figure 6.3: Social Power Plus sections providing tips, social interaction possibilities via the pinboard, and the regional energy saving competition.

Specifically dealing with digitally-enabled transition processes, a review on the effectiveness of ICT-based health behaviour change interventions (Morrison, 2015) has remarked the importance of embedding social support features into the digital interventions themselves, arguing that such features can work as realistic substitutes for face-to-face interaction, provided that they are properly moderated and sufficiently flexible to adapt to their users' changing needs. Social features can be especially effective when users perceive insufficient peer support outside the intervention, which is the case of SPP, since dominant social norms do not support energy-sufficient routines.

The pinboard is structured in eight topics, plus an additional one regarding "Anything about Social Power Plus". To incentivise app users to interact with each other via the pinboard and to fully exploit it as an enabler of social learning opportunities and shaper of new shared norms and conventions, it is directly connected with challenges. At the end of the two-week period, in fact, the challenge can be concluded. For this purpose, no automatic control procedure verifies if household members have actually met the challenge commitment. Instead, to complete the challenge app users are asked to upload in the "pinboard" section a short message or picture about their experience when tackling the challenge, their reflections on the social conventions associated with the related practice, and possibilities to re-craft it to save energy and emissions. Through this mechanics, the pinboard can remain alive over time —provided that app users are willing to share their experiences with other peer users, most likely unknown to each other.

To favour such exchanges and support the creation of a (virtual) community feeling and enhance the sense of belonging to it, pinboard-mediated social interactions are also coupled with less structured interaction possibilities, via monthly meetings. Those meetings, originally programmed as in-person meetings and then transformed into online meetings due to the COVID-19 pandemics, offer opportunities for exchange with and clarification by the project team (technical problems related to the proper working of the whole system are not uncommon), but also for informal interaction between app users, that are free to join them at any time during the two-hour evening time slot devoted to each meeting. These meetings were managed by the ZHAW research group and were not recorded. Therefore, unfortunately I cannot include the related materials in my analysis about the effectiveness of the specific app features.

A "booster" feature is then activated once a month: leveraging the sense of belonging to one's region, a weekly "regional energy saving competition" is launched between the households living in different regions. All households of the same region are automatically put into their region's team and the app automatically computes the amount of energy saved by the regional teams (average of savings by each team members). The highest saving region wins the regional competition. No real-life prizes are available, though notifications in the regional energy saving competition section as well as in the pinboard congratulate the winner team, thus providing a virtual reward and public recognition of the obtained results, which are expected to help keep the interest in SPP high.

App users get a regular feedback on the effects of their action, through the goal achievement feedback that is shown in the home page (Figure 6.1): if they achieve their energy saving goal, a congratulation message appears; otherwise, the message incites them to keep efforts to save energy. This allows to activate a *contingency management* process, using the language by Prochaska and Velicer (1997), that rewards households for their actions. Finally, in order to support maintenance of the new behaviour over time, a notification system provides by-monthly reminders about energy-saving topics, such as short news about energy-related events, or additional tips and recommendations over a selection of topics. Overall, the combination of challenges, tips, pinboard, regional energy competitions, and notifications is expected to support households throughout the action and maintenance phases, until new behaviours are set and permanently implemented.

6.1.5 Bridging over behavioural and social practice theories

Challenges and pinboard features are key to the design of SPP. From the perspective of Theory of Planned Behaviour, challenges directly contribute to the increase in perceived behavioural control, which drives the intention to behaviour change. Learning elements provided by peer households via the pinboard directly affect the evolution of both perceived behavioural control and subjective norms. Through the challenges, however, the app's focus moves from energy consumption data per se, to what energy is used for (Butler et al., 2018) —namely, routines that are performed in the household. Indeed, challenges question collective conventions about the functions and needs that energy consuming good and services are expected to meet (Royston, Selby, et al., 2018). They allow to hands-on create new competences, with the final aim of re-shaping current social norms and conventions about the specific household routines they refer to.

From this perspective, therefore, challenges and social interactions mediated by the pinboard have high affinity with policy interventions inspired by Social Practice Theories (SPTs). As they act on the dynamics of consumption and attempt to shift collective conventions of normality within the community of SPP households, challenge and pinboard features could be framed as a way to bridge over the incommensurable difference between behavioural and practice approaches (Shove, 2011), which was also proposed and attempted by other authors (Spurling et al., 2013; Kurz et al., 2015; Spangenberg and Lorek, 2019; Hess, Samuel, et al., 2018; Hess, I. Schubert, et al., 2022). From an SPT perspective, in fact, challenge performance and social interaction and learning elements conveyed by the pinboard can be regarded as the enablers of the evolution of competences and of shared meanings and social norms, and they synergistically contribute to re-craft the practices in the direction suggested by the challenges.

Challenges and pinboard features might also be regarded as an attempt to directly address SPT-informed critiques by Strengers (2014) about the tendency to ground the energy transition on smart technologies and on the request to individuals to play a key, active role in such a transition, by becoming "smart" as well. According to the author, the conceptualisation of individuals and households as "interested, immersed, and engaged in managing their demand" (Strengers, 2014, p. 25), which is dominant in the energy sector, cannot stand to the challenges of real-life.

This approach targets "Resource Man" individuals, that are interested in their own energy data, understand it, and are open to changing their energy consumption patterns by responding to the provision of information about them and their impacts. Resource Men are not necessarily males; however, male figures fit well with the image of male-dominated engineering and economics sectors, and therefore these approaches are expected to capture the attention of men more than women. But how many people do really meet this categorisation? According to the author, most of individuals and households simply do not understand energy consumption data, do not know what a kWh is, and are even less capable to guess the implications of a ton of CO_2 saved or emitted. Furthermore, household energy consumption routines need to tightly fit with work or school schedules by household members: by focusing on a rational discourse through the provision of information (energy consumption feedback), the dynamic and social practices within which energy consumption is performed risk being overlooked, and with them the elements that either enable or prevent change.

Strengers (2014) therefore suggests to move the focus on practices that are performed in the home by household members —which by the way are still predominantly women. She invites to consider what is really happening in the household, to explicitly consider how household practices are influenced by constraints coming from outside the household itself (which is in line with the suggestion of "opening the household box" by Raven, Reynolds, et al., 2021), and, more broadly, to re-imagine a different type of everyday life, which has a lower energy footprint, relies on slower times, and is more relaxed.

Furthermore, Strengers (2014) also suggests to overcome the typical individualistic approach of smart energy processes, in which Resource Men operate in isolation from one another —or, in the best case, when they interact with other social entities, they do so by sharing their energy performances on social networks, by comparing themselves with other Resource Men. The risk the author envisions is that interventions targeting Resource Men only appeal to a small group of the population that is already interested in energy data and in the related costs, and that is keen on using new technologies to meet their energy demand. She argues that, with those population groups, interventions aimed at providing information feedback and at leveraging smart technologies will work and will result in decreased energy consumption. The problem is however how to deal with Non-Resource Men people —most likely the majority of the population— that are instead not motivated by engineering-inspired feedback approaches, and would therefore ignore them. For this purpose, she suggests to move the focus on how to promote new ways of living, namely on how to change household practices, inside and outside the home, in order to support the needed decrease in energy demand and CO_2 emissions.

This is exactly what SPP tries to do through the challenges, which focus on practices that are routinely performed in the household, and with the pinboard, which invites households to share their thoughts, impressions, difficulties, or enabling conditions about the process of engaging on new ways of living inspired by the challenges themselves. However, SPP includes such challenges and pinboard exchanges as additional features within a broader process that is still mainly informed by behavioural theories (from Theory of Planned Behaviour to the Transtheoretical model, including gamification and nudges) and focuses on the provision of energy consumption feedback information. Therefore, even though it attempts to go beyond it, SPP does not completely move away from the concept of Resource Man.

6.2 Specific research questions

The RQ4 research question I address through the SPP case aims at identifying which app features can foster higher user engagement and thus provide greater support to the reduction of energy consumption and CO_2 emissions. From both behavioural and SPT theoretical perspectives, I am especially interested in understanding if the interaction possibilities made available by the pinboard, which is the “novelty” of SPP as the majority of persuasive apps only offer individual features, are actually exploited and appreciated by app users, and if, in combination with challenges, they manage to trigger peer-to-peer social learning processes between app users.

Reed et al. (2010) developed an operational definition of social learning, as “a change in understanding that goes beyond the individual, to become situated within wider social units or communities of practice through social interactions between actors within social networks”. Accordingly, for a learning process to be considered “social”, three conditions have to be met. First, a change has to occur in the individual. This might be a rather superficial change, such as the recall of new information, or a deeper change in attitudes or epistemological beliefs. Second, change has to go beyond the individual, and to occur within wider emerging communities of practice. Third, change has to occur through a process between actors within a social network, either via direct interaction between them, such as through conversation or other analogic or digital media, including web-based applications. Besides active interactions, the process of knowledge acquisition between peers can also occur through passive observation of others, as stated by the well-known social learning theory by Bandura and Walters (1977).

Based on an extensive review on peer effects, namely the situations that occur when an individual's attitudes, values, or behaviours are influenced by others who are perceived to be similar, Wolske et al. (2020) have recently theorised when and how peer effects are likely to support change. According to the authors, peer effects can be driven by two types of processes: normative social influence processes, that leverage social norms, mostly in a passive fashion, namely by providing opportunities for observing others' behaviours; and social learning processes between peers, that leverage active inter-personal and persuasive communication channels, for instance through one-to-one conversations or group meetings. In particular, Wolske et al. (2020) have argued that energy consumption behaviours, which they suppose to be more convenient to change, can be effectively influenced by peer observation or social comparison, with little cognitive elaboration or critical analysis (passive social interaction). More complex processes, such as the decision to retrofit a building or to install PV panels, are instead supposed to require dedicated and active communication channels.

Against this background, I tackle RQ4 by first verifying if, within the community of app users, features of the SPP app manage to activate social interaction processes that support energy sufficiency in daily household routines. Then, I consider the user experience and assess the users' level of engagement with each app feature. Finally, I collect insights about households' self-reported energy routines after use of the SPP app, in order to investigate if they change towards energy sufficiency, thus lowering energy consumption and CO_2 emissions. Specifically, I address the three following questions:

- RQ4-1: Do challenges and pinboard features in the SPP app activate social interaction processes supporting energy sufficiency in daily household routines?
- RQ4-2: What SPP intervention techniques (app features) encourage greater engagement in energy sufficient routines by household members?
- RQ4-3: Are SPP users' reported household consumption routines more energy-sufficient after use of SPP?

6.3 Research design

At the time of writing, the SPP app is being tested in a one-year long policy intervention involving voluntary households located in the regions of Schaffhausen, Wil, and Winterthur. Households of the treatment group were recruited in Fall 2021 among the customers of the utility company operating in each region. Open communication campaigns based on posts on the utilities' printed and digital customer newsletters and their social networks were organised; furthermore, customised, personal printed letters were sent by the utilities themselves. No relevant incentives for participation to field test activities were offered and no disguise was used: all recruitment materials transparently hinted at the energy and carbon saving impacts expected by using the SPP app.

Two key eligibility requirements were set: first, households had to be connected to one or more smart meters for the automatic collection of all their energy consumption data, including heating. In the region of Schaffhausen, where smart meters have not been rolled out yet, households had to be willing to install a smart sensor delivering the same information as the smart meter, developed on purpose. Since for the regions of Winterthur and Schaffhausen no gas metering infrastructure was in place, this implied requiring houses to be heated via heat pumps or direct electric heating systems, whose consumption can be measured via electricity smart meters or sensors. The second requirement was that smart meters/sensors had to only measure energy consumption data of the single household —namely, they had to be equipped with a decentralised heating system. As in Switzerland this condition is mostly met by independent house buildings, the latter eligibility criterion was indicated in all recruitment materials.

The recruitment period concluded with 341 project applicants. Since they self-selected themselves based on the communication material and voluntarily applied to join the project, they were expected to be already in the contemplation stage identified by the Transtheoretical model. Namely, they were supposed to be households interested in, or at least open to, saving energy in their daily routines. A few of the applicants had to be rejected since either they did not fully meet the eligibility requirements, or sensor/smart meter connection problems emerged after a technical check by the utility companies, or they were equipped with a photovoltaics power plant and meters/sensors were not able to detect the share of their electricity demand satisfied by the photovoltaic plant itself. Ultimately, 220 households could join the SPP policy intervention (treatment group) and thus potentially enter the SPP community.

The SPP intervention was organised in two phases: for the first three months (phase 1: February, 1 – May, 1 2022), households used the "full version" of the SPP app, which includes the release of new challenges and tips every two weeks and a regional energy saving competition at the end of each month. Then, eight additional months of intervention have been performed (phase 2: May, 2 – December, 31 2022). During such a period, the SPP app was still available to the households of the SPP community, though no new challenges or tips were offered and no regional competitions were held. Households could keep checking their consumption, set their energy saving goal and check its level of achievement, engage in past challenges and benefits of past tips. Also exchanging comments and suggestions via the pinboard was still possible. Every two weeks they received an in-app notification, acting as a reminder about the app's features. Phase 2 thus offered a "light version" of the SPP app and was designed as a "maintenance" phase, aimed at consolidating changes obtained by means of the full set of activities performed during intervention's phase 1.

Within the SPP project, the quantitative energy saving effect of app use is assessed by analysing energy consumption data provided by smart meters, under a quasi experimental evaluation scheme. A control group of similar households, identified ex-post via matching techniques, is considered in order to estimate the effect of app use. As I have done for the enCompass and Social Power cases, energy consumption is measured before and after the intervention in both treatment (app users) and control group, and the average treatment effect on the treated is estimated via panel regression models. Due to constraints on the availability of baseline data and on data for the control group, the estimate of the SPP causal impact on energy consumption is only possible by comparing energy consumptions recorded during the whole calendar years 2021 (baseline) and 2022 (stages 1 and 2 of the SPP intervention). Analyses on such data will thus be performed in Spring 2023 and are not included in this dissertation.

The SPP project also performs a three-wave survey targeting households of the treatment group, which allows to investigate the evaluation of the app's features from the perspective of its users. The questionnaires are administered immediately before the start of the intervention (January 2022), at the end of phase 1 (May 2022), and at the end of phase 2 (January 2023). Furthermore, a rich set of detailed data analytics indicators automatically collected through the app can be used to investigate how household members interacted with the app, while analysis of the pinboard posts allows to investigate how households interacted with each other.

For the analyses I perform in this dissertation, I focus on SPP app users only and on responses to the first two survey waves, coupled with data about the type and frequency of interactions with the app features during intervention's phase 1. I analyse these data-sets with a mixed-methods approach, as summarised in Table 6.2, which specifies how I tackle each specific research question I consider for this case.

Table 6.2: Data sources I use to tackle the specific research questions for SPP.

Research question	Topic of analysis	Data source	Method
RQ4-1: Do challenges and pinboard features in the SPP app activate social interaction processes supporting energy sufficiency in daily household routines?	Number and content of pinboard interactions	Pinboard messages	Quantitative, Qualitative.
RQ4-2: What SPP intervention techniques (app features) encourage greater engagement in energy sufficient routines by household members?	Users' feedback on app's features Evolution over time of app use	Survey after 3-month intervention App's data analytics system	Quantitative, Qualitative. Quantitative.
RQ4-3: Are SPP users' reported household consumption routines more energy-sufficient after use of SPP?	Changes in energy consumption routines	Surveys before and after 3-month intervention	Quantitative, Qualitative.

Regarding RQ4-3, note that, as at this stage of analyses a control group is lacking, any changes in self-reported routines I should find cannot be causally attributed to use of the SPP app. Changes might in fact be due to contextual events affecting the whole SPP community, such as the Russian war in Ukraine that blew up during SPP's field intervention and the related fear for an energy crisis throughout Europe. Estimate of the causal impact of SPP will be performed in Spring 2023 through analyses on the energy consumption data, accounting for a matched control group of comparable households.

The survey questionnaires I analyse in this dissertation were delivered online via the Qualtrics software. Invitation to answer them was sent via email to the responsible household member for the SPP project. At each wave, up to three reminders were sent to favour getting survey responses; to guarantee comparability between the two waves, the same household member was asked to answer the two questionnaires. Before analysis, I checked survey responses in order to remove respondents who had never logged in the Social Power Plus SPP app, as actually they had not received the treatment. Also, I only considered respondents to both survey waves.

Data analytics and pinboard messages were instead automatically made available by the app's data management system. I considered data-analytics and pinboard messages related to the period February, 1 2022 – May, 1 2022 period (phase 1). All pinboard exchanges took place in German, though my analyses are performed on their literal translation in English, automatically obtained via the DeepL Pro tool. For all such data-sets, I compute descriptive and inferential statistics via the R software tool (graphical representations by "ggplot2" and "igraph" packages) and perform qualitative analyses on open-answer questions and pinboard posts with the NVivo 12 software tool.

6.3.1 Social interactions

To verify whether challenges and pinboard features activate social interaction processes, I look for both active and passive interaction processes. For passive social interactions, namely social interactions based on peer observation, I consider the number, evolution over time, and content of pinboard posts. Then, I classify each post based on the topic it refers to and on its purpose, such as for instance reporting an experience, asking a question to other peers, or expressing a commitment for the future. By adopting a qualitative content analysis approach, I perform such a classification inductively, as long as I explore the pinboard topics themselves, without a pre-defined coding list.

To verify if active social interactions occur via peer-to-peer, pinboard-mediated exchanges between members of the SPP community, I then look for presence of direct questions and related answers. Furthermore, I look for presence of direct mentions between members of the SPP community, which, as in many ICT-based chat communication systems, are obtained by mentioning the pseudonym of the target user, followed by the symbol "@". Despite all pinboard messages are publicly visible to all app users, presence of the symbol "@" indicates that novel direct interaction channels have been created by the app, which are worth being formally mapped in a network. I thus explore the number and intensity of connections between each user and its peers by means of network plots, by performing a basic social network analysis which summarises the structure and intensity of active social interactions through a graph.

6.3.2 User engagement

To identify the intervention techniques (i.e., the app features) that encourage households' high levels of engagement on energy sufficient routines, I rely on the survey at the end of phase 1 and on the analysis of in-app actions. Two batteries of survey questions respectively refer to the evaluation of current app's features and the suggestion for novel app features for future app versions, and are therefore well-suited to my analysis.

The other piece of information that I rely on to assess the users' level of engagement with the different features is the app's data analytic system, which allows to track all the relevant in-app activities: registration of new accounts, app openings, setting of energy-saving goals (for both heating and non-heating purposes), starting the challenge, completing a challenge, publishing a post in the pinboard. I analyse the evolution over time of each in-app activity type, by accounting for the number of app users performing each of them and by computing the related descriptive statistics.

6.3.3 Energy consumption routines

To measure presence and extent of changes in energy consumption routines, I perform a comparison between baseline and phase 1 survey responses, by investigating self-reported routines addressed by the challenges. For this purpose, I refer to question items from the Swiss Household Energy Demand Survey SHEDS (Weber et al., 2017), which deal with indoor temperature thermostat setting (as a proxy for energy consumption for heating), number of weekly baths, showers, cycles of washing machine, tumble dryer, and dishwasher, and daily hours of use of digital entertainment/working devices (laptops, tablets, and TVs). I also explore how household members cope with lower thermostat settings, by means of question items inspired to the work by Matschoss et al. (2021).

6.4 Results

Out of the 220 eligible households accepted to join the SPP intervention and therefore to access the treatment with the SPP app, 213 accounts were registered in the app, corresponding to 203 different households (multiple accounts on the app were in fact possible for the same household). Overall, 17 households (7.8% of the accepted applicants) did not comply with the treatment, ad they did not install the app, or at least never logged in. Then, attrition affected the survey waves: we received 199 complete responses to the first questionnaire (baseline) and only 140 complete responses to the second one (after phase 1). By filtering out the responses by households that only answered one of the two surveys and the responses by households that never logged on the Social Power Plus app (though felt the moral obligation to answer the first and in, some cases, even the second questionnaire), I obtain an analytical sample of 112 respondents, which corresponds to

55% of the sample of households that at least once logged in the Social Power Plus app. Their characteristics are reported in Figure 6.4: survey respondents are mostly males (83%), middle-aged (average age is 53 years old) and highly educated (more than half of the participants having a university or PhD degree). Household composition is quite balanced: about half of the households consists of families with kids and the remaining half consists of single adults (more than half of which are older than 65 years).

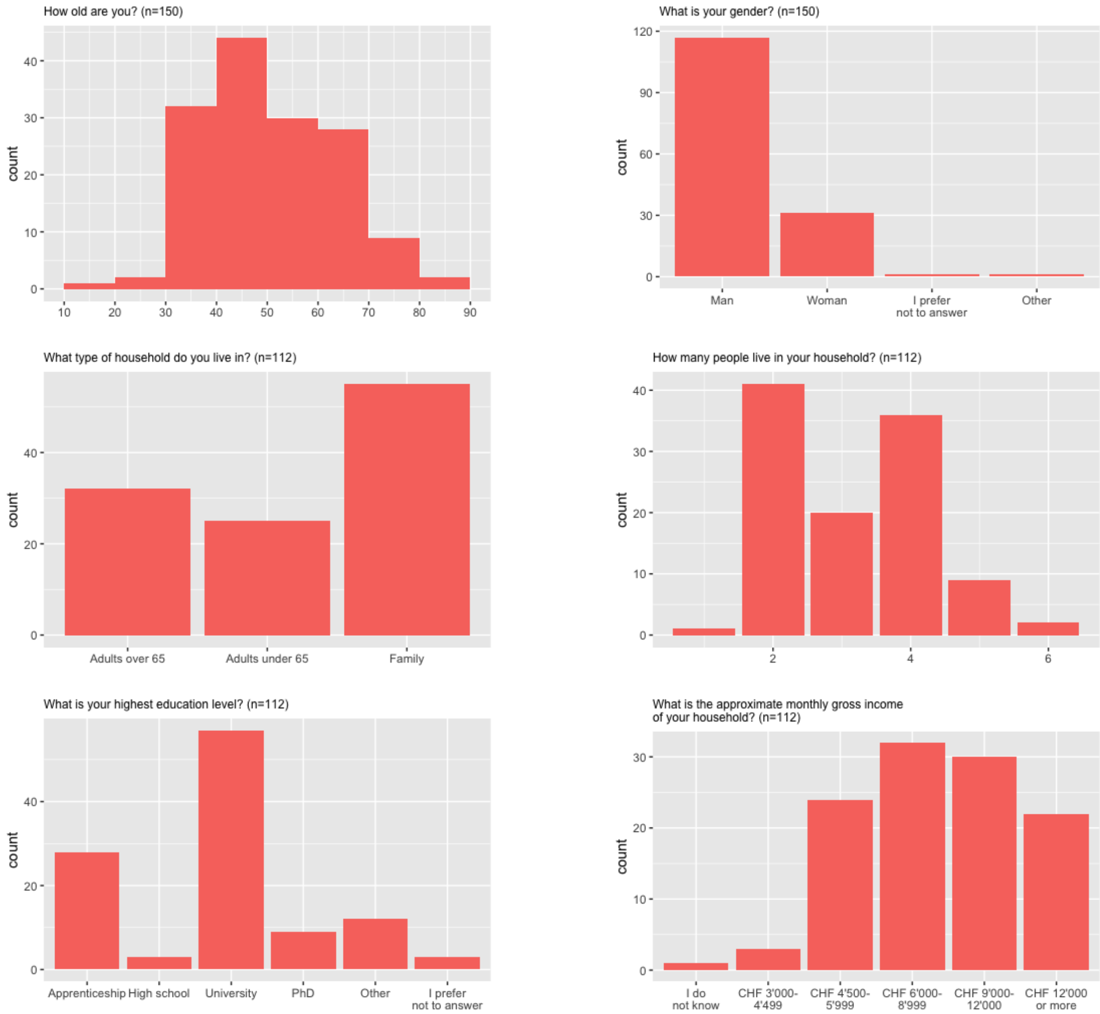

Figure 6.4: Characteristics of respondents to both surveys (frequencies, n = 112).

