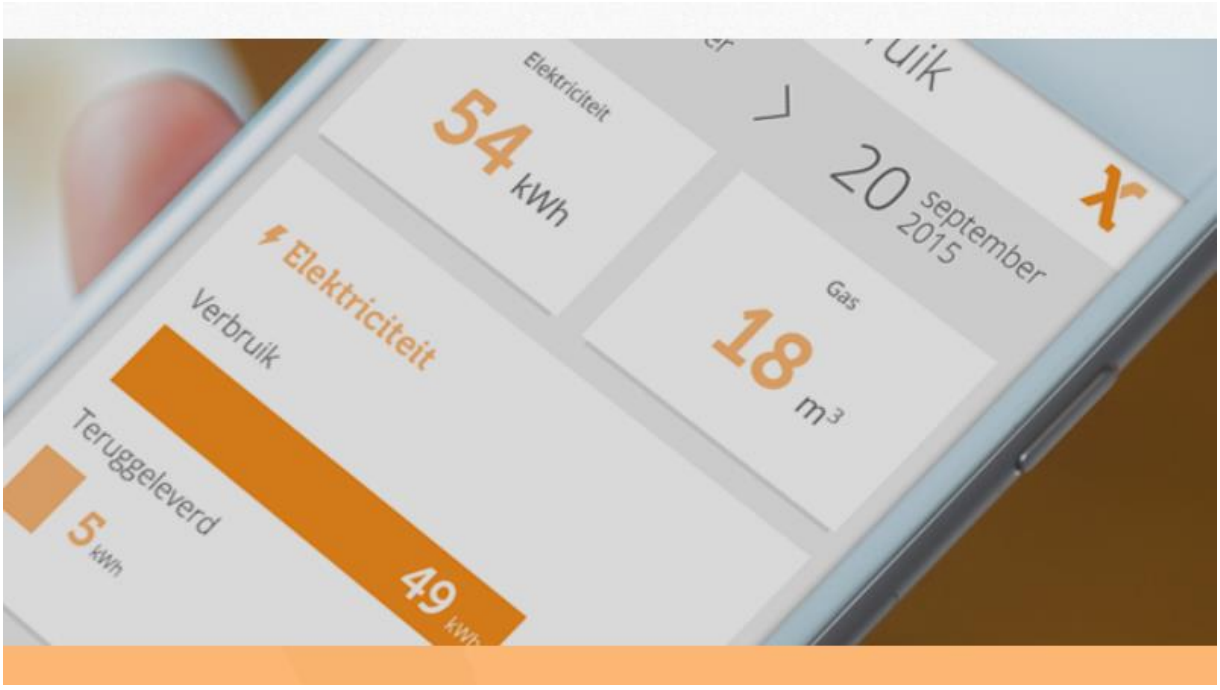
MARCH 23, 2020

Gamified energy apps within the Dutch market

CAN GAME DESIGN ELEMENTS LOWER HOUSEHOLD ENERGY CONSUMPTION ?



LUUK KOEDAM
UTRECHT UNIVERSITY , MASTER ENERGY SCIENCE
Student number: 5572312

SUPERVISORS
Dr. Tarek Alskaif, and Dr. ir. Iannis Lampropoules

Preface

This thesis project is conducted- and this report is written by me (Luuk Koedam). However, without the help of many, it would not have been possible to conduct this research. Therefore, I want to thank some people before the research and thesis report are presented. First, I want to thank all the participants for participating in the experiment and helping me every time it was needed during the experiment. Thanks to them, the research shows interesting results that can be an inspiration for future research. Second, I want to thank my supervisor Dr. Tarek Alskaif for all his help in setting up the research, supporting me during the whole process, and providing comments on the thesis report. In advance, I also want to thank him, and Dr. ir. Iannis Lampropoules for reading and evaluating my thesis report.

Third, I want to thank Dr. ir. Jesus Rosales Carreon for helping me finding a thesis subject that fits my interest and connecting me with my supervisor Dr. Tarek Alskaif. Lastly, I want to thanks Prof. dr. van Sark for allowing me to participate in the program of the master Energy science. It was a long journey with some significant ups and downs. However, this never happened if I did not get the opportunity to start.

Summary

Next to transport, households are the largest energy-consuming sector. The sector accounts for 27.2% of the total final energy consumption in the EU. Referring to the Paris agreement (United Nations, 2015) and the new 'Green Deal' of the European Commission (becoming energy neutral by 2050), it is crucial to lower the energy consumption of the household sector as far as possible. Buildings can be modified and equipped with RES installations and extra insulation. However, for applying these measures, significant investments are required and the implementation is time-consuming. Next to these measures, human decision making and behavior are critical in lowering the energy consumption of households. Residents need to use energy more efficiently and to enable change in behavior, they first need to get insights into their energy consumption. Energy apps that are connected to a smart meter can provide these insights, and by implementing game design elements in these apps, the behavior of users can possibly be influenced.

This research is inspired by other studies that show that specially designed gamified environments are able to motivate consumers to conserve energy. Nevertheless, these environments are not available to the general public. This reveals a research gap and the focus of this research. It needs to be investigated to what extent widely available energy apps include game design elements, and what effect these elements have on the behavior of its users. Therefore, to investigate to what extent game design is implemented in energy apps, the available energy apps in the Dutch market are evaluated based on the framework of AlSkaif at al. (2018). This framework consists of five game design types and 18 game design elements. In total, 30 apps with the goal of energy efficiency within households are evaluated. These apps were found by using a similar method as a research in the U.S. in which energy apps were evaluated (Beck, Chitalia, & Rai, 2019).

For the present evaluation, the game design types and elements were clearly defined, and the apps were thoroughly examined. To enable a comparison of to what extent a particular energy app is gamified, a formula has been created and has been used to grant a game design score to all apps. The best scoring app received only 9.8 of the maximum 28.8 points. On average, the apps only received a score of 3.8 points, indicating that the energy apps in the Dutch market are not that gameful. These are interesting findings for the experiment to test the effect of the game design elements. Even the best-performing apps do not make use of the most effective game design elements used in previously done experiments. Therefore, the experiment is split into three months, all with a different focus. During the first month (December 2019), the participants got access to one of the three selected apps, and no further game design elements were added. During the second month (January 2020), game design elements that trigger the intrinsic motivation were added by the researcher (tips, data-driven reports, and a happy holiday e-card). During the third month (February 2020), elements which trigger the extrinsic motivation were added (competition, leaderboard, and prizes).These elements can have a different effect on each individual participant, and to make the experiment even more interesting, the participants were sorted into three groups, all with a different social connection to each other. One group consisted of 11 households that lived in the same apartment building, one group with nine participants consisted of friends of the researcher, and the nine participants in the last group did not have a social connection with each other. During the experiment, it was monitored whether these participants are affected by others in their group. This creates an exciting experiment with a 3x3 design, three different energy apps, three months to test different types of motivation, and three different social groups. The experiment only focused on electricity consumption since not all of the participants use natural gas for heating