6.4.1 Social interactions

In total, N= 257 messages were posted on the pinboard: n= 193 were by members of the SPP community and n=64 were by three app administrators of our research team, included myself. All such messages were always visible to all SPP community members, thus at least enabling "passive" social interactions (the observation of others) for those who did not actively engage in posting messages. By considering the topic of the posts by app users, and excluding posts related with technical problems experienced with app use, the most frequent post category deals with heating-related routines (Figure 6.5).

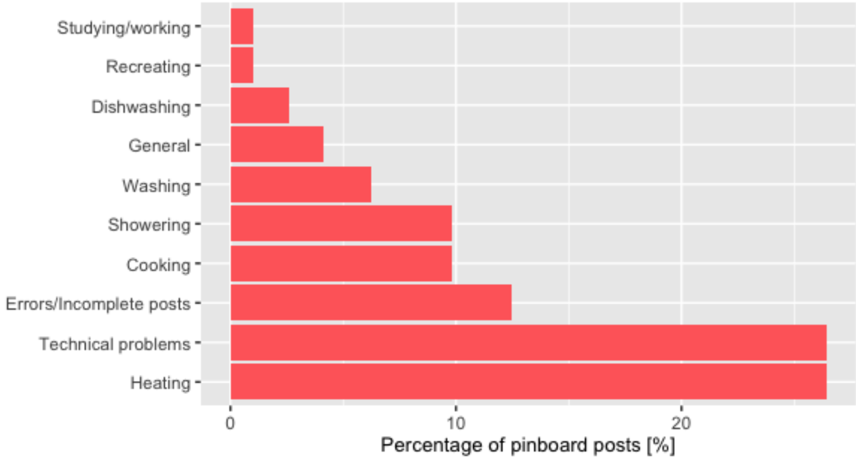

Figure 6.5: Classification of pinboard posts by members of the SPP community (n_{users} = 193).

Four factors might have driven a larger interest for heating-related routines, compared with those tackled by other challenges: challenges about heating were launched at the very beginning of the field intervention, when app use was higher by most households. Further, SPP displays energy consumption data for heating/hot water purposes and for other purposes separately, thus giving relevance to the heating topic. Also, the challenge description into the app indicated that reducing indoor temperature has a very high effect in energy saving. And finally, members of the SPP community consist solely of households living in independent houses, that are directly responsible for use of their heating system, and therefore possibly highly interested in how to manage it at best.

To understand the types of exchanges that happened in the pinboard, the posts were inductively categorised according to both their content and the type of message they expressed. Many posts reported issues households experienced in energy data-transmission or questioned the reliability of the energy consumption break-down in heating and non-heating purposes —and were therefore accompanied by the responses by app administrators. Overall, 78 posts out of 193 (about 40% of them) met the goals the pinboard was developed for, namely sharing positive or negative experiences with peers of the SPP community and laying the ground for an evolution of the related social norms and conventions. A selection of such posts aimed at exemplifying their content is reported in Table 6.3. Most of such posts refer to actions that could immediately be implemented by their peer households, since they do not require investments in technical infrastructure or equipment. A limited sub-set of posts, instead, adopts highly technical, expert-based perspectives, and refers to measures that require investments in energy efficiency of the building components or its heating system.

To assess if the scope and coverage of such interactions were sufficiently wide and rich to trigger an active social learning process, I verify if, besides reporting a few of their own experiences and making them available for the observation by peer households, members of the SPP community also actively interact with each other. For this purpose, I analyse the pinboard posts, looking for presence of mentions ("@" symbol) and questions/answers

Table 6.3: A selection of pinboard posts dealing with SPP community members' experiences.

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, 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	—	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.
Dish-washing	We have connected our dishwasher to the hot water system because we do 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 have a very small lawn, 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.

between members of the SPP community. Presence of either mentions or questions that are answered in fact hints at an active dialogue between community members.

Overall, 53 pinboard posts include either a direct mention to another member of the SPP community or a question or an answer. Of them, 24 are posted by app users, the rest (29) by the administrators. The graph of active user interactions mediated by the pinboard is represented in Figure 6.5. Nodes represent members of the SPP community that either sent/received a mention or interacted with questions or answers (administrators are represented in red) and edges indicate presence of a mention/question/answer between the couple of nodes they connect. Larger width edges imply presence of more exchanges between the related nodes.

Besides a large number of active messages from the three administrators to other members of the SPP community, which give rise to a "star network topology", Figure 6.6 shows there are also a few interactions that bilaterally involve couples of SPP community members, without passing from the administrators. Namely, in some cases the pinboard appears to have originated the peer-to-peer "mesh-type" active interactions that it was designed for, in order to activate a social learning process between peer app users. However, the graph also shows that the number of households involved in such a network is limited: only 35 different households are involved in active social interactions via mentions, questions or answers (including messages to/from the app administrators), which is about 17% of the 203 registered households of the SPP community. Also, the frequency of such interactions is limited: on average, every app user made 0.22 mention/question/answer posts and was mentioned/questioned/answered by 0.20 posts. The users that at least once mentioned/questioned/answered other households, on average made 2.7 mentions to other users, and those that at least once received a mention/question/answer, on average received 1.85 mentions.

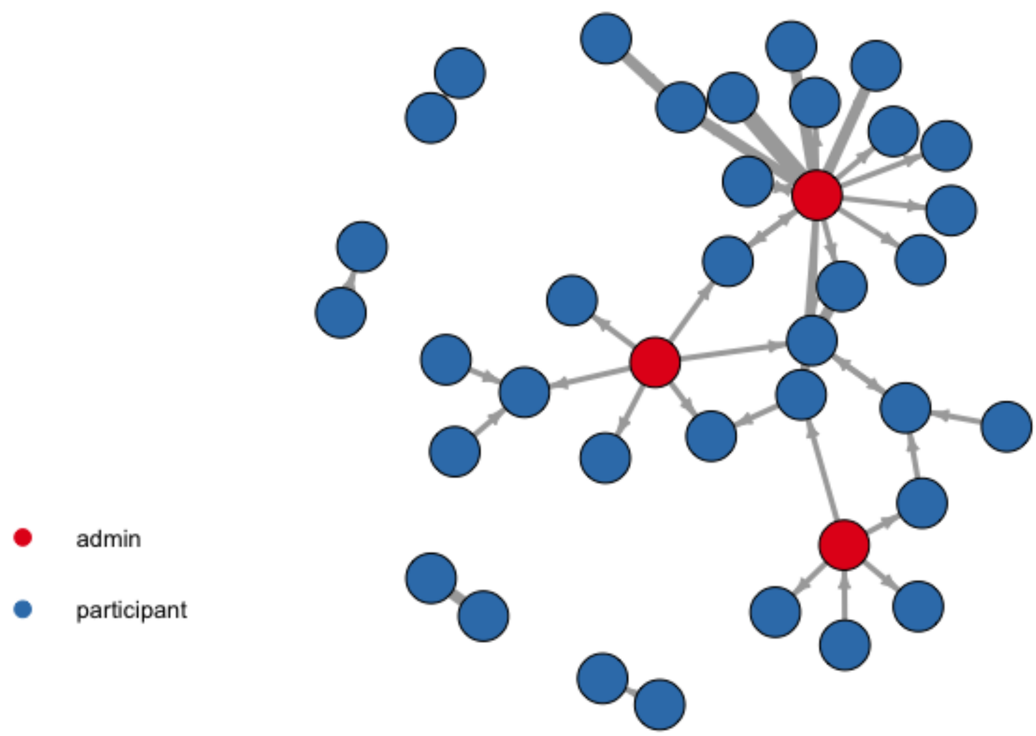

Figure 6.6: Pinboard-mediated active interactions between members of the SPP community.

To get additional insights on pinboard-mediated active interactions between members of the SPP community, I analyse in details the content of question and answer messages. Among all pinboard posts, seven questions explicitly targeted peer members of the community, and six of them received a direct answer by other community members via the pinboard (Table 6.4 reports them integrally).

These interactions, which correspond to a total of about 8% of the pinboard posts, involve 11 of the 203 households of the SPP community, namely about 5% of them. As shown by Table 6.4, which reports them integrally, all but one questions deal with the heating topic, and focus on smart home systems and technical settings of heater or heat pump and smart home systems. The question not dealing with heating asks for support in understanding atypical consumption peaks observed in the app's energy consumption feedback plots. To properly answer these questions, knowledge and competences by average households are definitely not sufficient, and heating system developers of installers are needed. Nevertheless, households with the needed competences were part of the SPP community and were willing to share their knowledge and provide tentative answers.

Elements collected by analysis of the pinboard messages themselves are not sufficient to provide a definitive answer to my research question RQ4-1. On the one hand, in fact, the pinboard seems to have achieved its goal to create a protected space for households to actively interact with their peers about energy-related topics and at least allow for passive observation of peer's energy practices. Namely, it has enabled novel interactions that, in its absence, would not have occurred. From Social Practice Theory perspectives, to some extent the pinboard allowed for alternative discourses, discussions about collective conventions, as well as novel competences about household energy routines to emerge.

On the other hand, however, the small sample of actively engaged members of the SPP community and the low amount of questions, answers, and mentions exchanged between peer users seem to suggest that the active social learning process I expected to find, triggered by active social interactions between members of the SPP community, has not occurred for a large part of users. Based on the collected material, I thus tend to exclude that the pinboard has triggered an active social learning process. The collected elements, however, do not allow me to exclude that a passive social learning process has started to take place, mediated by observation of peers' experiences through the pinboard. No information is in fact available about the number of times each pinboard post was read or about the number of peers that read each pinboard message. To get insights on passive social learning processes, I thus turn to research question RQ4-2, aimed at assessing the level of engagement by members of the SPP community with the different app features.

6.4.2 User engagement

The interactions by each member of the SPP community with key app features are automatically computed by the in-app monitoring system. The evolution of in-app actions over time by all app users is reported in Figure 6.7. Also, Figure 6.8 represents

Table 6.4: Pinboard questions and answers provided by peer users (pseudonyms 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)
@ 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)	@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)
@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)	@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)
@BH: What kind of models [of programmable thermostatic valves, that user BH described in a post before, editor's note] 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)	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)
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)	*[Not answered by the community]*
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)	From an energy-saving point of view, this is worthwhile from the first day onwards. It depends on many factors, 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. 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)
Consumption No heating. We are absent, boiler off via fuse and RCCB [Residual Current Circuit Breaker, a safety measure for electrical circuits, editor's note] I am now completely amazed to observe that there is regularly a consumption peak of 2 to 4 kw/h 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)	@AH what are your large consumers that could draw over 2kW? Boiler? HP [heat pump, editor's note]? 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)

the distribution of in-app actions between users and Table 6.5 reports descriptive statistics on the number of in-app actions per user, over the three-month intervention period.

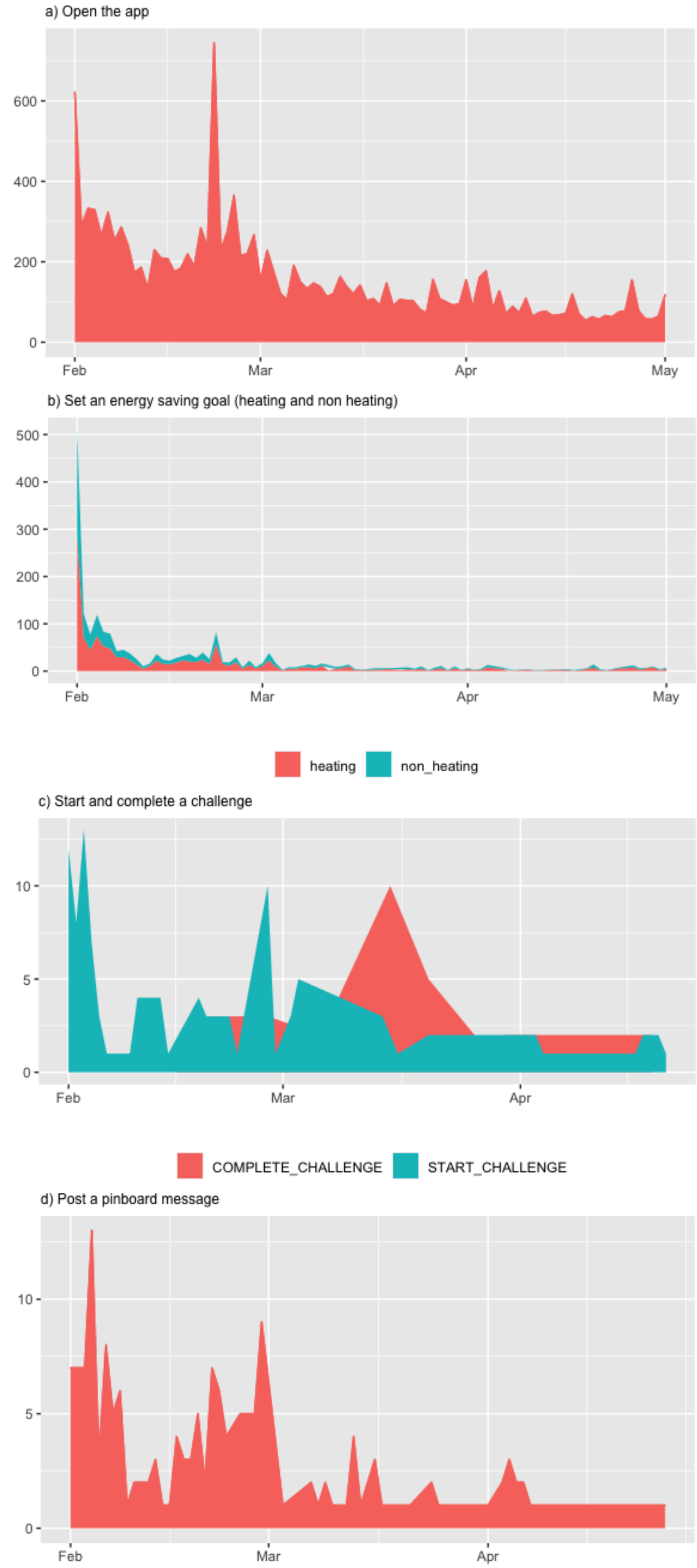

Figure 6.7: Evolution over time of the number of key in-app actions, considering all app users.

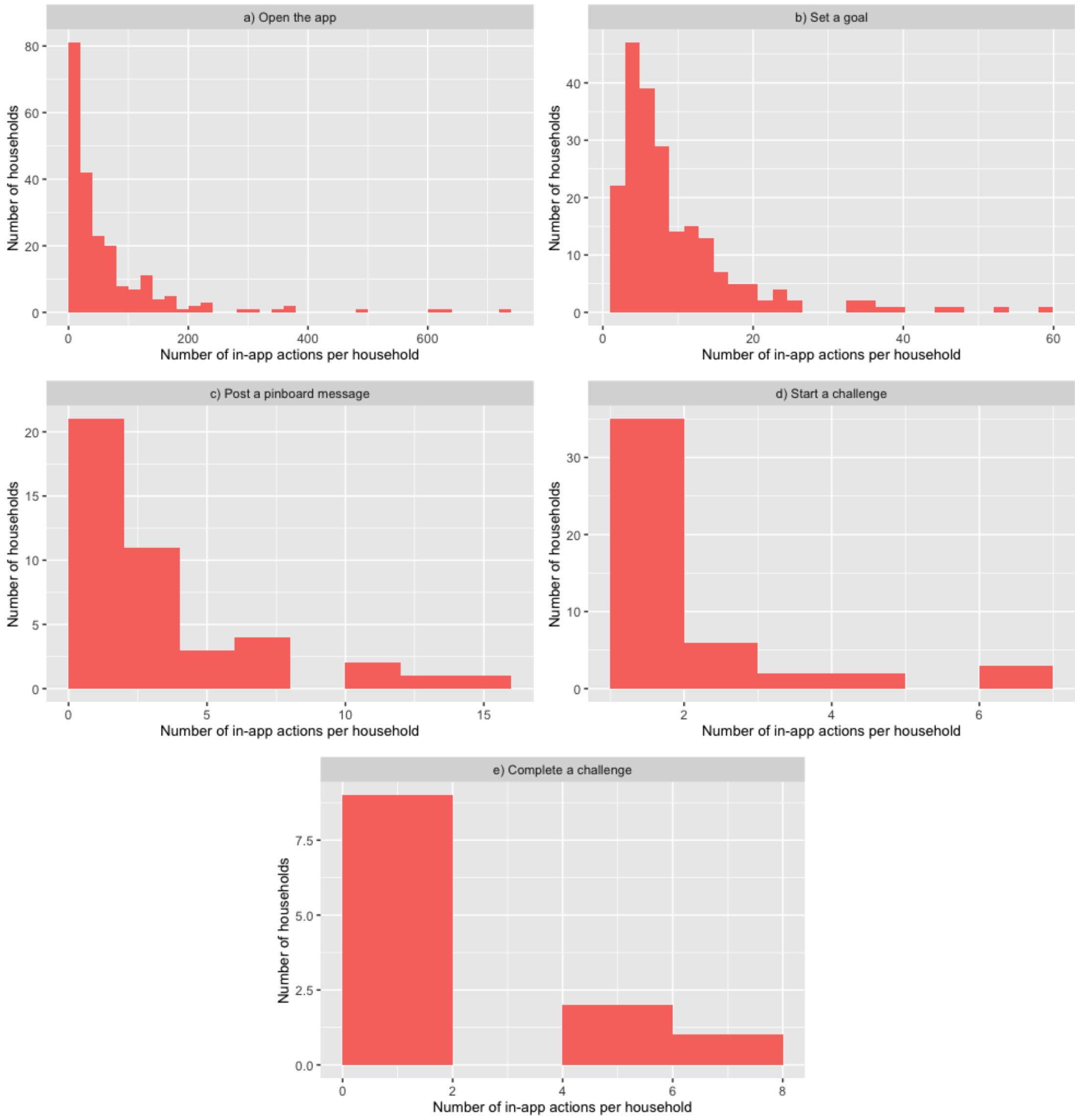

Figure 6.8: Distribution of key in-app actions by app users.

Though the number of app openings remained quite high over time, with about 80 openings per day even in the last month of the intervention, the plots suggest that the level of engagement with SPP's specific features was limited for most of the SPP community members. This is confirmed by the descriptive statistics reported in Table 6.5, which also shows the number of users that performed each action. Individual goal setting was by far the most frequent action performed. App users were in fact requested to set a goal for non-heating and one for heating energy consumption at the start of app use; then, they were free to modify the goals whenever they wanted. Data show that goals were indeed changed quite often in the first month of app use (on average 7.1 additional times) and that changes in the heating goal were more frequent than changes in non-heating goal (Figure 6.7). In contrast, the percentage of households that engaged in challenges or posted messages in the pinboard was definitely smaller, respectively equal to 21% and 25% of the SPP community. Furthermore, only 6% of households formally completed a challenge by briefly commenting on their experience in the pinboard. Nevertheless, among those that performed such in-app actions, the number of interactions was still satisfactory: households that posted at least one pinboard message, on average posted 3.8 messages each; households that at least started one challenge, on average started 2.4 challenges each, and households that at least completed one challenge, on average completed 2.5 challenges each.

Table 6.5: Descriptive statistics on the number of in-app actions per household.

		Open the app	Set a goal	Post a pinboard message	Start a challenge	Complete a challenge
Households that performed the action at least once	Num	203	203	42	48	12
	%	100.0	100.0	20.7	23.7	5.9
Actions per household (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

The above piece of information already provides insights for research question RQ4-2, aimed at identifying which SPP intervention techniques (and thus app features) encourage greater engagement by household members. To further tackle RQ4-2, I also refer to the app users' direct feedback on the app features, collected via the survey performed at the end of the three-month intervention (Figure 6.9, n = 112).

In line with the above figures, survey respondents indicate appreciation for individual-level features (goal setting and tips). Instead, the features based on social interaction (pinboard, in-person meetings) or leveraging social-level determinants (regional energy saving competition) are rated as the least useful in motivating to save energy, in keeping the level of engagement high, and in learning something new about energy-saving. Specifically regarding pinboard posts, 74% of respondents declare they did not like to post messages in the pinboard and 67% of them declare interaction with other members of the SPP community did not motivate them to save energy.

These responses hint at the fact that app users did not feel engaged into active social interaction process. Responses also indicate that passive social interaction processes have only occurred in a limited number of cases: 36% of the respondents declare they liked to read about their peers' experiences in the pinboard, while 45% of them declare they did not. Combined with the 19% of respondents indicating a neutral position, this implies that 64% of the members of the SPP community were marginally involved by the pinboard, even in a passive way. The majority of respondents also indicates that challenges neither inspired them to experiment with new ways to save energy nor motivated them to energy saving, and that the breakdown into heating and non-heating energy consumption was not sufficiently helpful for them to save energy: when they had applied to SPP, they expected to get more detailed energy consumption feedback, including the break-down at the appliance level, possibly also in real-time.

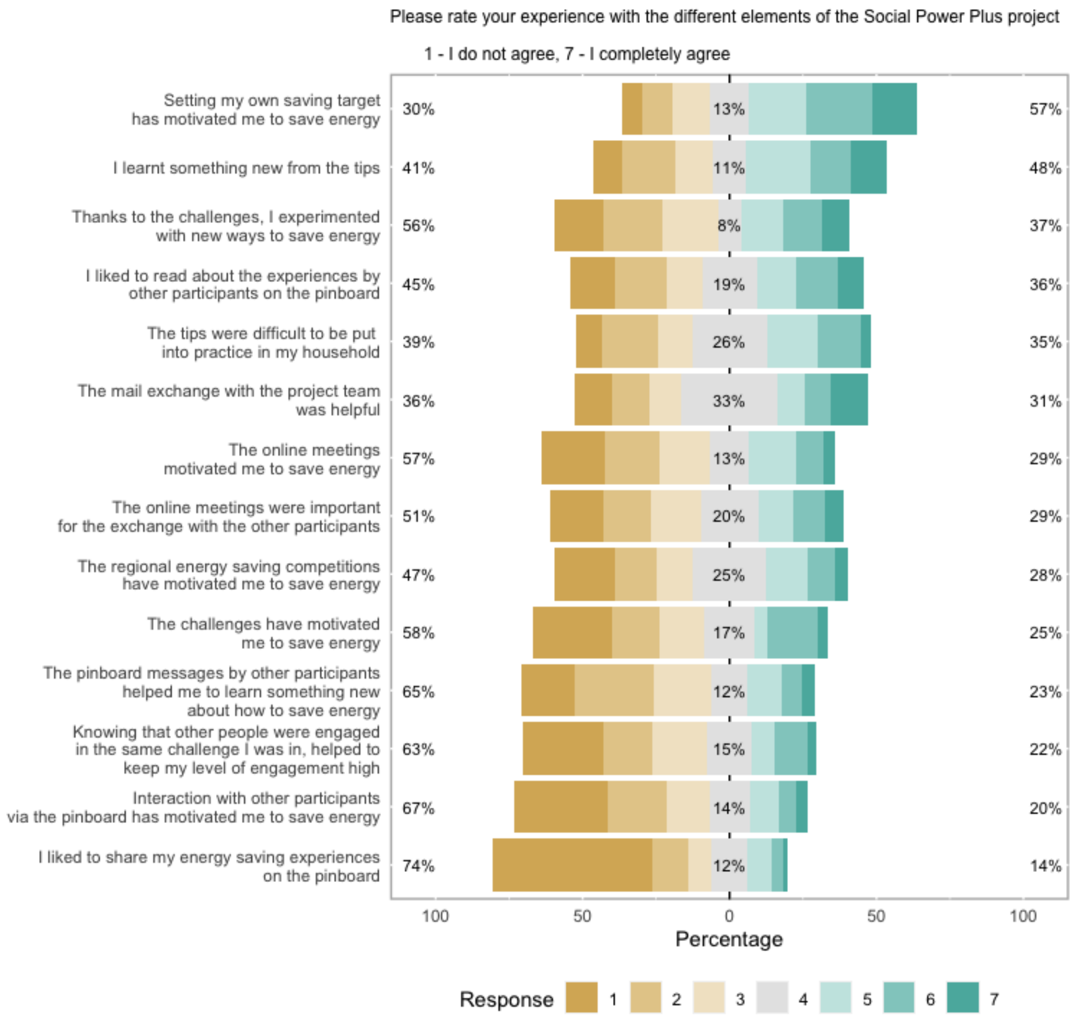

Figure 6.9: Feedback on SPP's app features according to responses to the survey after the intervention (n = 112).

Open-text comments provided by survey respondents about general feedback on the app's features are in fact very clear and explicit in confirming that their interest resides mostly on individual energy consumption features, particularly on the provision of detailed energy consumption feedback, which they would then use for further individual analysis and optimisation. In the open fields available for general app feedback, 44 households (nearly 40% of the 112 respondents), complained they would have expected 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 rough and you have no idea how the use of individual devices affects the energy consumption"*;
- 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"*;

- 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 with the best possible temporal resolution"*.

Such expectations seem to be grounded in the faith in the role of technology and energy efficiency measures, more than in the belief that energy sufficiency measures and a global reduction in energy demand are needed. This attitude is explicitly indicated by a few open feedback comments, by SPP members that have already implemented energy efficiency interventions in their houses, 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 considered 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"*;
- *"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 test winner economy shower head."*

Finally, features aimed at fostering interactions and social learning processes are very seldom mentioned in the open questionnaire comments. Though two households 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"*), to the opposite one household stated the explicit lack of interest for reading about suggestions by their peers (*"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"*).

6.4.3 Energy consumption routines

Overall, the above results are in contrast with my initial expectations about the activation of social learning processes among peer members of the SPP community and their level of engagement with the different SPP features. However, I still tackle research question RQ4-3, aimed at analysing whether, after use of the Social Power Plus app, household members report changes in energy consumption routines.

Answering to a general question investigating whether survey respondents had tried new energy saving activities at home, 52% of them answered affirmatively. Indeed, collected data comparing responses before and after app use ("baseline" vs "after phase 1" questionnaires), reported in Table 6.6, show a decrease in both self-reported temperature settings and weekly/daily frequency of all energy consuming household routines.

Table 6.6: Energy consumption self-reported routines before and after use of SPP app.

	Baseline			After phase 1			Effect size	t	p value
	n	Mean	SD	n	Mean	SD	Cohen's d	statistics	
At what average temperature (°C) do you heat your living room during the day?	110	21.06	0.87	107	20.57	0.92	0.56	5.7991	6.97E-08 ***
On average, how many showers do you take per week?	111	5.73	2.60	111	5.63	2.91	0.04	0.4285	0.6691
On average, how many baths do you take per week?	110	0.52	1.20	110	0.38	0.74	0.13	1.3887	0.1845
How many times per week does your household use the oven?	112	4.01	2.08	112	3.46	1.64	0.29	3.0999	0.0024 ***
How many times per week does your household use the dishwasher?	110	5.23	2.76	111	4.81	2.62	0.25	2.6360	0.0096 ***
How many times per week does your household use the washing machine?	112	4.42	2.48	112	4.17	2.57	0.15	1.5909	0.1145
How many times per week does your household use the tumble dryer?	102	2.56	2.42	99	2.28	2.27	0.23	2.2776	0.0250 **
On average, how many hours per day are TVs running in your home?	106	3.00	2.17	105	2.50	1.94	0.26	2.6988	0.0081 ***
On average, how many hours per day are computers running in your home?	111	6.81	6.20	108	5.72	6.25	0.25	2.5956	0.0108 **
On average, how many hours per day are tablets running in your home?	88	3.65	4.78	91	3.07	5.18	0.09	0.8227	0.4131

Statistical significance: 0.1 *, 0.05 **, 0.01 ***

I tested whether such decreases are statistically significantly different, against the null hypothesis H_0 that they are not. According to paired t-tests, temperature settings and frequency of most of the routines results to be significantly lower after use of the SPP app (at the 0.05 or 0.01 significance level, depending on the routine), which suggests that a reduction in energy demand for household routines has actually taken place, towards higher energy sufficiency. The decrease in the average value of thermostat temperature setting is also characterised by the highest effect size, measured via Cohen's d parameter (0.56, intermediate effect size). Interestingly, such a decrease is observed for the routine that was discussed the most in pinboard posts, namely "heating" (setting thermostat value). This relevant change in routines might be related with external factors (such as the fear of an energy crisis due to the Ukraine war) or passive social interaction processes among peer app users. I suppose however that this is also related with the same reasons I put forward to explain the high amount of posts on the topic of heating: it was tackled in the first challenges, the effect of indoor temperature reduction on energy saving was reported to be very high in the challenge description, the app offers breakdown feedback on heating energy consumption, and households living in individual houses can directly and easily influence their heating consumption by modifying their thermostat setting.

The observed reduction in temperature settings is worth a specific comment. A change in energy demand for heating in fact is strictly inter-twined with the temperature level that is considered to offer sufficient comfort —and, more generally, on trade-offs between comfort and money spent for energy supply (Sovacool, Osborn, et al., 2020). The statistical significance and intermediate effect size of the change in thermostat setting is partially in contrast with findings by a recent analysis by Sovacool, Demski, et al. (2021), who argue that policies aimed at encouraging households to reduce the thermostat setting temperature may have limited effectiveness, since individuals expect to have comfortable, warm homes —indeed, householders are not interested in heating *per se*, but in the comfort heating provides. The conclusion by Sovacool, Demski, et al. is that heating decarbonisation can more easily be obtained by providing individuals with the needed comfort level (via energy efficiency solutions), rather than by relying on individuals to voluntary reduce their heating demand (via energy sufficiency solutions). The above data on household routines, instead, show that sufficiency changes in heating demand are actually possible. Future research based on electricity consumption data will however estimate the quantitative outcomes of the SPP intervention in terms of the amount of saved energy and of the related effect size (quasi-experimental controlled intervention), and then compare such outcomes with the findings by Sovacool, Demski, et al. (2021). In fact, despite the self-reported change in household temperature setting, I cannot exclude that the corresponding energy saving effect is negligible, which would ultimately support findings by Sovacool, Demski, et al. Furthermore, changes in heating routines might be related with the Russian war in Ukraine. If so, by considering the change in energy consumption by both the treatment and the control group, quasi-experimental analyses would allow the phenomenon to emerge.

Observed decreases in the number of weekly showers, baths, and washing machine cycles, as well as the daily use of tablets, are instead not significant, even at the 0.1 level. Even though a decrease in frequency is not the only way to reduce energy demand, which could also be obtained by shorter or lower temperature showers and baths or lower temperature cycles, these results may hint at the difficulty of shifting cultural conventions on cleanliness-related topics, which are at the core of many works inspired by Social Practice Theories (Shove, 2003; Hand et al., 2005), also within real-life interventions in living lab frameworks (Godin et al., 2020). The lack of (or very limited) discussion on social conventions via the pinboard might have contributed to hinder such a shift. This suggests that different strategies should be enacted to favour more engagement by householders in questioning and challenging current norms and conventions.

Considering current building structures and technologies, the largest amounts of energy savings that households can obtain are related with heating demand (Kemmler and Spillmann, 2021). From the behavioural perspective, the highest savings stem from lower setting of temperature thermostats, therefore they are worth further investigation.

To investigate if and how households changed their routines when they set lower temperatures, in the questionnaire at the end of phase 1 we also asked the additional set of questions reported in Figure 6.10, inspired to the work by Matschoss et al. (2021).

Households were requested to indicate if, since they turned heating down, they performed much less, more often, or with the same frequency as before, a set of actions aimed at keeping themselves warm. Household members 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 up the temperature again. However, they declared a decrease in use of warmer clothes, which would be an easy solution to keep bodies warm in less heated rooms, and in fact emerged in surveys by Matschoss et al. (2021) as the most used way to keep warm under lower indoor temperature settings. The fact that, as above indicated, most of the members of the SPP community believe in technology and energy-efficiency approaches, more than energy sufficiency ones, might explain why they seem to have experimented only with a small set of practices aimed at keeping their bodies warm: they were not experienced with such practices and the weak social pressure exerted by SPP challenges and by the peer group of the SPP community was not sufficient to stimulate them to create novel competences.

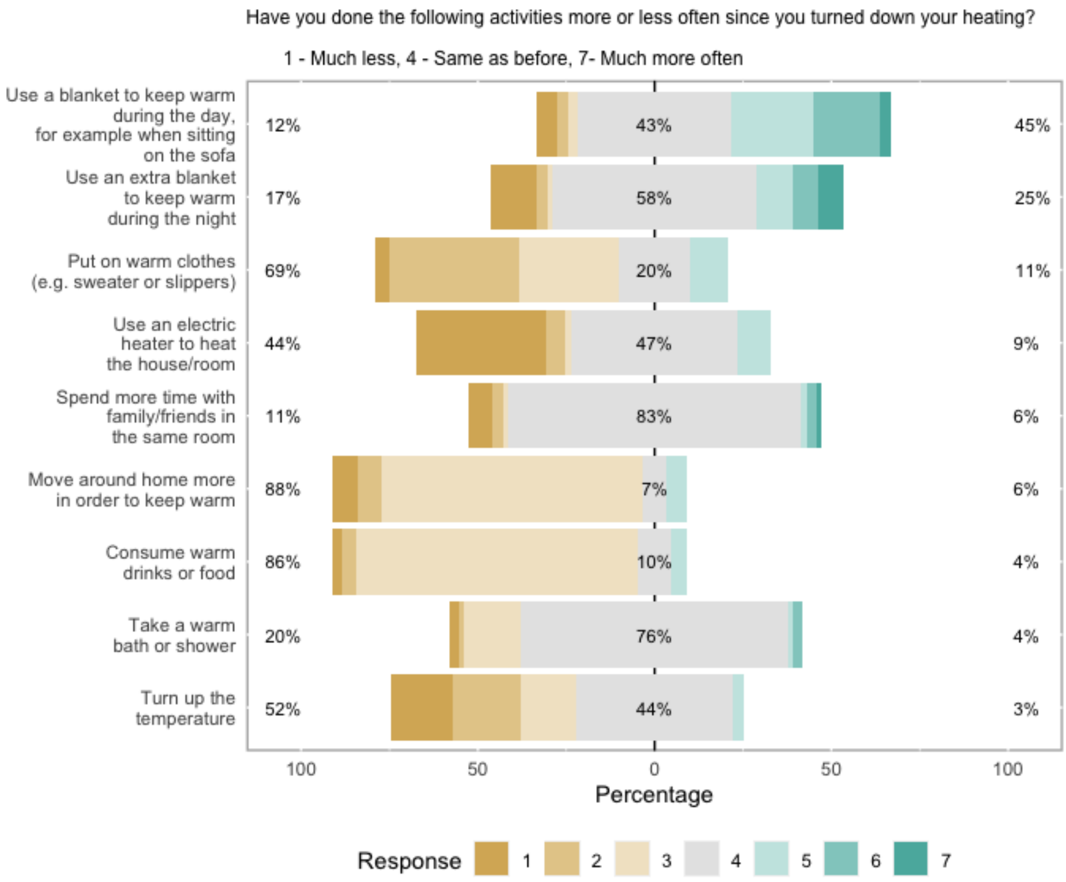

Figure 6.10: Evolution of household practices when indoor heating is turned down, according to survey responses by households that declared to have turned heating down (n=74).

6.4.4 The role and potential of the Social Power Plus app

Analysis of the open-ended survey feedback about household routines provides additional insights about the behaviour change process and the role and potential by the SPP app. Specifically, households' comments suggest that members of the SPP community actually differ than the target households I expected to engage, and for which the SPP's app features were designed. With regard to the Transtheoretical model of behaviour change, it seems in fact that most of SPP households are not simply contemplating change towards energy saving: they are already highly experienced with change, having in some cases already invested in technical energy efficiency interventions on their home's building envelope and on the related energy systems, and being mostly familiar with the amounts of energy consumed by their house, thanks to self-monitoring processes they had already activated themselves well before use of SPP. This is clearly reported by open comments in the surveys such as the following ones:

- *"If someone has no idea how to save, the app is ok. But if you have optimised 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..."*;
- *"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"*.

Namely, SPP app users seem to exactly coincide with the "Resource Man" conceptualised by Strengers (2014), from which with the SPP research team we tried to distance ourselves in the design of the app features. On the one hand, this might be a consequence of the eligibility requirements imposed by current development of metering infrastructure: to join the SPP intervention, households were required to live in independent houses, and in the regions of Winterthur and Schaffhausen they were also required to be equipped with heat pumps or direct electric heating systems. The presence of a heat pump, in particular in non-recent buildings, is not the common standard in the building sector yet. Being equipped with a heat pump is thus the outcome of a specific choice by the household, which might be a predictor of higher pro-environmental attitudes than average. Furthermore, following the Theory of Planned Behaviour, such attitudes would in turn drive intention to perform more energy efficient (and possibly also energy sufficient) behaviours than average households, already before joining SPP. This would also explain the high interest by SPP community members in detailed and highly broken-down energy consumption feedback and in customised recommendations to further optimise their energy saving results.

On the other hand, presence of a majority of "Resource Men" in the SPP community might also be the direct outcome of the self-selection process behind app use: apps like SPP might not manage to raise the interest by its actual user target group, made of households in the contemplation stage. Rather, they might mostly stimulate technically interested individuals, that are already experienced with energy efficiency at home. I cannot even exclude, in particular, that the very fact that the SPP intervention was based on an app, automatically restricted the potential audience of interested households and individuals to those belonging to the Resource Man category. Especially, by considering the way previous apps were designed, the fact that they were grounded in engineering and optimisation approaches, and the types of feedback features they were endowed with, the fact that during household recruitment the SPP team specifically advertised access to the SPP app as an opportunity to save energy, might have raised the interest by those who expected to find detailed energy consumption data and customised suggestions on how to further optimise their consumption amounts, by means of additional technology-based and energy-efficient interventions to the physical and material components of their house. Namely, I cannot exclude that app-based interventions cannot but be accompanied by biases towards the engagement of "Resource Men" instead of households that had no previous engagement with their energy consumption data and with the related optimisation through technology-based, energy efficiency measures. If so, bias towards engagement of a particular category of households would be intrinsically related with use of persuasive apps to support the energy and climate transition, and would unavoidably appear "by design", when apps are leveraged to support change —unless their clear difference with respect to previous apps is crystal-clear indicated during initial communication activities aimed at recruiting app users.

6.5 Conclusions

The results I found indicate that some changes occurred in household energy routines after use of the SPP app. I am well aware that self-reported behaviours might be subject to social desirability bias and that the lack of a control group implies the self-reported change in routines cannot be causally attributed to use of the SPP app. It might in fact be due to external, contextual events affecting the whole SPP community, such as for instance the fear for an energy crisis induced by the Russian war in Ukraine that started during SPP's field intervention. Analyses on the automatically measured energy consumption data, accounting for the control group as well, will provide insights on these aspects and clarify if and to what extent use of the SPP app contributes to energy and carbon saving (and thus to changes in routines). For the time being, based on the available material, I cannot exclude that at least some part of the reported change is due to use of the SPP app.

Furthermore, the collected material provides me with sufficient material to answer my research question RQ4 about which app features can foster higher user engagement and thus provide greater support to the reduction in energy consumptions and CO_2 emissions. The resulting insights seem to suggest that, differently from my expectations, the possible change induced by the SPP app is marginally related with both active and passive social interaction processes enabled and mediated by the SPP pinboard. Survey responses in fact indicate that household members mostly appreciated SPP's individual features, suggesting that individual feedback on energy consumption, together with individual goal setting options, have played a major role in supporting the evolution of household routines. Most of the members of the SPP community declared their eagerness for even more detailed energy consumption feedback, under the expectation that it allows them to further optimise their consumption and guarantee personal well-being with limited routine change, under a mostly technology-based energy efficiency framework. The majority of app users also declared they did not like to either read about their peers' experiences or to write about their own one's in the pinboard. These responses thus strengthen the conclusion that the self-reported changes in routines are marginally related even with passive pinboard observation. Instead, they might mostly be related with individual app features providing feedback on consumption and goal setting opportunities.

Overall, the collected elements suggest that social interactions enabled and mediated by the SPP app did not manage to activate a social learning process on a broader scale. These results may definitely be influenced by two key factors. On the one hand, SPP social interaction features were mediated by the app and were expected to occur between strangers that had no previous connections between each other. The lack of previous relations between members of the SPP community may explain the limited interest in the pinboard features. Future interventions might therefore preferably address households that already have personal connections —possibly also leveraging such connections to increase the number of app users. Note, however, that such a strategy would not always be feasible. In the case of Switzerland, for instance, smart meters have not been widely rolled-out and the number of households already equipped with them is still limited. Further, if the app aims at addressing energy consumption for heating purposes, households living in building blocks with centralised heating systems would not be eligible to app use, as their specific energy consumption for heating purposes would not be measurable by the smart meter.

On the other hand, the specific characteristics of the app users, who self-selected themselves based on the communication material on the project, where individual energy consumption feedback was a very prominent incentive, may have influenced their expectations regarding the SPP app and the way they interacted with it. Elements from pinboard posts and survey open responses in particular suggest that members of the SPP community differ from the target households we expected to engage, and for which the SPP's app features were designed. With regard to the Transtheoretical model of behaviour change, it seems that most of the members of the SPP community were well-beyond the contemplation stage for which the app was designed, and were already in the action —if not maintenance— stage. Also, most of them seem to be characterised by higher pro-environmental awareness than average households. This is partially due to the current limited development of the smart metering infrastructure and the need for fulfilling eligibility requirements (living in single households equipped with heat pumps) and partially to pure self-selection reasons. And finally, most of them appear to have a strong faith in technology, which drives them to mostly think in terms of energy efficiency and detailed monitoring of their performances, rather than be interested in sharing their experiences with other peers about energy sufficiency measures.

Overall, the collected data suggest high affinity between the majority of members of the SPP community and the concept of "Resource Man" by Strengers (2014), which is indeed quite far from the target group of app users we had in mind when we co-designed the app features. In fact, with the SPP team we had specifically introduced the challenge and pinboard features exactly with the aim of moving away from Resource Men only interested in accessing their household's energy consumption data and optimising them, mostly through energy efficiency approaches.

With a different group of members of the SPP community, would my conclusions about social interactions have been different? Further research might address this question. A small group of participants in fact actively interacted with the pinboard and the challenges and shared experiences and knowledge around energy saving at home. The community-based app features implemented through the SPP challenges and the pinboard might therefore have a potential for a specific target group as those SPP users: further research might aim at profiling them and at estimating the average treatment effect on them.

However, under the current status of the metering infrastructure, the SPP experience tells about the difficulty of engaging with such actual target group of users and brings me to question the overall potential and usefulness of the SPP app. It seems in fact it has limited potential to change routines (and thus energy consumptions and CO_2 emissions) of most of those who are voluntary and spontaneously interested to use it: they have already implemented most technical changes and made their house and appliances more energy-efficient, which leaves less room for reducing their energy consumption and emissions by changes in daily practices informed by energy sufficiency frameworks.

Indeed, insights from the SPP experience suggest that, based on the way they are designed, advertised, and generally received by the population, persuasive apps tend to mostly attract "Resource Men" households and not to manage to raise the interest by its actual target users, who would probably benefit the most from app use. These learnings from the SPP case therefore suggest that further research and reflections are needed to understand if and how participation of the actual target group can be ensured within app-based interventions. Ultimately, it leads me to question whether persuasive apps are really worth the effort and are beneficial to energy and climate transition processes.

7 Discussion

> *Every time man makes a new experiment he always learns more. He cannot learn less.*
>
> — **R. Buckminster Fuller**
> Architect, system theorist and designer

In this chapter I perform a general discussion of the results I obtained by the analysis of the three cases, with respect to the four overarching questions that have driven my research. I cross-compare their outcomes between each other and further compare them with interventions reported in the scientific literature. I also reflect on the limitations of the persuasive apps I have analysed regarding their upscaling potential, their capability to engage a large share of the population that has limited interest in energy and climate topics, and their overall cost-benefit effectiveness. I close the chapter with a reflection on the role of the context and its potential impact on energy and carbon saving policies.

7.1 Cross-comparison of the cases

The three cases I have analysed aim at supporting the energy and climate transition in households. They leverage an app to provide energy consumption data feedback, coupled with a number of persuasive, gamified features that, as decades of theoretical and applied research have shown, can potentially increase feedback's behaviour change effectiveness. All the three app cases were designed by the respective research teams before my work for this dissertation and I was only involved in the co-design process for the design of the features of the Social Power Plus app. From the theoretical perspective, all three app cases are grounded in the Theory of Planned Behaviour, namely they assume that a given behaviour is intentionally performed, driven by individual attitudes, subjective norms and perceived behavioural control. Also, they acknowledge that behaviour change is a process occurring through stages. For this reason, they provide features aimed at acting on behavioural determinants and at supporting progress from a stage to the next one. Within this general framework, each of the apps has its peculiar characteristics, as summarised in Table 7.1.

In summary, enCompass focuses on the individual level. It offers goal setting features and provides customised recommendations based on the energy consumption patterns learnt by the app itself, also leveraging gamified features such as points, badges, leaderboards and real-life prizes based on progress towards one's goal. It also offers features covering all five behaviour change stages identified by the Transtheoretical Model of behaviour change, starting from "consciousness raising" process in the Pre-contemplation phase

Table 7.1: Summary of the key features of the three persuasive apps I analysed.

Characteristics	enCompass	Social Power	Social Power Plus
Type of energy consumption addressed	Heating (heat pumps, direct electricity), appliances and lighting	Appliances and lighting	Heating (heat pumps, direct electricity, gas), appliances and lighting
Individual goal setting	Yes	No (goals at team level)	Yes
Disaggregation of consumption feedback	No	No	Heating/Non-heating
Customisation of recommendations	Yes	No	No
Use of external rewards (points, real-life prizes)	Yes	Yes	No
Reference to household practices	No	Yes (through individual challenges)	Yes (through individual challenges)
Social influence	No	Indirect interaction (team level goal and feedback)	Direct interaction (pinboard)
Duration of intervention	One year	Three months (plus nine months of "low-interaction" app availability)	Three months (plus nine months of "low-interaction" app availability)

via the provision of eco-feedback to enhancing feelings of "helping relationship" via the provision of customised notifications in the Maintenance stage. Also, enCompass attempts to cover for all residential-related energy consumption types, by accounting for both heating and non-heating related consumption (lighting and appliances). For heating, however, only consumption due to heat-pumps or direct electricity heating system can be considered —namely, enCompass is only fed by electricity smart meters. The app was made available to 55 self-selected households in the Swiss municipality of Contone for one entire year between years 2018 and 2019.

Conversely, the Social Power app enriches the individual focus and attempts to also include a social dimension. For this purpose, it introduces the idea of "teams of households" and engages them in either collaborative or competitive energy saving processes, depending on the app version, by attributing team level energy saving goals and providing team level feedback on consumption. Though gamified features are offered to favour teams' progress (badges, points), teams are not provided with direct interaction possibilities: team members can only see the overall team progression towards the goal, besides their individual level consumption feedback. Social Power also introduces the idea of tackling specific energy consumption practices that are performed in the household, by means of challenges and non-customised tips addressing use of energy consumption appliances typically available in households. It only addresses electricity consumption for appliances and lighting, without considering heating or hot water production. Finally, also Social Power offers features that can deal with all behaviour change stages identified by the Transtheoretical Model, from "consciousness raising" to "helping relationships", via push notifications. The app was made available to 108 households in the Swiss municipalities of Massagno and Winterthur, over a period of three months in Spring 2016, plus nine additional months of “low-interaction” app availability. Push-notifications were only sent during the three-month “high-interaction” intervention period.

The Social Power Plus app again focuses on the provision of feedback at the individual level; however, it leverages social influence techniques through the pinboard feature (in-app virtual forum), that invites to share experiences, suggestions, and difficulties with other app users that are engaged in the same process. Specifically, Social Power Plus invites households to engage in challenges that tackle eight energy consuming practices, by adopting a perspective that aims at bridging between Theory of Planned Behaviour and Social Practice Theories. The aim is to foster active and passive social interactions around those practices via the pinboard and favour the evolution of shared meanings around energy consumption practices, towards energy sufficiency. Differently than the other two apps, Social Power Plus acknowledges that its users are likely to have already started their path along the behaviour change stages identified by the Transtheoretical Model. Therefore, its features start from the Contemplation stage, through a “self-reevaluation” process based on energy consumption feedback, and then support users until the Maintenance stage, through (non-customised) reminders provided via push notifications. The app was made available to 220 households in the Swiss regions of Wil, Schaffhausen and Winterthur for a three-month “active use” period in Spring 2022, followed by additional nine months during which no new challenges were released, though all features were still available, including by-weekly notifications aimed at avoiding relapse. Finally, similarly to enCompass, Social Power Plus targets all types of residential energy-related consumption, by including heating (also produced through gas) and providing users with estimated break-down in heating/non-heating consumption.

Outcomes of the analyses I performed on the three cases are summarised in Table 7.2, which reports the average treatment effect on the treated (ATT), the effect size (through Cohen’s d) and the level of statistical significance of quantitative analyses I performed. I now discuss these results, by first commenting on the effects during the intervention (RQ1), including the heterogeneity analyses I performed (RQ3). Then, I deal with the critical lack of maintenance of intervention effects over time (RQ2) and finally focus on the level of engagement by different app features (RQ4).

7.2 Short-term effects of app use

Overall, findings from the cases show that interventions based on the enCompass and Social Power app were effective in producing savings in households’ energy consumption and carbon emissions (RQ1). On average, during the intervention period use of enCompass reduced households’ consumption and emissions by 4.95% (small effect size measured through Cohen’s d) and use of Social Power reduced them by 9.23% (medium effect size). In both cases, the obtained savings are comparable with the upper-end of the range identified by early smart meter feedback studies reviewed by Darby et al. (2006), Fischer (2008) or Delmas et al. (2013) —with the difference that, for my two cases, I

Table 7.2: My overarching research questions and the results I obtained.

RQ	Research question	enCompass	Social Power	Social Power Plus
RQ1	Can the use of behaviour change apps produce a reduction in residential energy consumption and related CO_2 emissions of households?	During the one-year intervention: ATT: -4.95% **, d: -0.35 (small).	During the three-month intervention: ATT: -9.23% ** (-6.00** kWh/week), d: -0.51 (medium).	— (Not addressed)
RQ2	If app use is found to produce a reduction in energy consumption and CO_2 emissions during or immediately after the intervention, is such a reduction also observed long after the end of the intervention?	No statistically significant ATT is found in the long-term (during one and two years after the end of the intervention).	No statistically significant ATT is found in the long-term (one year after the start of the intervention).	— (Not addressed)
RQ3	Are the effects on energy consumption and CO_2 emissions constant, on varying observable characteristics of households? Or does heterogeneity in observed characteristics of households lead to heterogeneous effects as well?	Households using electricity for appliances have a statistically significant different ATT than households with boiler and heat pump: ATT: -14.46% ***, d: -0.91 (large). No statistically significant differences in ATT are found on varying the level of app use.	No statistically significant differences in ATT are found on varying the location, the type of household and the type of house.	— (Not addressed)
RQ4	Which app features can foster higher user engagement, thus providing greater support to the reduction of energy consumption and CO_2 emissions?	— (Not addressed)	— (Not addressed)	Apart for specific user groups, individual feedback and goal setting are more engaging than features leveraging social interaction. Neither active nor passive social learning take place.

Significance levels: * 0.1, ** 0.05, *** 0.01

have tried to overcome the methodological limitations that had led Delmas et al. (2013) to argue that higher and more promising effects were obtained by weaker studies from the methodological point of view.

7.3 Heterogeneity of effects of app use

For the cases of enCompass and Social Power I also performed analyses on the heterogeneity of the effects among sub-samples of app users (RQ). In most of the cases, no statistically significant differences have emerged between the considered sub-samples. I cannot exclude that those results also depend on the small sample sizes considered in enCompass and Social Power and therefore on the very small size of their sub-samples I considered for heterogeneity analyses: small sample sizes might have negatively affected the power of the analyses I performed and their capability to detect statistically significant effects. However, for the time being I can only draw my conclusions based on them —and suggest that, if heterogeneous effects are deemed relevant and plausible, future research should perform similar analyses by engaging larger samples of app users, to verify whether similar or conflicting results are obtained.

7.3.1 Heterogeneity on varying the level of app use

I performed analyses on the effect of app use on varying the amount of interactions with the app only for the case of enCompass. No statistically significant differences have emerged: the intensity of the effect has not varied with the intensity of app use. This result suggests that behaviour change might have been stimulated by the very fact of having accessed a persuasive app (and possibly also by being part of a research project). If so, following Theory of Planned Behaviour, behaviour change would have stemmed from an increase in subjective norms and perceived behavioural control around energy saving, triggered by the possibility to access the energy saving app, much more than by specific features it is endowed with.

7.3.2 Heterogeneity on varying house(hold) characteristics

Regarding Social Power, no statistically significant differences in the effects have emerged, either considering the location or the type of household or the type of house. Regarding enCompass, instead, heterogeneity analyses have shown that the effect on households that only use electricity for lighting and appliances is statistically different from the effect on households using electricity also for hot water and/or room heating purposes. Indeed, in the former sub-sample of households the average treatment effect on the treated (ATT) is equal to a 14.49% reduction in energy consumption and carbon emissions, which corresponds to a large effect size.

As Social Power was only addressing the latter category of households using electricity only for lighting and appliances, this result indicates that enCompass was more effective than Social Power (ATT equal to 4.95%, with small effect size). A possible explanation of such higher effectiveness might be attributed to the presence, in the enCompass app, of customised recommendations and goal setting features: customisation possibilities might have provided users with a more engaging experience, since, as predicted by the Self-Determination Theory, they might have favoured higher feelings of autonomy. Thanks to higher levels of "integrated regulation", users' motivation to use the app and implement its suggestions might have been higher. In turn, following the Theory of Planned Behaviour, users' perceived behavioural control and/or subjective norms might have been higher. These results also seem to suggest that the social influence features, that were central to Social Power, were less effective in driving higher savings than the customised features operating at the individual level. This result is aligned with reviews by Abrahamse, Steg, et al. (2005), Osbaldiston and Schott (2012), and Karlin et al. (2015), who found that individual feedback and commitment opportunities through goal setting techniques are more promising behavioural intervention techniques than social-influenced based ones. This result might however be due to the specific way the social interaction features were implemented in Social Power. The automatic attribution to teams engaged in the collective effort towards energy saving, either in the competitive or in the collaborative fashion, was in fact questioned by the app users themselves, for being not sufficiently clear and intuitive. Therefore, the Social Power app might not have properly exploited the social influence features it aimed to leverage.

No statistically significant differences have emerged from heterogeneity analyses on Social Power users, by either considering the location (Massagno vs Winterthur), the type of household (family vs single adult), and the type of home (apartment vs house). From a policy-making perspective, these results suggest that, not only the Social Power treatment can be generalised across different types of household and house types, but also across locations and regional contexts. This thus speaks in favour of scaling-up the intervention through replication to other regions.

Overall, outcome of the heterogeneity analyses on enCompass suggests that households that do not use electricity for heating purposes (either of spaces or of hot water) should preferably be targeted for use of persuasive apps, as in this category of households the largest effect sizes were obtained. Or, more simply, these results suggest to only target via persuasive apps the non-heating share of households' energy consumption. In non energy retrofitted buildings, however, the share of energy consumption for non-heating purposes is definitely lower than the share of energy consumption for heating purposes (Citherlet and Defaux, 2007; Wang et al., 2018). Specifically regarding Switzerland, recent estimates by Kemmler and Spillmann (2021) indicate that 64.5 % of Swiss households' energy consumption is due to heating purposes, 15.5 % to hot water, and only the remaining 20% is due to non-heating purposes (p. 48). These estimates of energy consumption break-down in Swiss residential buildings thus indicate that, by only intervening on consumption for non-heating purposes, the overall absolute amount of saved energy would be low. Choosing to only focus on non-heating household consumption would therefore be not fully beneficial to the energy and climate transition. Furthermore, this would require to envision different types of intervention to effectively intervene on hot water and heating demand. Focusing on interventions that are more deeply entrenched with daily household practices, as suggested by Social Practice Theories, might be an alternative strategy to address the heating-related energy demand.

7.4 Long-term effects of app use

Despite the statistically significant effects in the desired direction that were estimated during the intervention period, for both enCompass and Social Power the statistical significance disappears in the long-term, from one year after the end of the interventions onwards (RQ2). This implies that I cannot draw conclusions on the long-term intensity and direction of the effect on the treated. By considering the practical significance, however, the estimated ATTs of the interventions even indicate small increases in energy consumption and carbon emissions, for both enCompass and Social Power. Findings from these cases namely show that, during the maintenance phase identified by the Transtheoretical model of behaviour change, relapse to previous behaviour has occurred.

I cannot exclude that the observed relapse is due to a change in the characteristics of the households participating in the enCompass and Social Power projects, such as for instance the birth or arrival of a new family member, the purchase of a new highly energy consuming appliance such as a fish tank, or the purchase of an electric vehicle. However, I assume that these situations have not systematically occurred in the majority of the involved households and thus consider the obtained long-term results as the evidence of the lack of intervention effectiveness.

Indeed, this is not uncommon. Despite the literature I reported in Chapter 3 indicates that long-term effects of behavioural interventions have seldom been analysed, the concept of relapse is included in any stage model of behaviour change —including the one by Li et al. (2011), that only considers the two stages of *discovery* and *maintenance* of new behaviour: during maintenance, falling back to discovery-stage energy consuming behaviour and routines is always possible. With reference to Theory of Planned Behaviour, it appears that feelings of increased perceived behavioural control and/or subjective norms that had occurred throughout the intervention progressively disappeared as long as the intervention was discontinued. This also suggests that the novel energy consumption routines suggested by the app —which were indeed put into practice during its use— have not been fully internalised by the household members, up to the point that they have become fully automatised.

7.4.1 Supporting long-lasting change

Relapse may indeed be due to a number of reasons. Though literature is clear it may happen, there is no clear agreement on the reasons why it may occur —which actually reflects the lack of agreement among scholars and the large number of theories that have been developed around behaviour and behaviour change. The Self-Determination Theory might however be particularly relevant to explain relapse situations: it suggests that, when extrinsic motivational factors such as those provided by an app remain at the level of "external regulation" and are not integrated into the self (Ryan and Deci, 2000a), effects of an intervention tend to disappear. As indicated by Ryan and Deci (2000b), such a lack of internalisation is more likely to appear if the intervention fails in providing a context supportive of autonomy, competence, and relatedness: from this perspective, internalisation of behaviour, which lays the grounds for self-determined, long-lasting behaviours, is only likely to occur if individuals feel competent to perform a given behaviour and if the social context provides support for autonomy, within collective engagement processes.

Dealing with the post-action phase of the "Model of Action Phases" by Bamberg (2013), Ohnmacht et al. (2017) have suggested that relapse might be avoided by keeping the provision of feedback and reminders and by activating community-based strategies. The continued provision of feedback has however been offered by Social Power (which kept providing hourly consumption feedback for nine months after the active intervention phase has ended, however without notification reminders), without maintenance of the effect in the long-term. Indeed, as Hargreaves, Nye, et al. (2010) and Hargreaves, Nye, et al. (2013) have indicated about In-Home-Displays providing feedback on smart meter consumption, feedback may tend to become "backgrounded" into everyday routines and to fade over time. A similar conclusion can be drawn from the work by Geelen et al. (2019), that provides insights from interviews with participants to an app-based feedback intervention in the Netherlands: at first, the app was checked frequently, to get insights on the household's consumption patterns. Once such patterns were understood, then the app was mostly used only from time to time, to monitor whether the levels of consumption remained on the same patterns. After an initial "learning period", the information provided by the app was perceived as of low relevance and limited usefulness. The app thus tended to be ignored, remaining in the back of the smartphone, without being ever opened: "out of sight, out of mind" (Hargreaves, Nye, et al., 2013).

When they are not committed to deliberatively open the app on their smartphone, individuals do not receive the feedback or the other features provided by the app. Temptation to relapse to old behaviour is thus more likely to happen. As spontaneous commitment is unlikely to occur, Ohnmacht et al. (2017) recommended to regularly send reminders via push notifications, to trigger users to check their consumption feedback (possibly through customised notifications if consumption data shows an increase with respect to a baseline) or simply to open the app and keep interacting with its features. Such a strategy has for instance been implemented in Social Power Plus: by-monthly push notifications are in fact sent to app-users during the "low-interaction intensity" period of availability of app use, after the three-month "high-interaction intensity" intervention. The analyses on the energy consumption data collected in 2021 and 2022 for the treatment and control households, which will be performed in Spring 2023, will provide insights on the effectiveness of this strategy.

Notifications provided into the smartphone itself are however in competition with a myriad of other stimuli that are offered by the smartphone itself. Individuals are in fact growingly affected by *information overload* phenomena (Milgram, 1970), which lead them to disregard what they perceive as low-priority inputs. When information supply exceeds information processing capacity, individuals become highly selective and ignore large amounts of the information they receive (Eppler and Mengis, 2008). In the specific context of overabundant sensational stimuli offered by smartphones and digital devices, research has clearly shown that individuals tend to make inconsistent and irrational choices, that favour immediate gratification instead of long-term utility maximisation (Gui, Shanahan, et al., 2021). Reminder push notifications provided by persuasive apps are in conflict with many other push notifications by other apps installed on the smartphone, included social media. The risk such notifications are ignored is thus high, as other apps may be perceived as more immediately gratifying than persuasive apps. When persuasive apps adopt gamification techniques, they may send rewards through badges, leaderboards, or goal achievement feedback. However, such rewards may not be sufficient to get the individual's highly selective attention. Furthermore, highly relying on push notifications might exacerbate problems of digital wellbeing (Gui, Fasoli, et al., 2017), favouring the appearance of problems of smartphone dependence and negatively affecting the individuals' cognitive capacity (Ward et al., 2017).

Based on these considerations and on the literature I reviewed and summarised in Chapters 2 and 3, I therefore think that the other suggestion by Ohnmacht et al. (2017), namely to resist relapse by building on community-based strategies, may be more effective. This approach was for instance proposed by McKenzie-Mohr (2000a) and McKenzie-Mohr (2000b), who suggested to enact community-level strategies that bring individuals together (either through in-person contact or through digital channels) and build confidence among them that the desired behaviour will be performed by many members of the community. A possibility might also be to go beyond the idea of simply sharing a collective endeavour and to perform interventions that explicitly act on the material and cultural contexts surrounding individual households, by accounting for non-energy policies, decisions, or practices that can influence energy-related behaviour, and by engaging with the relevant actors that can influence and shape them. This might for instance be attempted through a living lab approach aimed at questioning and challenging current social norms and conventions around energy-consuming practices.

Attempts of interventions combining use of Social Practice Theories with living lab approaches, in which individuals are brought together to challenge existing household practices and conventions and to co-create and test novel solutions to support the energy and climate transition, have been for instance identified (Heiskanen et al., 2018) and already implemented (Sahakian, Rau, et al., 2021). According to their authors, these interventions can also be used to address heating-related energy consumption and therefore might be beneficial to tackle the heating-related issue that I have discussed in the previous section, while striving for long-term changes.

7.5 App features and user engagement

The idea to perform interventions that explicitly address daily practices, with the aim of changing them, is not new. For instance, in one of her early works on Social Practice Theory, Shove (2003) argued that the conventions and social norms on personal hygiene, cleanliness, and comfort, should first of all be addressed in order to save energy in households —rather than focusing on behaviour change at the individual level. Similarly, Hargreaves, Nye, et al. (2010; 2013) have remarked that people tend not to accept changes that they perceive as renouncing to something, which would be the case for a reduction in indoor temperatures —perceived as a reduction in thermal comfort at home. This perception might be counteracted by the provision of tips, aimed at increasing competence on how to achieve energy saving without decreasing one's comfort. This was exactly the aim of the tips provided by all three apps I have analysed with respect to heating or showering activities and, more broadly, to all challenge topics. More radically, however, and again inspired by Social Practice Theories, Hargreaves (2018) has suggested to fundamentally rethink the role of energy feedback, by broadening analysis to include dynamics of everyday life and the socio-technical system underlying it. This could for instance be achieved by working with community groups of peer networks, rather than individuals, and by exploring novel ways to design forms of feedback that encourage reflection on social conventions, habits and routines, rather than only on energy consumption data per se.

Overall, even though this suggestion comes from a different theoretical perspective grounded in Social Practice Theories, this is very much aligned with the above suggestion by Ohnmacht et al. (2017), grounded in social and environmental psychology theories, to develop community-based strategies. And this is what was attempted by Social Power, through the team-level collaborative and competitive goal and feedback features, and especially by Social Power Plus, through the challenges and pinboard features. Design choices that were made in Social Power, first, and in Social Power Plus, later, assumed in fact that features related with social influence and social interactions, mediated by a digital technology such as the app, might "make the difference" and favour long-lasting behaviour change. However, the way social interaction features were developed in Social Power was not sufficiently clear and appealing to activate change: neither the composition of teams was clear to app users nor within-team active or passive interaction possibilities were offered, so that ultimately social influence features were not really triggered. Social Power Plus, instead, was specifically designed with the aim of leveraging the effect of social interaction features, which were therefore carefully designed. Nevertheless, analyses I performed on the case of Social Power Plus show that social interactions mediated by the pinboard feature did not stimulate high user engagement and did not manage to activate a social learning process on a broader scale: most of the app users were only limitedly engaged with the pinboard feature aimed at providing a venue for open discussion and sharing of experiences about the processes of changing routine behaviours related with the challenge topics. Also, the evidence provided by survey responses, especially the open-text comments, and by the materials published on the pinboard, indicates that the majority of app users were more interested in their own individual consumption data, rather than in challenging or re-negotiating collective conventions and norms with other app users.

According to findings by Social Power Plus, and to address research question RQ4 about which features foster greater user engagement, it seems in fact that individual app features were more appreciated by app users than social app features. The evaluations provided by app users, indicating that they were more engaged by individual features than by features leveraging social influence aspects, seem to challenge the idea that a focus on app-based community-based strategies, in the way that was attempted and implemented by Social Power Plus, can produce long-lasting changes.

Further research is therefore definitely needed to explore these topics. In fact, on the one hand the suggestions by Ohnmacht et al. (2017) and Hargreaves (2018) were not informed by strict experimental evidence of effectiveness. On the other hand, my current findings from Social Power Plus might be affected by limitations. In fact, the conclusion that individual features are more appreciated by app users than community-based strategies might be biased by the specific composition of the sample of participants to Social Power Plus. Most of those households had in fact already implemented energy efficiency interventions aimed at insulating their house and optimising their energy consumption: they were already in the action or maintenance stage, according to the Transtheoretical model of behaviour change. Therefore, they were less open to reducing their residual energy demand by means of behavioural interventions aimed at re-thinking their daily practices. It is not clear yet whether a different sample of households, a bit behind in the behaviour change process, would have made the same evaluations. Besides waiting for the outcome of the quantitative analyses of impact of Social Power Plus, further research would therefore benefit by replicating use of the Social Power Plus in other contexts, and especially with different types of households, to either consolidate or reverse the findings I obtained.

The low interest in the social interaction features offered by Social Power Plus might however also be due to a lack of clarity by the research team in framing the needed energy and climate transition as a collective endeavour, within which all societal actors, including the three utility companies supporting the intervention and governmental and non-governmental representatives of the regions where the intervention took place, need to collaborate in order to rethink and re-shape current societal organisation, expectations, cultural norms, and material needs. Therefore, also from this perspective I suppose that a living lab approach, envisioning the strict collaboration of different private, public and civil society actors, might have better supported the communication of the urgency to tackle together a common challenge, and thus might have given more visibility, value, and intrinsic utility to the social interaction features offered by the pinboard.

From my viewpoint, however, key issues with the living lab approach reside in how to ensure the participation by a large share of individuals, beyond those that have already clearly shown awareness and intrinsic interest towards energy and climate topics, and possibly have also implemented pro-environmental behaviours in their daily routines. Namely, one crucial difficulty would remain open and challenge the whole process: the difficulty of going “beyond the converted”.

7.6 Going beyond "converted"

In all the three case I have analysed, invitation to join app use was offered as an “opt-in” opportunity: app use was made freely available to any interested households, provided that they accepted technical eligibility requirements. For enCompass and Social Power Plus such requirements certainly played a role in the self-selection of households. In the former case, in fact, they were requested to accept installation of a set of temperature and humidity sensors to feed the algorithms for customised recommendations. In the latter case, in the region of Schaffhausen households had to accept installation of similar sensors to collect electricity consumption data, and in the regions of Winterthur and Schaffhausen households had to be equipped with heat pumps or electric heating systems, as no gas smart meters were installed.

I assume that the requirement of being equipped with an heat pump introduced some bias in the energy saving motivation of intervention participants, compared with average households: even though they are increasingly installed in new buildings and in energy retrofits of existing buildings, in Switzerland heat pumps are not the building standard yet. Due to their higher cost with respect to gas heaters, I suppose that they are installed by individuals with higher than average pro-environmental attitudes (and income). I also assume that the need for installing a sensor acted as an additional barrier to the decision to join project activities: it is likely that households that decided to join project activities were more motivated to save energy than average households, either for environmental or for monetary reasons. This is why, I argue that it is likely that those apps mostly raised the interest of already motivated and climate and energy-aware households. This also implies that they had limited room for reducing their consumptions and emissions, compared to average households, as it is likely they had already implemented either behavioural or technical measures to reduce energy consumption —or both.

The case of Social Power also provides me with a further indication: in the Region of Massagno (one of the two regions were the app was tested), the requirements to join project activities were minimal, as electricity smart meters were widely spread across the region and no specific requirements to measure heating consumption were needed (as the app only focused on electricity consumption for non-heating purposes). Well, also in this case getting 100 participants, which was the initial target for the number of households to be involved, was difficult. In the region of Massagno, in fact, project activities started with only 54 households that applied to join app use, and then only 23 households complied with app download and registration of an account. The number of interested app users was therefore very small. I suppose that those 23 households that downloaded the app and at least registered on it, did so as their motivation was higher than average. Hence, I cannot exclude that also those households had already performed behavioural (and maybe even technical) energy saving interventions before the start of project activities and thus that also in their case the room for further energy saving activities was smaller.

Does this mean that these apps were "preaching to the converted"? If those who could definitely benefit the most from their use are indeed not sufficiently interested to access and use them, the overall value of these apps would be questionable. Namely, the apps would fail to reach their actual target group, and therefore, even assuming that they would manage to activate a long-lasting change, their impact would in any case be limited. Also, this would indicate a clear obstacle to possible large-scale deployments 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. Persuasive apps would in fact 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 persuasive apps are universally offered under "opt-out" frameworks, learnings from my case studies suggest that apps 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 A. Nilsson et al. (2014), who, in a randomised intervention providing electricity consumption feedback through In-Home-Displays to Swedish households, found that lack of interest in energy saving, together with insufficient understanding of the information provided by the display itself, was an important barrier 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 the persuasive apps, I expect that their practical impact would not be higher than the ones resulting from my case studies. Instead, it might possibly be even lower: households would have more room for implementing energy saving actions; however, as most of them would tend to ignore the app's features, the average energy saving effect due to app availability would be lower than the effect I found by considering self-selected, interested households.

Future research would be needed to scrutinise whether differences exist between the target population and the participants to the three interventions I analysed, by identifying to how much they amount and on which household characteristic they occur. For this purpose, new questionnaires might be administered to the participating households, by using 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, representative of the Swiss population (`https://www.sccer-crest.ch/research/swiss-household-energy-demand-survey-sheds/`, Weber et al., 2017). In any case, the problem of "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 samples are made by volunteers who actively decided to join interventions due to their interest in energy topics, findings by 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 its authors do not explicitly discuss it, the problem also remains for other approaches such as the one attempted by Sahakian, Rau, et al. (2021), that aimed at challenging social practices through living labs. The authors in fact 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 lab activities, which were performed through in-person meetings and group interviews. Under such a demanding plan of activities, I 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 an heterogeneous sample and at including traditionally less represented groups, such as unemployed or single-parent households), I presume 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 broad, 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.

7.7 Cost-benefit effectiveness of persuasive apps

An additional discussion element stemming from analysis of my case studies, which is closely connected with the above one about "converted" and "non-engaged" households, deals with the overall number of app users. As I indicated above, in the case of Social Power in Massagno the number of voluntary app users was definitely lower than expectations. Nonetheless, this was the case also for the other regions, where technical eligibility requirements constrained the size of the eligible population. In all the three case studies and regions that I considered, in fact, the number of households participating in project activities was lower than planned at the research design stage.

By considering these difficulties in recruiting app users, one may wonder what is the cost-effectiveness of persuasive apps. One of the beneficial characteristics of ICT and app-based interventions, in fact, is their scalability: provided that sufficient server space is available to process and store consumption and any app-related data, use of apps can potentially be extended to more users with no (or very little) marginal economic cost. What happens, however, if the number of app users remains limited? Particularly, what is the cost-benefit impact of persuasive apps? This is a highly relevant practical question for policy-makers, who need to choose between possible interventions aimed at supporting the energy and climate transition within a framework of limited available resources. Particularly, comparing the costs associated with the apps' development and maintenance activities and the benefits they deliver offers a valuable support to public decision-making activities: if the social and environmental benefits delivered by these apps are lower than the social and environmental costs associated with their development and maintenance over time, relying on them as tools to support the energy and climate transition would at least be questionable.

A cost-benefit analysis by Wemyss, Cellina, Grieder, et al. (2022), still unpublished and currently under review, estimated the number of app users needed to achieve the break-even point for the Social Power app. Specifically, the analysis estimated the number of app users that are needed for the monetary estimate of the direct climate benefits of app use (in terms of reduced CO_2 equivalent emissions valued at the social cost of carbon) to outweigh the app's development and maintenance monetary costs. The latter costs were provided by the professional software developers, while the average treatment effect resulting from app use, in terms of percentage reduction of electricity consumption, was used to indicate the app's climate bebefits. To get conservative estimates, the overall electricity saving effect was assumed to be equal to zero after eighteen months from the start of app use, and to linearly decrease over time between month three (immediately after app use, when the short-term estimate I reported in Section 5.6 were computed) and month eighteen. Further, the related amount of saved electricity was translated into saved tons of CO_2 equivalent emissions, by considering current electricity emission factors for Switzerland and for Germany —a country with a much higher carbon footprint than Switzerland (see Section 1.5). Finally, a monetary estimate for the social cost of carbon was selected, namely an estimate of the monetary impacts on the environment, the economy, and human health of the emission of one ton of CO_2 equivalent in the atmosphere. Also, a discount rate was chosen[1].

The resulting number of users needed to achieve break-even point is huge, if compared with current number of app users: for Switzerland, depending on assumptions on the development and maintenance costs, break-even for the Social Power app would be reached with a number of users varying between 36'000 and 190'000. For Germany, whose electricity carbon footprint is higher, and therefore the same percentage reduction in consumed electricity has a higher contribution to the reduction of carbon emissions, the number of users should be between 14'000 and 73'000. Since for enCompass the treatment effect was of comparable order of magnitude as Social Power (and so were its costs), a number of users of similar magnitude would be needed as well. These figures indicate that a massive number of app users should be pursued to achieve break-even —which looks to be out of reach in the contexts that I have examined, considering the difficulties experienced during recruitment activities for app users.

However, one has to consider that any climate mitigation measure has a cost. Therefore, the number of "break-even" users might also be computed by considering the climate benefits produced by app use against a cost threshold that policy-makers can accept to pay for. Doing so, the number of users needed for break-even would be lower. Still, these analyses suggest that, unless a definitely large upscaling in the number of app users occurs, from a cost-benefit perspective these apps are not a wise policy choice.

7.8 A changing context

Current socio-political contextual conditions, characterised by the energy crisis triggered by the Russian war in Ukraine, might indeed shift the ground. Motivation to save energy has so far been limited to a small part of the population, driven by pro-social and pro-environmental attitudes and the related subjective norms. The current lack of availability of gas resources and the consequent increase in all energy prices might however drive an increase in motivation by a broader share of the population: first of all, in fact, the sharp increase in the cost of energy experienced in Switzerland in the last months of 2022 and also expected for 2023[2] might drive an increase in individual motivation to energy-saving, for purely economic reasons.

Furthermore, the way energy saving is framed by mass media and by the popular press may also drive an evolution in social norms. For instance, at the end of August 2022 the

[1] A value of 175 USD per ton of CO_2 equivalent was chosen, which is within the currently most frequently used estimates according to the expert survey by Pindyck (2019). Regarding the discount rate, a 3% value was chosen, which is common for social cost-benefit analyses performed in industrialised countries.

[2] `https://www.elcom.admin.ch/elcom/it/home/documentazione/medienmitteilungen.msg-id-90237.html`, last accessed on January, 27 2023.

Swiss Federal Office of Energy launched a communication campaign aimed at spreading the message that "energy is scarce" and inviting the population to activate energy saving measures within the household. The campaign, funded by an overall two-million Swiss francs investment, aims at reducing overall energy consumption, in order to guarantee the energy demand across the country can be fully met, even under gas shortages from Russia and the shortages in electricity import from French nuclear power plants that are undergoing maintenance work. The campaign provides the population with five tips about how 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)[3].

At the time of writing this dissertation, assessment of the effects of the war combined with the campaign have not been analysed yet, as energy consumption data are still manually read in many regions of the country and readings will only be available at the start of 2023, for calendar year 2022. Indeed, a large-scale shift in shared social meanings and conventions about consuming (and saving) energy in households may be taking place. By favouring the diffusion of a novel social context and culture supportive to low-energy consumption practices, the combined effect of the war and of the energy-saving campaign might increase household motivation to save energy.

It is also likely that, in such a changing context, the share of households willing to experience use of persuasive apps gets higher, particularly among previously "non engaged" households, who would now be motivated to start using the app and significantly reduce their consumption. Launching an app-based intervention aimed at energy sufficiency in the next months might thus provide different outcomes than the cases I have analysed, in terms of both higher effect sizes, statistical significance and maintenance of the energy saving effect over time. It may also happen that, in such a supportive context for energy sufficiency, "social" app features such as those allowing exchange of experiences in virtual forums (e.g. the pinboard offered by Social Power Plus) would manage to engage app users much more than, so far, they have managed to do. Results I obtained in my analyses might thus be different, if the related interventions were repeated in the specific contextual energy crisis situation of years 2022 and 2023 (and possibly even later).

[3] `https://www.dont-waste.ch`, last accessed on January, 27 2023.

8 Conclusions

> *Never doubt that a small group of thoughtful, committed citizens can change the world; indeed, it's the only thing that ever has.*
>
> — **Margaret Mead**
> Anthropologist

In this chapter I summarise the analyses I performed, by referring to the research questions I dealt with throughout the dissertation, and present the results I obtained. I highlight the overall contribution that my research offers to understanding the effectiveness of persuasive, gamified apps and their role and contribution to the energy and climate transition. I also reflect on the limitations of my activities, identify points for improvement for future applied research, and focus on open issues requiring additional investigation. I conclude by identifying promising venues for future policy interventions. I hope that the outcomes of this dissertation help to deliver a concrete and long-lasting impact for a more sustainable, low-carbon and energy sufficient society.

8.1 Contributions of this work

The work I performed for this dissertation aimed at shedding light on the contribution of persuasive, gamified apps to the grand societal challenge of mitigating climate change and supporting the energy and climate transition goals identified at the national and international level. Whether and why these apps are effective or not in delivering energy and carbon emissions savings is in fact an open research issue in the existing literature so far: as I have shown through Chapters 2 and 3, it is not clear yet whether investing in them is a valuable and beneficial policy effort. Scientific literature on persuasive, gamified apps has in fact shown that, like most of non app-based behavioural interventions developed in the last fifty years, their effects have not been properly estimated. A number of methodological limitations were in fact found to affect the procedures used to estimate their outcomes and assess their energy and climate impact.

Specifically, this work unfolded through the analysis of three persuasive apps aimed at reducing residential energy consumption in households, that were developed in Global North countries between 2016 and 2022: enCompass, Social Power and Social Power Plus. These apps have been field tested with self-selected voluntary households (treatment groups) across different Swiss regions, with the support of the local utility companies. The design of the apps' features has been performed by the respective research teams before and outside the work for this dissertation, apart for the case of Social Power Plus, for which I was involved of the broader process for the co-design of its features.

Regarding enCompass (Chapter 4) and Social Power (Chapter 5), I first showed how their features are grounded in prominent behavioural theories, how they relate to key behaviour change techniques, and how they fit with persuasive systems design principles. I then estimated their effectiveness through quasi-experimental research designs, by identifying matched comparable control groups among non-treated households served by the same utility companies. For this purpose, I performed Two-Ways Fixed Effects panel regressions models, through which I estimated the average treatment effect on the treated (ATT) and its effect size, both in the short-term (during the intervention itself) and in the long-term (after the end of the intervention, for one or two follow-up years depending on the case). Also, I verified whether heterogeneous effects can be identified, on varying the available observed characteristics of the households (families/sigle adults), of their houses (building type: apartment/independent house, purpose of use of electricity: only appliances and lighting, also heating of water, also heating of spaces), or the location where they live (municipality of Massagno/Winterthur). For enCompass I also looked for possible different treatment effects depending on the intensity of app use.

For Social Power Plus (Chapter 6), not only I contributed to the evaluation of its effects, but I was also involved in the design of its features. The latter were in fact the outcome of a co-design process performed with a group of potential app users engaged in a “living lab” activity. Again, in this work I showed how the app’s features are grounded in behavioural theories and specified which behaviour change techniques and persuasive principles they exploit. Ongoing research activities envision a quasi-experimental panel data regression analysis also for the case of Social Power Plus. In this case, however, the timeframe of this dissertation is not compatible with reporting and discussing these activities, as the energy consumption data required for these analyses will only be available at the start of year 2023. Instead of estimating the ATT, therefore, I analysed the Social Power Plus app and the related intervention to explore the interactions between the app and its users and to understand which app features —namely, which behavioural intervention techniques— managed to raise more engagement by its users. Also, I assessed whether either active or passive social learning processes between app users were triggered by app use. Analysis of this case therefore provided me with specific insights about how app-based behavioural interventions can (and cannot) produce their outcomes.

The research methodologies and approaches that I adopted to analyse these cases address three critical limitations that had emerged from the analysis of previous literature on persuasive, gamified apps and, more broadly, behaviour change interventions (see Section 3.7). First of all, whereas most of past app-based interventions were poorly grounded in behavioural theories, I explicitly related all app features to theoretical determinants of the energy consumption behaviour change process. This allowed me to get more insights on the mechanics through which interventions can have higher effectiveness. I also explored evaluation of app features from the perspective of app users, thus enhancing the capability to understand why intervention are effective (or not). Second, I adopted quasi-experimental research designs, by considering baseline consumption values and control groups of households that are matched with observable characteristics of the treatment group households. This is a more rigorous methodology than those applied in many previous interventions, even though it does not achieve yet the "gold standard" represented by randomised controlled trials. Moreover, apart for the case of Social Power Plus (wherein, due to constrains on the dissertation delivery time, I can only rely on self-reported data), in the other two cases I rely on objective energy consumption data directly provided by the utility companies via smart meters. This implies the results are not affected by the social desirability bias or specification and measurement errors that usually affect survey-based data collection processes. Third, I analysed long-term effects of app use, by collecting energy consumption data for the considered treatment and control groups of households. Despite previous research has clearly and repeatedly called for investigation of the long-term impacts of behavioural interventions, long-term analyses are still rare. From a policy-making perspective, this is highly critical, as short-term effects might suggest larger behaviour change impacts than they ultimately are —as in fact I found it happened for the cases of enCompass and Social Power. Alternative intervention types might be capable to produce more durable effects, at comparable intervention cost and effort.

8.1.1 Summary of results

Overall, results indicate that, during the intervention period, both the enCompass and the Social Power app were effective in changing energy consumption and CO_2 emissions of households, as shown by the ATTs I estimated. On average, enCompass reduced energy consumption and CO_2 emissions for heating and non-heating purposes by 4.95 % with respect to the consumption baseline measured one year before the intervention (small effect size, d= 0.35). Social Power reduced electricity consumption and CO_2 emissions for non-heating purposes by 9.23 % with respect to the baseline (medium effect size, d= 0.51). In both cases, the ATTs are statistically significant at the 0.05 level.

By only considering households that solely use electricity for non-heating purposes, enCompass managed to reduce energy consumption and CO_2 emissions by 14.46 % with respect to the baseline (large effect size, d= 0.91), with a 0.01 significance level. This therefore suggests that targeting households' energy consumption for purposes different from heating would result in higher energy savings in relative terms. However, as the share of energy consumption for non-heating purposes is much lower than the share of energy consumption for heating purposes (see for instance recent updated statistics for Swiss households reported by Kemmler and Spillmann, 2020), doing so the overall absolute amount of saved energy would be lower.

Despite the promising and statistically significant results in the short-term, by considering follow-up energy consumption data (respectively, one and two years after the intervention for enCompass and one year after the intervention for Social Power), statistical significance disappears. Evidence about the long-term effectiveness of the intervention is thus lacking. Practical significance indicated by the long-term estimates of the average treatment effect on the treated, however, even shows a relapse to pre-intervention (if not higher) consumption levels.

Finally, regarding which app feature and intervention techniques are more effective in engaging app users (and thus in producing the short-term effects resulting from panel data regressions), the case of Social Power Plus suggests that features focusing on the individual level stimulate more engagement than features acting at the social level. App users in fact indicated their positive evaluation of energy consumption feedback and goal setting features (indeed asking for much more detailed feedback on the consumption of their appliances then it was offered by the app), while social influence features (particularly those aimed at supporting active and passive social learning) were the least appreciated by app users.

Overall, despite the Social Power Plus case is still open, and estimates of its long-term treatment effect that will be available in Spring 2023 might confute my current results, these findings tend to dampen enthusiasm about policies focusing on the use of persuasive, gamified apps. While my analyses confirm that app-based interventions were effective in reducing energy consumption and carbon emissions during the intervention period, they also confirm the problem of long-term effectiveness that had emerged in previous literature. Moreover, insights on the limited level of user engagement offered by social influence features seem to challenge the large body of literature that values social influence techniques leveraging peer interactions as beneficial for a long-lasting change (see Section 3.3.3). Based on my current findings and insights from previous related works and their theoretical background, I would therefore conclude that, acting in isolation, interventions leveraging persuasive gamified apps are not effective in driving tangible change for the energy and climate transitions. Nevertheless, I still envision a fruitful way in which persuasive gamified apps could contribute to such transitions. I will introduce it in the final section of this dissertation, after having summarised the limitations of my work.

8.2 Limitations of this work

My work is still affected by some of the methodological limitations that affected previous works, that I have summarised in Section 3.8. First of all, I adopted quasi-experimental approaches, which are weaker than truly experimental approaches in regards to the capability to produce robust causal impact estimates. Moreover, despite the care I devoted to matching comparable control groups to the treatment group households, only a very limited set of variables was available for me to perform such a matching, both for enCompass and Social Power. Therefore, the matched groups might actually largely differ from the treatment groups in many unobservable variables that I could not access. This may affect the internal validity (Vine et al., 2014; Frederiks, Stenner, Hobman, and Fischle, 2016) of my 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 my results.

Another relevant limitation lies in the problem of self-selection of intervention participants, who in all cases were volunteers answering a public call to join project activities. This may affect both the internal and the external validity of my quasi-experimental results (Vine et al., 2014; Frederiks, Stenner, Hobman, and Fischle, 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 interventions were not representative of the interventions' target populations, 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 in the three cases, 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 cases I analysed this was not possible due to the very small sample sizes of households that had registered for project participation. Randomisation within those samples would in fact have implied to halve (or in any case largely reduce) the size of the treatment groups, that were already very small. This would have very likely precluded the possibility to find any statistically significants results, therefore from the very beginning I 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 I have discussed above.

The small size of the samples is, indeed, a critical aspect in itself, which weakens all three cases I have analysed. In particular, the lack of statistical significance I found for the heterogeneity analyses that I performed in order to assess treatment effects for different sub-groups of household types might be due to the small sample sizes and a related insufficient statistical power.

Also, all the three app-based interventions I analysed include a number of different app features (namely, intervention techniques), aimed at motivating and supporting change along the different stages of the behaviour change process. The quasi-experimental research designs I adopted do not allow to disentangle effects by a single intervention technique or app feature. Through the procedures I adopted, I could only obtain aggregated estimates of the overall effectiveness of apps' features and could not identify which of them were more effective —and should therefore be recommended to policy-makers— and which ones should instead be discarded. 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.

Finally, for the case of Social Power Plus I relied on questionnaire data, in order to collect evaluations on the different features from the perspective of the app users themselves. This may have introduced both specification and measurement errors. Also, I used self-reported data to characterise households' routines before and after the Social Power Plus intervention. Responses might have been affected by social desirability bias, as the goals of the intervention (reducing energy consumption and carbon emissions) had been clearly communicated in project-related materials and by the app itself. Further, as I have remarked in Chapter 6, the lack of a control group (at its best, randomly allocated; at its worst, matched through statistical techniques) does not allow to draw causal conclusions on the relation between the observed changes in energy consumption routines and the use of the app, and not even between changes in routines and specific app features. Further work already programmed for the Social Power Plus project will however allow to overcome these latter limitations: when energy consumption data for year 2022 will be available, I will perform quasi-experimental analyses to estimate the average treatment effect by the Social Power Plus app, by using similar methodologies as the ones I used for enCompass and Social Power, with matched control groups.

8.3 Suggestions for future research and policy-making

Insights from the analyses of the three enCompass, Social Power and Social Power Plus app-based interventions that I performed for this dissertation suggest that behavioural effects stemming from use of persuasive gamified apps are only transient, and that individual level intervention techniques based on goal setting and monitoring of energy consumption at the household level are more effective in engaging app users. The fact that energy saving effects, which were found to take place and to be statistically significant, are short-lived and are not maintained over time, suggests that different intervention strategies need to be adopted, if the aim is to support a durable transition to a low-carbon and energy-sufficient society. At the light of the results of my work, no direct policy suggestions regarding how to scale-up the diffusion of persuasive, gamified apps such as the ones I analysed would in fact be sensible, neither from the energy and carbon viewpoint, nor from the cost-effectiveness viewpoint.

My suggestion for future research and policy-making is to keep exploring the social dimension and social interactions between households and individuals, by fitting use of persuasive apps within broader participatory processes, that explicitly aim at challenging and questioning the material and cultural contexts that support and shape individual action. This is for instance aligned with recommendations by Whitmarsh et al. (2021), who complain that in typical psychological approaches used to tackle climate change "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 creating 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, Nye, 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 open innovation that, through co-creation processes, operate as intermediaries among citizens, research organisations, companies, and government agencies, to create sustainable impact (see Section 6.1.1). 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, Rau, et al., 2021). To my 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. My suggestion would be to also integrate the latter, 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; then, meetings might be moved online, alternating in-person and online meetings. Apps might have a role into 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 in (nearly) real-time evolution of household consumption, and offer opportunities for social interaction via digital forums. Namely, 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 and communicated 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", according to the conceptualisation of social practices by Shove, Pantzar, 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 actively 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 urgency and need for collective solutions, thus being more open to implement and try them out.

As participation to living lab activities would be highly time and resource intensive, in terms of both intellectual and physical effort, also in this case voluntary, self-selected participants to these activities would undoubtedly be already highly motivated individuals —more motivated than the average population, close to the "converted" individuals I refer to in Section 7.6. However, it is likely that such individuals would be less interested in their own and specific energy consumption data than the case of Social Power Plus. In fact, apps, consumption feedback and monitoring, efficiency 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. Indeed, participants to living lab activities would be different from the majority of the population. Their higher motivation in this case would however turn into a strength, as, together with the broad networks of contacts by each living lab participant, 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) (Section 2.1). As noted by Raven, Van den Bosch, 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 abominable 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.

Together with a broad network of researchers in Switzerland, I will have opportunities to explore the potential of the living lab approach for the transition to an energy-sufficient, low-carbon society in a eight-year, Swiss-wide research project, that has just been funded by the Swiss Federal Office of energy (`https://sweet-lantern.ch`). Transition processes take time: by the end of 2025 the project will be able to report intermediate estimates of impact in key household-related domains, from energy consumption in households to mobility and leisure activities.

Bibliography

Aarts, Henk and Ap Dijksterhuis (2003). "The silence of the library: Environment, situational norm, and social behavior." In: *Journal of personality and social psychology* 84.1, p. 18.

Abraham, Charles and Susan Michie (2008). "A taxonomy of behavior change techniques used in interventions." In: *Health psychology* 27.3, p. 379.

Abrahamse, Wokje and Linda Steg (2013). "Social influence approaches to encourage resource conservation: A meta-analysis". In: *Global environmental change* 23.6, pp. 1773–1785.

Abrahamse, Wokje, Linda Steg, Charles Vlek, and Talib Rothengatter (2005). "A review of intervention studies aimed at household energy conservation". In: *Journal of Environmental Psychology* 25.3, pp. 273–291.

Abras, Chadia, Diane Maloney-Krichmar, Jenny Preece, et al. (2004). "User-centered design". In: *Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications* 37.4, pp. 445–456.

Adaji, Ifeoma and Mikhail Adisa (2022). "A Review of the Use of Persuasive Technologies to Influence Sustainable Behaviour". In: *Adjunct Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization*, pp. 317–325.

Agnisarman, Sruthy, Kapil Chalil Madathil, and Laura Stanley (2018). "A survey of empirical studies on persuasive technologies to promote sustainable living". In: *Sustainable Computing: Informatics and Systems* 19, pp. 112–122.

Ajzen, Icek (1991). "The theory of planned behavior". In: *Organizational behavior and human decision processes* 50.2, pp. 179–211.

Ajzen, Icek and Thomas J Madden (1986). "Prediction of goal-directed behavior: Attitudes, intentions, and perceived behavioral control". In: *Journal of experimental social psychology* 22.5, pp. 453–474.

Akenji, Lewis (2014). "Consumer scapegoatism and limits to green consumerism". In: *Journal of Cleaner Production* 63, pp. 13–23.

Akenji, Lewis, Magnus Bengtsson, Viivi Toivio, et al. (2021). *1.5-degree lifestyles: Towards a fair consumption space for all*. Tech. rep. Hot or Cool.

Alcott, Blake (2008). "The sufficiency strategy: Would rich-world frugality lower environmental impact?" In: *Ecological Economics* 64.4, pp. 770–786.

Allcott, Hunt (2011). "Social norms and energy conservation". In: *Journal of public Economics* 95.9-10, pp. 1082–1095.

Allcott, Hunt and Sendhil Mullainathan (2010). "Behavior and energy policy". In: *Science* 327.5970, pp. 1204–1205.

Allcott, Hunt and Todd Rogers (2014). “The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation”. In: *American Economic Review* 104.10, pp. 3003–37.

Almirall, Esteve, Melissa Lee, and Jonathan Wareham (2012). “Mapping living labs in the landscape of innovation methodologies”. In: *Technology innovation management review* 2.9.

Anagnostopoulou, Evangelia, Efthimios Bothos, Babis Magoutas, Johann Schrammel, and Gregoris Mentzas (2018). “Persuasive technologies for sustainable mobility: State of the art and emerging trends”. In: *Sustainability* 10.7, p. 2128.

Anderson, Kyle, Kwonsik Song, SangHyun Lee, et al. (2017). “Longitudinal analysis of normative energy use feedback on dormitory occupants”. In: *Applied energy* 189, pp. 623–639.

Andersson, Alfred, Lena Winslott Hiselius, and Emeli Adell (2018). “Promoting sustainable travel behaviour through the use of smartphone applications: A review and development of a conceptual model”. In: *Travel behaviour and society* 11, pp. 52–61.

Andor, Mark A and Katja M Fels (2018). “Behavioral economics and energy conservation — A systematic review of non-price interventions and their causal effects”. In: *Ecological Economics* 148, pp. 178–210.

Antin, Judd (2012). “Gamification is not a dirty word”. In: *interactions* 19.4, pp. 14–16.

Aparicio, Andrés Francisco, Francisco Luis Gutiérrez Vela, José Luis González Sánchez, and José Luis Isla Montes (2012). “Analysis and application of gamification”. In: *Proceedings of the 13th International Conference on Interacción Persona-Ordenador*, pp. 1–2.

Arimura, Toshi H, Shanjun Li, Richard G Newell, and Karen Palmer (2012). “Cost-effectiveness of electricity energy efficiency programs”. In: *The Energy Journal* 33.2.

Assmann, Susan F, Stuart J Pocock, Laura E Enos, and Linda E Kasten (2000). “Subgroup analysis and other (mis) uses of baseline data in clinical trials”. In: *The Lancet* 355.9209, pp. 1064–1069.

Avineri, Erel (2012). “On the use and potential of behavioural economics from the perspective of transport and climate change”. In: *Journal of Transport Geography* 24, pp. 512–521.

Ayres, Ian, Sophie Raseman, and Alice Shih (2013). “Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage”. In: *The Journal of Law, Economics, and Organization* 29.5, pp. 992–1022.

Bagozzi, Richard, Zynep Gurhan-Canli, and Joseph Priester (2002). *The social psychology of consumer behaviour*. McGraw-Hill Education (UK).

Bamberg, Sebastian (2013). “Changing environmentally harmful behaviors: A stage model of self-regulated behavioral change”. In: *Journal of Environmental Psychology* 34, pp. 151–159.

Bandura, Albert (1977). “Self-efficacy: Toward a unifying theory of behavioral change.” In: *Psychological review* 84.2, p. 191.

Bandura, Albert and Richard H Walters (1977). *Social learning theory*. Vol. 1. Prentice Hall: Englewood cliffs.

Baptista, Gonçalo and Tiago Oliveira (2019). “Gamification and serious games: A literature meta-analysis and integrative model”. In: *Computers in Human Behavior* 92, pp. 306–315.

Barrett, John, Steve Pye, Sam Betts-Davies, et al. (2022). “Energy demand reduction options for meeting national zero-emission targets in the United Kingdom”. In: *Nature Energy*, pp. 1–10.

Beck, Ariane L, Sohum Chitalia, and Varun Rai (2019). “Not so gameful: A critical review of gamification in mobile energy applications”. In: *Energy Research & Social Science* 51, pp. 32–39.

Becker, Lawrence J (1978). “Joint effect of feedback and goal setting on performance: A field study of residential energy conservation.” In: *Journal of applied psychology* 63.4, p. 428.

Beermann, Vincent, Annamina Rieder, and Falk Uebernickel (2022). “Green Nudges: How to Induce Pro-Environmental Behavior Using Technology”. In: *Proceedings of the International Conference on Information Systems (ICIS)*. Association for Information Systems (AIS).

Berger, Michelle, Theresa Lange, and Bastian Stahl (2022). “A digital push with real impact – Mapping effective digital nudging elements to contexts to promote environmentally sustainable behavior”. In: *Journal of Cleaner Production*, p. 134716. DOI: https://doi.org/10.1016/j.jclepro.2022.134716.

Bergvall-Kåreborn, Birgitta, Debra Howcroft, Anna Ståhlbröst, and Anita Melander Wikman (2010). “Participation in living lab: Designing systems with users”. In: *IFIP Working Conference on Human Benefit through the Diffusion of Information Systems Design Science Research*. Springer, pp. 317–326.

Bertho, Béatrice, Marlyne Sahakian, and Patrick Naef (2021). “The micro-politics of energy efficiency: An investigation of ‘eco-social interventions’ in western Switzerland”. In: *Critical Social Policy* 41.2, pp. 188–207.

Bertoldi, Paolo (2022). “Policies for energy conservation and sufficiency: Review of existing policies and recommendations for new and effective policies in OECD countries”. In: *Energy and Buildings*, p. 112075.

Best, Benjamin, Anja Bierwirth, Almut Bonhage, et al. (Apr. 2022). *Energiesparen als Schlüssel zur Energiesicherheit - Suffizienz als Strategie*. DOI: 10.5281/zenodo.6405939.

Bird, Stephen and Lisa Legault (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, Stephan and Lee Szwec (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*. Springer, pp. 631–640.

Boncu, Ștefan, Octav-Sorin Candel, and Nicoleta Laura Popa (2022). “Gameful Green: A Systematic Review on the Use of Serious Computer Games and Gamified Mobile Apps to Foster Pro-Environmental Information, Attitudes and Behaviors”. In: *Sustainability* 14.16, p. 10400.

Brancati, Dawn (2018). *Social Scientific Research*. Sage.

Breukers, Sylvia, Ruth Mourik, and DuneWorks BV Eindhoven (2015). “End-user engagement for flexible energy consumption patterns”. In: *Sustainable Consumption Transitions Series Issue 3*, p. 294.

Bruce, Steve and Steven Yearley (2006). *The Sage dictionary of sociology*. Sage.

Brynjarsdottir, Hronn, Maria Håkansson, James Pierce, et al. (2012). “Sustainably unpersuaded: How persuasion narrows our vision of sustainability”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, pp. 947–956.

Buchanan, Kathryn, Riccardo Russo, and Ben Anderson (2015). “The question of energy reduction: The problem (s) with feedback”. In: *Energy Policy* 77, pp. 89–96.

Bugden, Dylan and Richard Stedman (2019). "A synthetic view of acceptance and engagement with smart meters in the United States". In: *Energy Research & Social Science* 47, pp. 137–145.

Bundesamt für Energie (2021). *Schweizerische Elektrizitätsstatistik 2020*.

Burgess, Jacquelin and Michael Nye (2008). "Re-materialising energy use through transparent monitoring systems". In: *Energy Policy* 36.12, pp. 4454–4459.

Burke, Matthew J (2020). "Energy-sufficiency for a just transition: A systematic review". In: *Energies* 13.10, p. 2444.

Butler, Catherine, Karen Anne Parkhill, and Paulina Luzecka (2018). "Rethinking energy demand governance: Exploring impact beyond 'energy'policy". In: *Energy Research & Social Science* 36, pp. 70–78.

Byerly, Hilary, Andrew Balmford, Paul J Ferraro, et al. (2018). "Nudging pro-environmental behavior: Evidence and opportunities". In: *Frontiers in Ecology and the Environment* 16.3, pp. 159–168.

Campbell, Donald T and Julian C Stanley (2015). *Experimental and quasi-experimental designs for research*. Ravenio books.

Cartwright, Nancy (2007). "Are RCTs the gold standard?" In: *BioSocieties* 2.1, pp. 11–20.

– (2009). "Evidence-based policy: What's to be done about relevance?" In: *Philosophical Studies* 143.1, pp. 127–136.

Castells, Manuel (1996). *The Information age: Economy, Society and Culture volume one. The Network Society*. Blackwell, Oxford.

Cellina, Francesca, Dominik Bucher, Francesca Mangili, et al. (2019). "A large scale, app-based behaviour change experiment persuading sustainable mobility patterns: Methods, results and lessons learnt". In: *Sustainability* 11.9, p. 2674.

Cellina, Francesca, Dominik Bucher, José Veiga Simão, Roman Rudel, and Martin Raubal (2019). "Beyond limitations of current behaviour change apps for sustainable mobility: Insights from a user-centered design and evaluation process". In: *Sustainability* 11.8, p. 2281.

Cellina, Francesca, Roberta Castri, Mario Diethart, et al. (2018). "Constraints on upscaling and social inclusion in smart city living lab experiments and ways to anticipate them: Lessons from four "smarter" labs". In: *Open Living Lab Days 2018 Research and Innovation Conference Proceedings*, pp. 11–25.

Cellina, Francesca, Roberta Castri, José Veiga Simão, and Pasquale Granato (2020). "Co-creating app-based policy measures for mobility behavior change: A trigger for novel governance practices at the urban level". In: *Sustainable Cities and Society* 53, p. 101911.

Cellina, Francesca, Giuseppe Vittucci-Marzetti, and Marco Gui (2021). "Self-selection and attrition biases in app-based persuasive technologies for mobility behavior change: Evidence from a Swiss case study". In: *Computers in Human Behavior* 125, p. 106970.

Chatterjee, Samir and Alan Price (2009). "Healthy living with persuasive technologies: framework, issues, and challenges". In: *Journal of the American Medical Informatics Association* 16.2, pp. 171–178.

Chatzigeorgiou, IM and GT Andreou (2021). "A systematic review on feedback research for residential energy behavior change through mobile and web interfaces". In: *Renewable and Sustainable Energy Reviews* 135, p. 110187.

Chow, Yang-Wai, Willy Susilo, James G Phillips, Joonsang Baek, and Elena Vlahu-Gjorgievska (2017). “Video Games and Virtual Reality as Persuasive Technologies for Health Care: An Overview”. In: *J. Wirel. Mob. Networks Ubiquitous Comput. Dependable Appl.* 8.3, pp. 18–35.

Christensen, Robin, Martijn J. L. Bours, and Sabrina M. Nielsen (2021). “Effect modifiers and statistical tests for interaction in randomized trials”. In: *Journal of Clinical Epidemiology* 134, pp. 174–177.

Cialdini, Robert B and Noah J Goldstein (2004). “Social influence: Compliance and conformity”. In: *Annu. Rev. Psychol.* 55, pp. 591–621.

Cialdini, Robert B, Carl A Kallgren, and Raymond R Reno (1991). “A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior”. In: *Advances in experimental social psychology*. Vol. 24. Elsevier, pp. 201–234.

Cialdini, Robert B, Raymond R Reno, and Carl A Kallgren (1990). “A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places.” In: *Journal of personality and social psychology* 58.6, p. 1015.

Citherlet, Stephane and T Defaux (2007). “Energy and environmental comparison of three variants of a family house during its whole life span”. In: *Building and environment* 42.2, pp. 591–598.

Cohen, Jacob (2013). *Statistical power analysis for the behavioral sciences*. Routledge.

Creutzig, Felix, Joyashree Roy, William F Lamb, et al. (2018). “Towards demand-side solutions for mitigating climate change”. In: *Nature Climate Change* 8.4, pp. 260–263.

Croissant, Yves and Giovanni Millo (2008). “Panel Data Econometrics in R: The plm Package”. In: *Journal of Statistical Software* 27.2, pp. 1–43. DOI: 10.18637/jss.v027.i02.

Dalén, Anders and Jan Krämer (2017). “Towards a user-centered feedback design for smart meter interfaces to support efficient energy-use choices”. In: *Business & Information Systems Engineering* 59.5, pp. 361–373.

Daly, Herman E (1991). *Steady-state economics: with new essays*. Island press.

Darby, Sarah et al. (2006). “The effectiveness of feedback on energy consumption”. In: *A Review for DEFRA of the Literature on Metering, Billing and direct Displays* 486.2006, p. 26.

Darby, Sarah (2010). “Smart metering: What potential for householder engagement?” In: *Building research & information* 38.5, pp. 442–457.

Darby, Sarah and Tina Fawcett (2019). “Energy sufficiency in policy and practice: The question of needs and wants”. In: *eceee 2019 Summer Study Proceedings*.

De Young, Raymond (1996). “Some psychological aspects of reduced consumption behavior: The role of intrinsic satisfaction and competence motivation”. In: *Environment and Behavior* 28.3, pp. 358–409.

Deci, Edward L and Richard M Ryan (2004). *Handbook of Self-Determination Research*. University Rochester Press.

– (2008). “Self-determination theory: A macrotheory of human motivation, development, and health.” In: *Canadian psychology/Psychologie canadienne* 49.3, p. 182.

Dell’Era, Claudio and Paolo Landoni (2014). “Living Lab: A methodology between user-centred design and participatory design”. In: *Creativity and Innovation Management* 23.2, pp. 137–154.

Della Valle, Nives and Paolo Bertoldi (2021). "Mobilizing citizens to invest in energy efficiency". In: *EUR–Scientific and Technical Research Reports. Ispra: Joint Research Centre*.

– (2022). "Promoting energy efficiency: Barriers, societal needs and policies". In: *Frontiers in Energy Research* 9.

Delmas, Magali A, Miriam Fischlein, and Omar I Asensio (2013). "Information strategies and energy conservation behavior: A meta-analysis of experimental studies from 1975 to 2012". In: *Energy Policy* 61, pp. 729–739.

Deterding, Sebastian (2012). "Gamification: Designing for motivation". In: *interactions* 19.4, pp. 14–17.

Deterding, Sebastian, Dan Dixon, Rilla Khaled, and Lennart Nacke (2011). "From game design elements to gamefulness: defining 'gamification'". In: *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*, pp. 9–15.

Dolan, Paul, Michael Hallsworth, David Halpern, et al. (2012). "Influencing behaviour: The mindspace way". In: *Journal of economic psychology* 33.1, pp. 264–277.

Douglas, Benjamin D and Markus Brauer (2021). "Gamification to prevent climate change: A review of games and apps for sustainability". In: *Current Opinion in Psychology* 42, pp. 89–94.

Druckman, Angela, Mona Chitnis, Steve Sorrell, and Tim Jackson (2011). "Missing carbon reductions? Exploring rebound and backfire effects in UK households". In: *Energy Policy* 39.6, pp. 3572–3581.

Druckman, Angela and Tim Jackson (2009). "The carbon footprint of UK households 1990–2004: A socio-economically disaggregated, quasi-multi-regional input–output model". In: *Ecological Economics* 68.7, pp. 2066–2077.

– (2010). "The bare necessities: How much household carbon do we really need?" In: *Ecological Economics* 69.9, pp. 1794–1804.

Dubois, Ghislain, Benjamin Sovacool, Carlo Aall, et al. (2019). "It starts at home? Climate policies targeting household consumption and behavioral decisions are key to low-carbon futures". In: *Energy Research & Social Science* 52, pp. 144–158.

Dunlap, Riley E and Kent D Van Liere (1978). "The new environmental paradigm". In: *The journal of environmental education* 9.4, pp. 10–19.

Edenhofer, Ottmar (2015). *Climate change 2014: Mitigation of climate change*. Vol. 3. Cambridge University Press.

ElHaffar, Ghina, Fabien Durif, and Laurette Dubé (2020). "Towards closing the attitude-intention-behavior gap in green consumption: A narrative review of the literature and an overview of future research directions". In: *Journal of Cleaner Production* 275, p. 122556.

Eppler, Martin J and Jeanne Mengis (2008). "The concept of information overload-a review of literature from organization science, accounting, marketing, mis, and related disciplines (2004)". In: *Kommunikationsmanagement im Wandel*, pp. 271–305.

European Environment Agency (2020). *Final Energy Consumption by Sector and Fuel in Europe*. Tech. rep. EEA, Copenhagen, Denmark.

Evans, D Southerton, A D McMeekin, A Warde, and D Southerton (2012). “Sustainable consumption, behaviour change policies and theories of practice”. In: *The Habits of Consumption: COLLeGIUM: Studies across Disciplines in the Humanities and Social Sciences*. University of Helsinki, pp. 113–129.

Farrow, Katherine, Gilles Grolleau, and Lisette Ibanez (2017). “Social norms and pro-environmental behavior: A review of the evidence”. In: *Ecological Economics* 140, pp. 1–13.

Faruqui, Ahmad, Kevin Arritt, and Sanem Sergici (2017). “The impact of advanced metering infrastructure on energy conservation: A case study of two utilities”. In: *The Electricity Journal* 30.3, pp. 56–63.

Ferraro, Paul J, Juan Jose Miranda, and Michael K Price (2011). “The persistence of treatment effects with norm-based policy instruments: Evidence from a randomized environmental policy experiment”. In: *American Economic Review* 101.3, pp. 318–22.

Festinger, Leon (1954). “A theory of social comparison processes”. In: *Human relations* 7.2, pp. 117–140.

Figge, Frank, William Young, and Ralf Barkemeyer (2014). “Sufficiency or efficiency to achieve lower resource consumption and emissions? The role of the rebound effect”. In: *Journal of Cleaner Production* 69, pp. 216–224.

Fischer, Corinna (2008). “Feedback on household electricity consumption: A tool for saving energy?” In: *Energy Efficiency* 1.1, pp. 79–104.

Fogg, Brian J (1998). “Persuasive computers: Perspectives and research directions”. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 225–232.

– (2003). *Persuasive Technology: Using Computers to Change What We Think and Do*. Morgan Kaufmann.

– (2009). “A behavior model for persuasive design”. In: *Proceedings of the 4th international Conference on Persuasive Technology*, pp. 1–7.

Følstad, Asbjørn (2008). “LIVING LABS FOR INNOVATION AND DEVELOPMENT OF INFORMATION AND COMMUNICATION TECHNOLOGY: A LITERATURE REVIEW.” In: *eJOV: The Electronic Journal for Virtual Organization & Networks* 10.

Forgas, Joseph P, Kipling D Williams, Kipling D Williams, et al. (2001). *Social Influence: Direct and Indirect Processes*. Vol. 3. Psychology Press.

Foster, Derek, Conor Linehan, Shaun Lawson, and Ben Kirman (2011). “Power ballads: Deploying aversive energy feedback in social media”. In: *CHI’11 Extended Abstracts on Human Factors in Computing Systems*. ACM, pp. 2221–2226.

Fraternali, Piero, Francesca Cellina, Sergio Luis Herrera Gonzales, et al. (2019). “Visualizing and gamifying consumption data for resource saving: Challenges, lessons learnt and a research agenda for the future”. In: *Energy Informatics* 2.1, pp. 1–13.

Frederiks, Elisha R, Karen Stenner, and Elizabeth V Hobman (2015a). “Household energy use: Applying behavioural economics to understand consumer decision-making and behaviour”. In: *Renewable and Sustainable Energy Reviews* 41, pp. 1385–1394.

– (2015b). “The socio-demographic and psychological predictors of residential energy consumption: A comprehensive review”. In: *Energies* 8.1, pp. 573–609.

Frederiks, Elisha R, Karen Stenner, Elizabeth V Hobman, and Mark Fischle (2016). “Evaluating energy behavior change programs using randomized controlled trials: Best practice guidelines for policymakers”. In: *Energy Research & Social Science* 22, pp. 147–164.

Frey, Bruno S and Reto Jegen (2001). “Motivation crowding theory”. In: *Journal of economic surveys* 15.5, pp. 589–611.

Friedrich, Katherine, Maggie Eldridge, Dan York, Patti Witte, and Marty Kushler (2009). “Saving energy cost-effectively: A national review of the cost of energy saved through utility-sector energy efficiency programs”. In: *American Council for an Energy-Efficient Economy. September*.

Froehlich, Jon (2015). “The Gameful World: Approaches, Issues, Applications”. In: ed. by Steffen P. Walz and Sebastian Deterding. MIT Press. Chap. Gamifying Green: Gamification and Environmental Sustainability, pp. 563–596.

Froehlich, Jon, Leah Findlater, and James Landay (2010). “The design of eco-feedback technology”. In: *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 1999–2008.

Fuchs, Doris, Marlyne Sahakian, Tobias Gumbert, et al. (2021). *Consumption corridors: Living a good life within sustainable limits*. Routledge.

Fuenfschilling, Lea, Niki Frantzeskaki, and Lars Coenen (2019). “Urban experimentation & sustainability transitions”. In: *European Planning Studies* 27.2, pp. 219–228.

Gardner, Gerald T and Paul C Stern (2008). “The short list: The most effective actions US households can take to curb climate change”. In: *Environment: science and policy for sustainable development* 50.5, pp. 12–25.

Geelen, Daphne, Ruth Mugge, Sacha Silvester, and Annemieke Bulters (2019). “The use of apps to promote energy saving: A study of smart meter–related feedback in the Netherlands”. In: *Energy Efficiency* 12.6, pp. 1635–1660.

Geels, Frank (2004). “From sectoral systems of innovation to socio-technical systems: Insights about dynamics and change from sociology and institutional theory”. In: *Research policy* 33.6-7, pp. 897–920.

Geels, Frank and Johan Schot (2007). “Typology of sociotechnical transition pathways”. In: *Research policy* 36.3, pp. 399–417.

Geels, Frank, Benjamin Sovacool, Tim Schwanen, and Steve Sorrell (2017). “Sociotechnical transitions for deep decarbonization”. In: *Science* 357.6357, pp. 1242–1244.

Geller, E Scott, Thomas D Berry, Timothy D Ludwig, et al. (1990). “A conceptual framework for developing and evaluating behavior change interventions for injury control”. In: *Health Education Research* 5.2, pp. 125–137.

Georgescu-Roegen, Nicholas (1975). “Energy and Economic Myths”. In: *Southern Economic Journal* 41.3, pp. 347–381.

Gerber, Alan S and Donald P Green (2012). *Field experiments: Design, analysis, and interpretation*. WW Norton.

Gertler, Paul J, Sebastian Martinez, Patrick Premand, Laura B Rawlings, and Christel MJ Vermeersch (2016). *Impact evaluation in practice*. The World Bank.

Giddens, Anthony (1984). *The constitution of society: Outline of the theory of structuration*. University of California Press.

Gillingham, Kenneth, David Rapson, and Gernot Wagner (2020). "The rebound effect and energy efficiency policy". In: *Review of Environmental Economics and Policy*.

Godin, Laurence, Senja Laakso, and Marlyne Sahakian (2020). "Doing laundry in consumption corridors: Wellbeing and everyday life". In: *Sustainability: Science, Practice and Policy* 16.1, pp. 99–113.

Goldemberg, José, Thomas B Johansson, Amulya KN Reddy, and Robert H Williams (1985). "Basic needs and much more with one kilowatt per capita". In: *Ambio*, pp. 190–200.

Goldthorpe, John H (2001). "Causation, statistics, and sociology". In: *European sociological review* 17.1, pp. 1–20.

Gollwitzer, Peter M (1990). "Action phases and mind-sets". In: *Handbook of motivation and cognition: Foundations of social behavior* 2, pp. 53–92.

Goodman, James (2022). "Social Movements and Climate Change: 'Climatizing' Society From Within". In: *Oxford Research Encyclopedia of Climate Science*.

Gui, Marco, Marco Fasoli, and Roberto Carradore (2017). "Digital well-being. Developing a new theoretical tool for media literacy research". In: *Italian Journal of Sociology of Education* 9.1.

Gui, Marco, James Shanahan, and Mina Tsay-Vogel (2021). "Theorizing inconsistent media selection in the digital environment". In: *The Information Society* 37.4, pp. 247–261.

Hajer, Maarten, Måns Nilsson, Kate Raworth, et al. (2015). "Beyond cockpit-ism: Four insights to enhance the transformative potential of the sustainable development goals". In: *Sustainability* 7.2, pp. 1651–1660.

Hamari, Juho (2019). "Gamification". In: *The Blackwell Encyclopedia of Sociology*. John Wiley Sons, Ltd, pp. 1–3. DOI: `https://doi.org/10.1002/9781405165518.wbeos1321`.

Hamari, Juho, Jonna Koivisto, and Tuomas Pakkanen (2014). "Do persuasive technologies persuade? A review of empirical studies". In: *International conference on persuasive technology*. Springer, pp. 118–136.

Hamari, Juho, Jonna Koivisto, and Harri Sarsa (2014). "Does gamification work? A literature review of empirical studies on gamification". In: *2014 47th Hawaii international conference on system sciences*. Ieee, pp. 3025–3034.

Hampton, Sam and Rob Adams (2018). "Behavioural economics vs social practice theory: Perspectives from inside the United Kingdom government". In: *Energy Research & Social Science* 46, pp. 214–224.

Hand, Martin, Elizabeth Shove, and Dale Southerton (2005). "Explaining showering: A discussion of the material, conventional, and temporal dimensions of practice". In: *Sociological Research Online* 10.2, pp. 101–113.

Hansen, Pelle Guldborg (2016). "The definition of nudge and libertarian paternalism: Does the hand fit the glove?" In: *European Journal of Risk Regulation* 7.1, pp. 155–174.

Hansen, Pelle Guldborg and Andreas Maaløe Jespersen (2013). "Nudge and the manipulation of choice: A framework for the responsible use of the nudge approach to behaviour change in public policy". In: *European Journal of Risk Regulation* 4.1, pp. 3–28.

Hargreaves, Tom (2011). "Practice-ing behaviour change: Applying social practice theory to pro-environmental behaviour change". In: *Journal of Consumer Culture* 11.1, pp. 79–99.

– (2018). "Beyond energy feedback". In: *Building Research & Information* 46.3, pp. 332–342.

Hargreaves, Tom and Lucie Middlemiss (2020). “The importance of social relations in shaping energy demand”. In: *Nature Energy* 5.3, pp. 195–201.

Hargreaves, Tom, Michael Nye, and Jacquelin Burgess (2010). “Making energy visible: A qualitative field study of how householders interact with feedback from smart energy monitors”. In: *Energy Policy* 38.10, pp. 6111–6119.

– (2013). “Keeping energy visible? Exploring how householders interact with feedback from smart energy monitors in the longer term”. In: *Energy Policy* 52, pp. 126–134.

Hartigan, John A and Manchek A Wong (1979). “Algorithm AS 136: A k-means clustering algorithm”. In: *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28.1, pp. 100–108.

Hartman, Raymon S (1988). “Self-selection bias in the evolution of voluntary energy conservation programs”. In: *The Review of Economics and Statistics*, pp. 448–458.

Hausman, Daniel M and Brynn Welch (2010). “Debate: To nudge or not to nudge”. In: *Journal of Political Philosophy* 18.1, pp. 123–136.

He, Helen Ai, Saul Greenberg, and Elaine M Huang (2010). “One size does not fit all: applying the transtheoretical model to energy feedback technology design”. In: *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, pp. 927–936.

Heckhausen, Heinz and Peter M Gollwitzer (1987). “Thought contents and cognitive functioning in motivational versus volitional states of mind”. In: *Motivation and Emotion* 11.2, pp. 101–120.

Hedin, Björn, Cecilia Katzeff, Elina Eriksson, and Daniel Pargman (2019). “A systematic review of digital behaviour change interventions for more sustainable food consumption”. In: *Sustainability* 11.9, p. 2638.

Heiskanen, Eva, Senja Laakso, Kaisa Matschoss, et al. (2018). “Designing real-world laboratories for the reduction of residential energy use: Articulating theories of change”. In: *GAIA-Ecological Perspectives for Science and Society* 27.1, pp. 60–67.

Henrich, Joseph, Steven J Heine, and Ara Norenzayan (2010). “Most people are not WEIRD”. In: *Nature* 466.7302, pp. 29–29.

Hermsen, Sander, Jeana Frost, Reint Jan Renes, and Peter Kerkhof (2016). “Using feedback through digital technology to disrupt and change habitual behavior: A critical review of current literature”. In: *Computers in Human Behavior* 57, pp. 61–74.

Hess, Ann-Kathrin, Robin Samuel, and Paul Burger (2018). “Informing a social practice theory framework with social-psychological factors for analyzing routinized energy consumption: A multivariate analysis of three practices”. In: *Energy Research & Social Science* 46, pp. 183–193.

Hess, Ann-Kathrin, Iljana Schubert, Robin Samuel, and Paul Burger (2022). “Changing routinized household energy consumption using the example of washing, cooking, and standby: A randomized controlled field experiment of home energy advice”. In: *Cleaner and Responsible Consumption*, p. 100052.

Hillgren, Per-Anders (2013). *PARTICIPATORY DESIGN FOR SOCIAL AND PUBLIC INNOVATION: LIVING LABS AS SPACES OF AGONISTIC EXPERIMENTS AND FRIENDLY HACKING*. Tech. rep. DESIS.

Holland, Paul W (1986). “Statistics and causal inference”. In: *Journal of the American Statistical Association* 81.396, pp. 945–960.

Huber, Martina Z and Lorenz M Hilty (2015). "Gamification and sustainable consumption: overcoming the limitations of persuasive technologies". In: *ICT innovations for sustainability*. Springer, pp. 367–385.

Huber, Patrick, Alberto Calatroni, Andreas Rumsch, and Andrew Paice (2021). "Review on deep neural networks applied to low-frequency nilm". In: *Energies* 14.9, p. 2390.

Icha, Petra and Gunter Kuhs (2021). *Entwicklung der spezifischen Kohlendioxid-Emissionen des deutschen Strommix 1990-2020*.

IEA (2021a). *Electricity Information: Overview*. Tech. rep. Accessed: 2021-08-06. International Energy Agency (IEA). URL: `https://www.iea.org/reports/electricity-information-overview`.

– (2021b). *Key World Energy Statistics*. Tech. rep. Accessed: 2021-08-06. Paris: International Energy Agency (IEA). URL: `https://www.iea.org/reports/key-world-energy-statistics-2021`.

– (2022). *World Energy Outlook 2022*. Tech. rep. Paris: International Energy Agency (IEA). URL: `https://www.iea.org/reports/world-energy-outlook-2022`.

Ivanova, Diana, John Barrett, Dominik Wiedenhofer, et al. (2020). "Quantifying the potential for climate change mitigation of consumption options". In: *Environmental Research Letters* 15.9, p. 093001.

Jackson, Michelle and David R Cox (2013). "The principles of experimental design and their application in sociology". In: *Annual Review of Sociology* 39, pp. 27–49.

Jackson, Tim (2005). "Motivating sustainable consumption". In: *Sustainable Development Research Network* 29.1, pp. 30–40.

– (2009). *Prosperity without growth: Economics for a finite planet*. Routledge.

Jalas, Mikko, Sampsa Hyysalo, Eva Heiskanen, et al. (2017). "Everyday experimentation in energy transition: A practice-theoretical view". In: *Journal of Cleaner Production* 169, pp. 77–84.

Jensen, Charlotte Louise, Gary Goggins, Inge Røpke, and Frances Fahy (2019). "Achieving sustainability transitions in residential energy use across Europe: The importance of problem framings". In: *Energy Policy* 133, p. 110927.

Joachain, Hélène and Frédéric Klopfert (2014). "Smarter than metering? Coupling smart meters and complementary currencies to reinforce the motivation of households for energy savings". In: *Ecological Economics* 105, pp. 89–96.

John, Peter, Graham Smith, and Gerry Stoker (2009). "Nudge nudge, think think: Two strategies for changing civic behaviour". In: *The Political Quarterly* 80.3, pp. 361–370.

Johnson, Daniel, Ella Horton, Rory Mulcahy, and Marcus Foth (2017). "Gamification and serious games within the domain of domestic energy consumption: A systematic review". In: *Renewable and Sustainable Energy Reviews* 73, pp. 249–264.

Jonsson, Daniel K, Stina Gustafsson, Josefin Wangel, et al. (2011). "Energy at your service: highlighting energy usage systems in the context of energy efficiency analysis". In: *Energy Efficiency* 4.3, pp. 355–369.

Jungell-Michelsson, Jessica and Pasi Heikkurinen (2022). "Sufficiency: A systematic literature review". In: *Ecological Economics* 195, p. 107380.

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

Kapp, Karl M (2013). *The gamification of learning and instruction fieldbook: Ideas into practice*. John Wiley & Sons.

Karlin, Beth, Joanne F Zinger, and Rebecca Ford (2015). “The effects of feedback on energy conservation: A meta-analysis.” In: *Psychological bulletin* 141.6, p. 1205.

Kassambara, Alboukadel and Fabian Mundt (2020). *factoextra: Extract and Visualize the Results of Multivariate Data Analyses*. R package version 1.0.7. URL: `https://CRAN.R-project.org/package=factoextra`.

Kaufman, Stefan, Alexander Saeri, Rob Raven, Shirin Malekpour, and Liam Smith (2021). “Behaviour in sustainability transitions: A mixed methods literature review”. In: *Environmental Innovation and Societal Transitions* 40, pp. 586–608.

Kemmler, A and T Spillmann (2020). *Energieverbrauch der Privaten Haushalte 2000–2020*. Tech. rep. Swiss Federal Office of Energy (SFOE).

– (2021). *Der Energieverbrauch der Privaten Haushalte 2000–2020: Ex-Post-Analyse nach Verwendungszwecken und Ursachen der Veränderungen*. Tech. rep. Swiss Federal Office of Energy (SFOE).

Kemp, René, Johan Schot, and Remco Hoogma (1998). “Regime shifts to sustainability through processes of niche formation: the approach of strategic niche management”. In: *Technology analysis & strategic management* 10.2, pp. 175–198.

Khanna, Tarun M, Giovanni Baiocchi, Max Callaghan, et al. (2021). “A multi-country meta-analysis on the role of behavioural change in reducing energy consumption and CO_2 emissions in residential buildings”. In: *Nature Energy* 6.9, pp. 925–932.

Kientz, Julie A, Eun Kyoung Choe, Brennen Birch, et al. (2010). “Heuristic evaluation of persuasive health technologies”. In: *Proceedings of the 1st ACM International Health Informatics Symposium*, pp. 555–564.

Kiesler, Charles A (1971). *The psychology of commitment: Experiments linking behavior to belief*. Academic Press.

Kimura, Hiroaki and Tatsuo Nakajima (2011). “Designing Persuasive Applications to Motivate Sustainable Behavior in Collectivist Cultures”. In: *PsychNology Journal* 9.1.

Klock, Ana Carolina Tomé, Isabela Gasparini, Marcelo Soares Pimenta, and Juho Hamari (2020). “Tailored gamification: A review of literature”. In: *International Journal of Human-Computer Studies* 144, p. 102495.

Kluger, Avraham N and Angelo DeNisi (1996). “The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory.” In: *Psychological bulletin* 119.2, p. 254.

Kollmuss, Anja and Julian Agyeman (2002). “Mind the gap: Why do people act environmentally and what are the barriers to pro-environmental behavior?” In: *Environmental Education Research* 8.3, pp. 239–260.

Koroleva, Ksenia, Mark Melenhorst, Jasminko Novak, et al. (2019). “Designing an integrated socio-technical behaviour change system for energy saving”. In: *Energy Informatics* 2.1, pp. 1–20.

Krath, Jeanine, Linda Schürmann, and Harald FO Von Korflesch (2021). “Revealing the theoretical basis of gamification: A systematic review and analysis of theory in research on gamification, serious games and game-based learning”. In: *Computers in Human Behavior* 125, p. 106963.

Krebs, Luana and Rolf Frischknecht (2021). *Umweltbilanz Strommixe Schweiz 2018*.

Kuijer, Lenneke and Conny Bakker (2015). “Of chalk and cheese: Behaviour change and practice theory in sustainable design”. In: *International Journal of Sustainable Engineering* 8.3, pp. 219–230.

Kurz, Tim, Benjamin Gardner, Bas Verplanken, and Charles Abraham (2015). “Habitual behaviors or patterns of practice? Explaining and changing repetitive climate-relevant actions”. In: *Wiley Interdisciplinary Reviews: Climate Change* 6.1, pp. 113–128.

Labanca, Nicola and Paolo Bertoldi (2018). “Beyond energy efficiency and individual behaviours: Policy insights from social practice theories”. In: *Energy Policy* 115, pp. 494–502.

Labanca, Nicola, Ângela Guimarães Pereira, Matt Watson, et al. (2020). “Transforming innovation for decarbonisation? Insights from combining complex systems and social practice perspectives”. In: *Energy Research & Social Science* 65, p. 101452.

Lapinski, Maria Knight and Rajiv N Rimal (2005). “An explication of social norms”. In: *Communication theory* 15.2, pp. 127–147.

Lawson, Tony and Joan Garrod (2012). *Dictionary of sociology*. Routledge.

Lehner, Matthias, Oksana Mont, and Eva Heiskanen (2016). “Nudging — A promising tool for sustainable consumption behaviour?” In: *Journal of Cleaner Production* 134, pp. 166–177.

Leminen, Seppo (2013). “Coordination and participation in living lab networks”. In: *Technology Innovation Management Review* 3.11.

Li, Ian, Anind K Dey, and Jodi Forlizzi (2011). “Understanding my data, myself: Supporting self-reflection with Ubicomp technologies”. In: *Proceedings of the 13th international conference on Ubiquitous computing*, pp. 405–414.

Lorek, Sylvia and Doris Fuchs (2013). “Strong sustainable consumption governance — Precondition for a degrowth path?” In: *Journal of Cleaner Production* 38, pp. 36–43.

Lorek, Sylvia and Joachim H Spangenberg (2014). “Sustainable consumption within a sustainable economy — Beyond green growth and green economies”. In: *Journal of Cleaner Production* 63, pp. 33–44.

– (2019). “Energy sufficiency through social innovation in housing”. In: *Energy Policy* 126, pp. 287–294.

Löschel, Andreas, Matthias Rodemeier, and Madeline Werthschulte (2020). *When nudges fail to scale: Field experimental evidence from goal setting on mobile phones*. Tech. rep. CESifo working paper.

Maciejewski, Matthew L (2020). “Quasi-experimental design”. In: *Biostatistics & Epidemiology* 4.1, pp. 38–47.

Maechler, Martin, Peter Rousseeuw, Anja Struyf, Mia Hubert, and Kurt Hornik (2021). *cluster: Cluster Analysis Basics and Extensions*. R package version 2.1.2. URL: `https://CRAN.R-project.org/package=cluster`.

Maréchal, Kevin (2010). “Not irrational but habitual: The importance of ‘behavioural lock-in’ in energy consumption”. In: *Ecological Economics* 69.5, pp. 1104–1114.

Markard, Jochen, Frank Geels, and Rob Raven (2020). “Challenges in the acceleration of sustainability transitions”. In: *Environmental Research Letters* 15.8, p. 081001.

Matschoss, Kaisa, Frances Fahy, Henrike Rau, et al. (2021). “Challenging practices: Experiences from community and individual living lab approaches”. In: *Sustainability: Science, Practice and Policy* 17.1, pp. 135–151.

Maxim, Alexandru, Costică Mihai, Constantin-Marius Apostoaie, et al. (2016). “Implications and measurement of energy poverty across the European Union”. In: *Sustainability* 8.5, p. 483.

McClelland, Lou and Stuart W Cook (1979). “Energy conservation effects of continuous in-home feedback in all-electric homes”. In: *Journal of Environmental Systems* 9.2.

McIntosh, Colin (2013). *Cambridge advanced learner’s dictionary*. Cambridge University Press.

McKenzie-Mohr, Doug (2000a). “Fostering sustainable behavior through community-based social marketing.” In: *American Psychologist* 55.5, p. 531.

– (2000b). “Promoting sustainable behavior: An introduction to community-based social marketing”. In: *Journal of Social Issues* 56.3, pp. 543–554.

Michie, Susan, Rachel N Carey, Marie Johnston, et al. (2018). “From theory-inspired to theory-based interventions: A protocol for developing and testing a methodology for linking behaviour change techniques to theoretical mechanisms of action”. In: *Annals of behavioral medicine* 52.6, pp. 501–512.

Michie, Susan and Andrew Prestwich (2010). “Are interventions theory-based? Development of a theory coding scheme.” In: *Health psychology* 29.1, p. 1.

Michie, Susan, Maartje M Van Stralen, and Robert West (2011). “The behaviour change wheel: A new method for characterising and designing behaviour change interventions”. In: *Implementation science* 6.1, pp. 1–12.

Milgram, Stanley (1970). “The Experience of Living in Cities”. In: *Science* 167.3924, pp. 1461–1468. DOI: 10.1126/science.167.3924.1461.

Millo, Giovanni (2017). “Robust standard error estimators for panel models: A unifying approach”. In: *Journal of Statistical Software* 82, pp. 1–27.

Millward-Hopkins, Joel, Julia K Steinberger, Narasimha D Rao, and Yannick Oswald (2020). “Providing decent living with minimum energy: A global scenario”. In: *Global Environmental Change* 65, p. 102168.

Mogles, Nataliya, Ian Walker, Alfonso P Ramallo-González, et al. (2017). “How smart do smart meters need to be?” In: *Building and Environment* 125, pp. 439–450.

Moloney, Susie and Yolande Strengers (2014). “‘Going green’?: The limitations of behaviour change programmes as a policy response to escalating resource consumption”. In: *Environmental Policy and Governance* 24.2, pp. 94–107.

Mols, Frank, S Alexander Haslam, Jolanda Jetten, and Niklas K Steffens (2015). “Why a nudge is not enough: A social identity critique of governance by stealth”. In: *European Journal of Political Research* 54.1, pp. 81–98.

Moran, Daniel, Richard Wood, Edgar Hertwich, et al. (2020). “Quantifying the potential for consumer-oriented policy to reduce European and foreign carbon emissions”. In: *Climate Policy* 20.sup1, S28–S38.

Morganti, Luca, Federica Pallavicini, Elena Cadel, et al. (2017). “Gaming for Earth: Serious games and gamification to engage consumers in pro-environmental behaviours for energy efficiency”. In: *Energy Research & Social Science* 29, pp. 95–102.

Morrison, Leanne G (2015). “Theory-based strategies for enhancing the impact and usage of digital health behaviour change interventions: A review”. In: *Digital Health* 1, p. 2055207615595335.

Moser, Corinne, Andreas Rösch, and Michael Stauffacher (2015). “Exploring societal preferences for energy sufficiency measures in Switzerland”. In: *Frontiers in Energy Research* 3, p. 40.

Muller, Adrian (2009). “Sufficiency — Does energy consumption become a moral issue?” In: *IOP Conference Series: Earth and Environmental Science (EES)*. Vol. 6. 26.

Murtagh, Niamh, Birgitta Gatersleben, and David Uzzell (2014). “20 60 20-Differences in energy behaviour and conservation between and within households with electricity monitors”. In: *PloS one* 9.3, e92019.

Nacke, Lennart E and Sebastian Deterding (2017). “The maturing of gamification research”. In: *Computers in Human Behaviour*, pp. 450–454.

Nash, Nick, Lorraine Whitmarsh, Stuart Capstick, et al. (2017). “Climate-relevant behavioral spillover and the potential contribution of social practice theory”. In: *Wiley Interdisciplinary Reviews: Climate Change* 8.6, e481.

Naus, Joeri, Bas JM van Vliet, and Astrid Hendriksen (2015). “Households as change agents in a Dutch smart energy transition: On power, privacy and participation”. In: *Energy Research & Social Science* 9, pp. 125–136.

Newzoo (2020). *Top Countries/Markets by Smartphone Penetration and Users*.

Nicholson, Scott (2015). “A recipe for meaningful gamification”. In: *Gamification in education and business*. Springer, pp. 1–20.

Nielsen, Kristian S., Susan Clayton, Paul C. Stern, et al. (2021). “How psychology can help limit climate change.” In: *American Psychologist* 76.1, p. 130.

Nielsen, Kristian S., Viktoria Cologna, Florian Lange, Cameron Brick, and Paul C. Stern (2021). “The case for impact-focused environmental psychology”. In: *Journal of Environmental Psychology* 74, p. 101559. DOI: `https://doi.org/10.1016/j.jenvp.2021.101559`.

Nielsen, Kristian S., Sander van der Linden, and Paul C. Stern (2020). “How behavioral interventions can reduce the climate impact of energy use”. In: *Joule* 4.8, pp. 1613–1616.

Nilsson, Andreas, Cecilia Jakobsson Bergstad, Liane Thuvander, et al. (2014). “Effects of continuous feedback on households’ electricity consumption: Potentials and barriers”. In: *Applied Energy* 122, pp. 17–23.

Nisa, Claudia F, Jocelyn J Bélanger, Birga M Schumpe, and Daiane G Faller (2019). “Meta-analysis of randomised controlled trials testing behavioural interventions to promote household action on climate change”. In: *Nature Communications* 10.1, pp. 1–13.

Norman, D.A. (2013). *The Design of Everyday Things*. The MIT Press. MIT Press.

Notter, Dominic A, Reto Meyer, and Hans-Jörg Althaus (2013). “The Western lifestyle and its long way to sustainability”. In: *Environmental Science & Technology* 47.9, pp. 4014–4021.

O’Rourke, Dara and Niklas Lollo (2015). “Transforming consumption: from decoupling, to behavior change, to system changes for sustainable consumption”. In: *Annual Review of Environment and Resources* 40, pp. 233–259.

O’Shaughnessy, Eric, Monisha Shah, David Parra, and Kristen Ardani (2022). “The demand-side resource opportunity for deep grid decarbonization”. In: *Joule*.

Ohnmacht, Timo, Dorothea Schaffner, Christian Weibel, and Helmut Schad (2017). “Rethinking social psychology and intervention design: A model of energy savings and human behavior”. In: *Energy Research & Social Science* 26, pp. 40–53.

Oinas-Kukkonen, Harri (2010). “Behavior change support systems: A research model and agenda”. In: *International Conference on Persuasive Technology*. Springer, pp. 4–14.

– (2013). “A foundation for the study of behavior change support systems”. In: *Personal and ubiquitous computing* 17.6, pp. 1223–1235.

Oinas-Kukkonen, Harri and Marja Harjumaa (2008). “A systematic framework for designing and evaluating persuasive systems”. In: *International conference on persuasive technology*. Springer, pp. 164–176.

– (2009). “Persuasive systems design: Key issues, process model, and system features”. In: *Communications of the association for Information Systems* 24.1, p. 28.

Orji, Rita and Karyn Moffatt (2018). “Persuasive technology for health and wellness: State-of-the-art and emerging trends”. In: *Health Informatics Journal* 24.1, pp. 66–91.

Osbaldiston, Richard and John Paul Schott (2012). “Environmental sustainability and behavioral science: Meta-analysis of proenvironmental behavior experiments”. In: *Environment and Behavior* 44.2, pp. 257–299.

Otto, Siegmar, Florian G Kaiser, and Oliver Arnold (2014). “The critical challenge of climate change for psychology: Preventing rebound and promoting more individual irrationality.” In: *European Psychologist* 19.2, p. 96.

Pallot, Marc, Brigitte Trousse, Bernard Senach, and Dominique Scapin (2010). “Living lab research landscape: From user centred design and user experience towards user cocreation”. In: *First European Summer School" Living Labs"*.

Peattie, Ken (2010). “Green consumption: behavior and norms”. In: *Annual review of Environment and Resources* 35.1, pp. 195–228.

Perski, Olga, Ann Blandford, Robert West, and Susan Michie (2017). “Conceptualising engagement with digital behaviour change interventions: A systematic review using principles from critical interpretive synthesis”. In: *Translational Behavioral Medicine* 7.2, pp. 254–267.

Pierce, James, Chloe Fan, Derek Lomas, Gabriela Marcu, and Eric Paulos (2010). “Some consideration on the (in) effectiveness of residential energy feedback systems”. In: *Proceedings of the 8th ACM conference on designing interactive systems*, pp. 244–247.

Pierce, James, Diane J Schiano, and Eric Paulos (2010). “Home, habits, and energy: Examining domestic interactions and energy consumption”. In: *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 1985–1994.

Pindyck, Robert S (2019). “The social cost of carbon revisited”. In: *Journal of Environmental Economics and Management* 94, pp. 140–160.

Pörtner, Hans-Otto, Debra C Roberts, H Adams, et al. (2022). *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press. DOI: 10.1017/9781009325844.

Princen, Thomas (2005). *The logic of sufficiency*. Mit Press Cambridge, MA.

Prochaska, James O and Wayne F Velicer (1997). “The transtheoretical model of health behavior change”. In: *American Journal of Health Promotion* 12.1, pp. 38–48.

Puntiroli, Michael and Valéry Bezençon (2020). “Feedback devices help only environmentally concerned people act pro-environmentally over time”. In: *Journal of Environmental Psychology* 70, p. 101459.

R Core Team (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: `https://www.R-project.org/`.

Rapp, Amon, Frank Hopfgartner, Juho Hamari, Conor Linehan, and Federica Cena (2019). “Strengthening gamification studies: Current trends and future opportunities of gamification research”. In: *International Journal of Human-Computer Studies* 127, pp. 1–6.

Raven, Rob, Eric Jolivet, Ruth M Mourik, and Ynke CFJ Feenstra (2009). “ESTEEM: Managing societal acceptance in new energy projects: A toolbox method for project managers”. In: *Technological Forecasting and Social Change* 76.7, pp. 963–977.

Raven, Rob, David Reynolds, Ruth Lane, et al. (2021). “Households in sustainability transitions: A systematic review and new research avenues”. In: *Environmental Innovation and Societal Transitions* 40, pp. 87–107.

Raven, Rob, Suzanne Van den Bosch, and Rob Weterings (2010). “Transitions and strategic niche management: Towards a competence kit for practitioners”. In: *International Journal of Technology Management* 51.1, pp. 57–74.

Raworth, Kate (2012). *A safe and just space for humanity: Can we live within the doughnut?* Oxfam.

Reckwitz, Andreas (2002). “Toward a theory of social practices: A development in culturalist theorizing”. In: *European Journal of Social Theory* 5.2, pp. 243–263.

Reed, Mark S, Anna C Evely, Georgina Cundill, et al. (2010). “What is social learning?” In: *Ecology and Society* 15.4.

Reno, Raymond R, Robert B Cialdini, and Carl A Kallgren (1993). “The transsituational influence of social norms.” In: *Journal of personality and social psychology* 64.1, p. 104.

Roberts, J Timmons and Bradley Parks (2006). *A climate of injustice: Global inequality, North-South politics, and climate policy*. Cambridge University Press.

Roberts, J Timmons, Julia K Steinberger, Thomas Dietz, et al. (2020). “Four agendas for research and policy on emissions mitigation and well-being”. In: *Global Sustainability* 3.

Rockström, Johan, Will Steffen, Kevin Noone, et al. (2009). “Planetary boundaries: Exploring the safe operating space for humanity”. In: *Ecology and Society* 14.2.

Røpke, Inge (2009). “Theories of practice — New inspiration for ecological economic studies on consumption”. In: *Ecological Economics* 68.10, pp. 2490–2497.

Royston, Sarah, Jan Selby, and Elizabeth Shove (2018). “Invisible energy policies: A new agenda for energy demand reduction”. In: *Energy Policy* 123, pp. 127–135.

Royston, Sarah, Elizabeth Shove, and Jan Selby (2017). “Invisible energy policy: A new agenda for research and intervention”. In: *Proceedings of the European Council for an Energy Efficient Economy Summer Study*.

RStudio Team (2022). *RStudio: Integrated Development Environment for R*. RStudio, PBC. Boston, MA. URL: `http://www.rstudio.com/`.

Ryan, Richard M and Edward L Deci (2000a). “Intrinsic and extrinsic motivations: Classic definitions and new directions”. In: *Contemporary educational psychology* 25.1, pp. 54–67.

– (2000b). “Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being.” In: *American psychologist* 55.1, p. 68.

Sachs, Jeffrey, Christian Kroll, Guillame Lafortune, Grayson Fuller, and Finn Woelm (2021). *Sustainable development report 2021*. Cambridge University Press.

Sachs, Wolfgang (1993). *Die vier E’s: Merkposten für einen maß-vollen Wirtschaftsstil*.

Sahakian, Marlyne, Henrike Rau, Eoin Grealis, et al. (2021). “Challenging social norms to recraft practices: A Living Lab approach to reducing household energy use in eight European countries”. In: *Energy Research & Social Science* 72, p. 101881.

Sahakian, Marlyne and Harold Wilhite (2014). “Making practice theory practicable: Towards more sustainable forms of consumption”. In: *Journal of Consumer Culture* 14.1, pp. 25–44.

Samadi, Sascha, Marie-Christine Gröne, Uwe Schneidewind, et al. (2017). “Sufficiency in energy scenario studies: Taking the potential benefits of lifestyle changes into account”. In: *Technological Forecasting and Social Change* 124, pp. 126–134.

Sandberg, Maria (2021). “Sufficiency transitions: A review of consumption changes for environmental sustainability”. In: *Journal of Cleaner Production* 293, p. 126097.

Sanders, Elizabeth B-N (2002). “From user-centered to participatory design approaches”. In: *Design and the social sciences*. CRC Press, pp. 18–25.

Sanders, Elizabeth B-N and Pieter Jan Stappers (2008). “Co-creation and the new landscapes of design”. In: *Co-design* 4.1, pp. 5–18.

– (2014). “Probes, toolkits and prototypes: three approaches to making in codesigning”. In: *CoDesign* 10.1, pp. 5–14.

Sareen, Siddharth and Håvard Haarstad (2021). “Digitalization as a driver of transformative environmental innovation”. In: *Environmental Innovation and Societal Transitions* 41, pp. 93–95.

Saunders, Harty D (1992). “The Khazzoom-Brookes postulate and neoclassical growth”. In: *The Energy Journal* 13.4.

Schanes, Karin, Jill Jäger, and Paul Drummond (2019). “Three scenario narratives for a resource-efficient and low-carbon Europe in 2050”. In: *Ecological Economics* 155, pp. 70–79.

Schatzki, Theodore R. (2021). *A Philosophical Account of the Constitution of Social Life and Change*. University Park, USA: Penn State University Press. DOI: `doi:10.1515/9780271023717`.

Schleich, Joachim, Corinne Faure, and Marian Klobasa (2017). “Persistence of the effects of providing feedback alongside smart metering devices on household electricity demand”. In: *Energy Policy* 107, pp. 225–233.

Schneidewind, Uwe and Angelika Zahrnt (2014). *The politics of sufficiency: making it easier to live the good life*. Oekom-Verl.

Schot, Johan, Laur Kanger, and Geert Verbong (2016). “The roles of users in shaping transitions to new energy systems”. In: *Nature Energy* 1.5, pp. 1–7.

Schubert, Christian (2017). “Green nudges: Do they work? Are they ethical?” In: *Ecological Economics* 132, pp. 329–342. DOI: `https://doi.org/10.1016/j.ecolecon.2016.11.009`.

Schuler, Douglas and Aki Namioka (1993). *Participatory design: Principles and practices*. CRC Press.

Schultz, P Wesley (1999). "Changing behavior with normative feedback interventions: A field experiment on curbside recycling". In: *Basic and applied social psychology* 21.1, pp. 25–36.

Schultz, P Wesley, Jessica M Nolan, Robert B Cialdini, Noah J Goldstein, and Vladas Griskevicius (2007). "The constructive, destructive, and reconstructive power of social norms". In: *Psychological science* 18.5, pp. 429–434.

Schwanen, Tim, David Banister, and Jillian Anable (2012). "Rethinking habits and their role in behaviour change: The case of low-carbon mobility". In: *Journal of transport geography* 24, pp. 522–532.

Schwartz, Shalom H (1977). "Normative influences on altruism". In: *Advances in experimental social psychology*. Vol. 10. Elsevier, pp. 221–279.

Schwartz, Shalom H and Judith A Howard (1981). "A normative decision-making model of altruism". In: *Altruism and helping behavior*, pp. 189–211.

Seaborn, Katie and Deborah I Fels (2015). "Gamification in theory and action: A survey". In: *International Journal of human-computer studies* 74, pp. 14–31.

Sedgwick, Philip and Nan Greenwood (2015). "Understanding the Hawthorne effect". In: *Bmj* 351.

Seidl, Roman, Corinne Moser, and Yann Blumer (2017). "Navigating behavioral energy sufficiency. Results from a survey in Swiss cities on potential behavior change". In: *PloS one* 12.10, e0185963.

Seligman, Clive and John M Darley (1977). "Feedback as a means of decreasing residential energy consumption." In: *Journal of Applied Psychology* 62.4, p. 363.

Sergici, Sanem and Ahmad Faruqui (2011). "Measurement and verification principles for behavior-based efficiency programs". In: *Brattle Group: Cambridge, MA, USA*.

Shih, Li-Hsing and Yi-Cin Jheng (2017). "Selecting persuasive strategies and game design elements for encouraging energy saving behavior". In: *Sustainability* 9.7, p. 1281.

Shove, Elizabeth (2003). "Converging conventions of comfort, cleanliness and convenience". In: *Journal of Consumer policy* 26.4, pp. 395–418.

– (2010a). "Beyond the ABC: Climate change policy and theories of social change". In: *Environment and planning A* 42.6, pp. 1273–1285.

– (2010b). "Social theory and climate change". In: *Theory, Culture & Society* 27.2-3, pp. 277–288.

– (2011). "On the difference between chalk and cheese—A response to Whitmarsh et al's comments on 'Beyond the ABC: Climate change policy and theories of social change'". In: *Environment and Planning A* 43.2, pp. 262–264.

– (2014). "Putting practice into policy: Reconfiguring questions of consumption and climate change". In: *Contemporary Social Science* 9.4, pp. 415–429.

Shove, Elizabeth, Mika Pantzar, and Matt Watson (2012). *The dynamics of social practice: Everyday life and how it changes*. Sage.

Shove, Elizabeth and Gordon Walker (2014). "What is energy for? Social practice and energy demand". In: *Theory, Culture & Society* 31.5, pp. 41–58.

Shukla, P.R., J. Skea, R. Slade, et al. (2022). "Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change". In: DOI: 10.1017/9781009157926.

Sorrell, Steve (2015). "Reducing energy demand: A review of issues, challenges and approaches". In: *Renewable and Sustainable Energy Reviews* 47, pp. 74–82.

Sorrell, Steve and John Dimitropoulos (2008). "The rebound effect: Microeconomic definitions, limitations and extensions". In: *Ecological Economics* 65.3, pp. 636–649.

Sorrell, Steve, John Dimitropoulos, and Matt Sommerville (2009). "Empirical estimates of the direct rebound effect: A review". In: *Energy Policy* 37.4, pp. 1356–1371.

Sorrell, Steve, Birgitta Gatersleben, and Angela Druckman (2020). "The limits of energy sufficiency: A review of the evidence for rebound effects and negative spillovers from behavioural change". In: *Energy Research & Social Science* 64, p. 101439.

Sovacool, Benjamin, Christina Demski, and Lance Noel (2021). "Beyond climate, culture and comfort in European preferences for low-carbon heat". In: *Global Environmental Change* 66, p. 102200.

Sovacool, Benjamin, Peter Newell, Sanya Carley, and Jessica Fanzo (2022). "Equity, technological innovation and sustainable behaviour in a low-carbon future". In: *Nature Human Behaviour*, pp. 1–12.

Sovacool, Benjamin, Jody Osborn, Mari Martiskainen, and Matthew Lipson (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.

Spaiser, Viktoria, Davide Luzzatti, and Aliki Gregoriou (2019). "Advancing sustainability: Using smartphones to study environmental behavior in a field-experimental setup". In: *Data Science* 2.1-2, pp. 277–290.

Spangenberg, Joachim H (2014). "Institutional change for strong sustainable consumption: sustainable consumption and the degrowth economy". In: *Sustainability: Science, Practice and Policy* 10.1, pp. 62–77.

Spangenberg, Joachim H and Sylvia Lorek (2019). "Sufficiency and consumer behaviour: From theory to policy". In: *Energy Policy* 129, pp. 1070–1079.

Spengler, Laura (2016). "Two types of 'enough': sufficiency as minimum and maximum". In: *Environmental Politics* 25.5, pp. 921–940.

Spreng, Daniel (2005). "Distribution of energy consumption and the 2000 W/capita target". In: *Energy Policy* 33.15, pp. 1905–1911.

Spurling, Nicola, Andrew McMeekin, Elizabeth Shove, Dale Southerton, and Daniel Welch (2013). *Interventions in practice: re-framing policy approaches to consumer behaviour*. Tech. rep. University of Manchester, Sustainable Practices Research Group.

Staats, Henk, Paul Harland, and Henk AM Wilke (2004). "Effecting durable change: A team approach to improve environmental behavior in the household". In: *Environment and Behavior* 36.3, pp. 341–367.

Steffen, Will, Katherine Richardson, Johan Rockström, et al. (2015). "Planetary boundaries: Guiding human development on a changing planet". In: *Science* 347.6223, p. 1259855.

Steg, Linda (2008). “Promoting household energy conservation”. In: *Energy Policy* 36.12, pp. 4449–4453.

Steg, Linda and Charles Vlek (2009). “Encouraging pro-environmental behaviour: An integrative review and research agenda”. In: *Journal of Environmental Psychology* 29.3, pp. 309–317.

Stern, Paul C (2000). “New environmental theories: toward a coherent theory of environmentally significant behavior”. In: *Journal of Social Issues* 56.3, pp. 407–424.

Stern, Paul C, Thomas Dietz, Troy Abel, Gregory A Guagnano, and Linda Kalof (1999). “A value-belief-norm theory of support for social movements: The case of environmentalism”. In: *Human ecology review*, pp. 81–97.

Stern, Paul C, Kathryn B Janda, Marilyn A Brown, et al. (2016). “Opportunities and insights for reducing fossil fuel consumption by households and organizations”. In: *Nature Energy* 1.5, pp. 1–6.

Stokes, Matt, Peter Baeck, and Toby Baker (2017). “What next for digital social innovation”. In: *NESTA. Mayo*.

Strengers, Yolande (2014). “Smart energy in everyday life: Are you designing for resource man?” In: *interactions* 21.4, pp. 24–31.

Stuart, Elizabeth A (2010). “Matching methods for causal inference: A review and a look forward”. In: *Statistical science: a review journal of the Institute of Mathematical Statistics* 25.1, p. 1.

Sunio, Varsolo and Jan-Dirk Schmöcker (2017). “Can we promote sustainable travel behavior through mobile apps? Evaluation and review of evidence”. In: *International journal of sustainable transportation* 11.8, pp. 553–566.

Suruliraj, Banuchitra, Makuochi Nkwo, and Rita Orji (2020). “Persuasive Mobile Apps for Sustainable Waste Management: A Systematic Review”. In: *International Conference on Persuasive Technology*. Springer, pp. 182–194.

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

Thøgersen, John and Folke Ölander (2003). “Spillover of environment-friendly consumer behaviour”. In: *Journal of Environmental Psychology* 23.3, pp. 225–236.

Thomas, Stefan, Johannes Thema, Lars-Arvid Brischke, et al. (2019). “Energy sufficiency policy for residential electricity use and per-capita dwelling size”. In: *Energy Efficiency* 12.5, pp. 1123–1149.

Tiefenbeck, Verena, Anselma Wörner, Samuel Schöb, Elgar Fleisch, and Thorsten Staake (2019). “Real-time feedback promotes energy conservation in the absence of volunteer selection bias and monetary incentives”. In: *Nature Energy* 4.1, pp. 35–41.

Toulouse, Edouard, Marlyne Sahakian, Katharina Bohnenberger, et al. (2019). “Energy sufficiency: how can research better help and inform policy-making?” In: *ECEEE 2019 Summer Study Proceedings*.

Triandis, Harry Charalambos (1977). *Interpersonal behavior*. Brooks/Cole Publishing Company.

Turner, Bryan S (2006). *The Cambridge Dictionary of Sociology.* Cambridge University Press.

Urry, John (2009). “Sociology and climate change”. In: *The Sociological Review* 57.2_suppl, pp. 84–100.

Urry, John (2010). “Consuming the planet to excess”. In: *Theory, Culture & Society* 27.2-3, pp. 191–212.

Valentine, Scott, Benjamin Sovacool, and Marilyn Brown (2019). *Policy for a Low-Carbon Future*. New York Chichester, West Sussex: Columbia University Press. DOI: doi:10.7312/vale18596.

Valkila, Noora and Arto Saari (2013). “Attitude–behaviour gap in energy issues: Case study of three different Finnish residential areas”. In: *Energy for Sustainable Development* 17.1, pp. 24–34.

Vine, Edward, Michael Sullivan, Loren Lutzenhiser, Carl Blumstein, and Bill Miller (2014). “Experimentation and the evaluation of energy efficiency programs”. In: *Energy Efficiency* 7.4, pp. 627–640.

Von Wirth, Timo, Lea Fuenfschilling, Niki Frantzeskaki, and Lars Coenen (2019). “Impacts of urban living labs on sustainability transitions: Mechanisms and strategies for systemic change through experimentation”. In: *European Planning Studies* 27.2, pp. 229–257.

Wachsmuth, Jakob and Vicki Duscha (2019). “Achievability of the Paris targets in the EU — The role of demand-side-driven mitigation in different types of scenarios”. In: *Energy Efficiency* 12.2, pp. 403–421.

Wallenborn, Grégoire, Marco Orsini, and Jeremie Vanhaverbeke (2011). “Household appropriation of electricity monitors”. In: *International Journal of Consumer Studies* 35.2, pp. 146–152.

Wang, Danhong, Jonas Landolt, Georgios Mavromatidis, Kristina Orehounig, and Jan Carmeliet (2018). “CESAR: A bottom-up building stock modelling tool for Switzerland to address sustainable energy transformation strategies”. In: *Energy and Buildings* 169, pp. 9–26.

Ward, Adrian F, Kristen Duke, Ayelet Gneezy, and Maarten W Bos (2017). “Brain drain: The mere presence of one’s own smartphone reduces available cognitive capacity”. In: *Journal of the Association for Consumer Research* 2.2, pp. 140–154.

Warde, Alan (2005). “Consumption and theories of practice”. In: *Journal of Consumer Culture* 5.2, pp. 131–153.

Wason, P.C. and J.ST.B.T. Evans (1974). “Dual processes in reasoning?” In: *Cognition* 3.2, pp. 141–154.

Weber, Sylvain, Paul Burger, Mehdi Farsi, et al. (2017). *Swiss household energy demand survey (SHEDS): Objectives, design, and implementation*. Tech. rep. IRENE Working Paper.

Wee, Siaw-Chui, Weng-Wai Choong, and Sheau-Ting Low (2021). “Can “Nudging” Play a Role to Promote Pro-Environmental Behaviour?” In: *Environmental Challenges* 5, p. 100364.

Weinmann, Markus, Christoph Schneider, and Jan vom Brocke (2016). “Digital nudging”. In: *Business & Information Systems Engineering* 58.6, pp. 433–436.

Weiser, Paul, Dominik Bucher, Francesca Cellina, and Vanessa De Luca (2015). “A taxonomy of motivational affordances for meaningful gamified and persuasive technologies”. In: *EnviroInfo and ICT for Sustainability 2015*. Atlantis Press, pp. 271–280.

Wemyss, Devon, Roberta Castri, Francesca Cellina, et al. (2018). “Examining community-level collaborative vs. competitive approaches to enhance household electricity-saving behavior”. In: *Energy Efficiency* 11.8, pp. 2057–2075.

Wemyss, Devon, Francesca Cellina, Manuel Grieder, and Felix Schlüter (2022). “Looking beyond the hype: Conditions affecting the promise of behaviour change apps as social innovations for low-carbon transitions”. Under review.

Wemyss, Devon, Francesca Cellina, Evelyn Lobsiger-Kägi, Vanessa De Luca, and Roberta Castri (2019). “Does it last? Long-term impacts of an app-based behavior change intervention on household electricity savings in Switzerland”. In: *Energy Research & Social Science* 47, pp. 16–27.

Wemyss, Devon, Evelyn Lobsiger-Kägi, Silvana Jud, and Francesca Cellina (2022). “Aligning interests: Evaluation of a co-creation process of an energy savings app with household members and researchers”. Under review.

Wever, Renee, Jasper Van Kuijk, and Casper Boks (2008). “User-centred design for sustainable behaviour”. In: *International journal of sustainable engineering* 1.1, pp. 9–20.

White, Katherine, Rishad Habib, and David J Hardisty (2019). “How to SHIFT consumer behaviors to be more sustainable: A literature review and guiding framework”. In: *Journal of Marketing* 83.3, pp. 22–49.

Whitmarsh, Lorraine, Wouter Poortinga, and Stuart Capstick (2021). “Behaviour change to address climate change”. In: *Current Opinion in Psychology* 42, pp. 76–81.

Wickham, Hadley (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Wiedmann, Thomas and Jan Minx (2008). “A definition of carbon footprint”. In: *Ecological economics research trends* 1, pp. 1–11.

Wilhite, Harold, Elizabeth Shove, Loren Lutzenhiser, and Willett Kempton (2000). “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*. Springer, pp. 109–126.

Willer, David and Henry A Walker (2007). *Building experiments: Testing social theory*. Stanford University Press.

Wolske, Kimberly S, Kenneth T Gillingham, and P Schultz (2020). “Peer influence on household energy behaviours”. In: *Nature Energy* 5.3, pp. 202–212.

Wood, Georgina, Dan van Der Horst, Rosie Day, et al. (2014). “Serious games for energy social science research”. In: *Technology Analysis & Strategic Management* 26.10, pp. 1212–1227.

Wooldridge, Jeffrey M (2010). *Econometric analysis of cross section and panel data*. MIT press.

– (2015). *Introductory econometrics: A modern approach*. Cengage learning.

– (2021). *Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators*. https://dx.doi.org/10.2139/ssrn.3906345.

Wynes, Seth and Kimberly A Nicholas (2017). “The climate mitigation gap: Education and government recommendations miss the most effective individual actions”. In: *Environmental Research Letters* 12.7, p. 074024.

Xu, Rui and Donald Wunsch (2005). “Survey of clustering algorithms”. In: *IEEE Transactions on neural networks* 16.3, pp. 645–678.

Zangheri, Paolo, Tiago Serrenho, and Paolo Bertoldi (2019). “Energy savings from feedback systems: A meta-studies’ review”. In: *Energies* 12.19, p. 3788.

Zhao, Qin-Yu, Jing-Chao Luo, Ying Su, et al. (2021). "Propensity score matching with R: Conventional methods and new features". In: *Annals of translational medicine* 9.9.

Zichermann, Gabe and Christopher Cunningham (2011). *Gamification by design: Implementing game mechanics in web and mobile apps*. O'Reilly Media, Inc.

Zimmermann, Sina, Andreas Hein, Thomas Schulz, Heiko Gewald, and Helmut Krcmar (2021). "Digital Nudging Toward Pro-Environmental Behavior: A Literature Review." In: *PACIS*, p. 226.

List of Figures

List of Tables

Declaration

I declare that I have completed this dissertation solely and only with the help of the references I mentioned.

Milano (Italy), Academic Year 2021/2022

Francesca Cellina

Francesca Cellina, registration number 854291

Colophon

This thesis was typeset with $\LaTeX\ 2_\varepsilon$. It uses the *Clean Thesis* style developed by Ricardo Langner. The design of the *Clean Thesis* style is inspired by user guide documents from Apple Inc. Download the *Clean Thesis* style at `http://cleanthesis.der-ric.de/`.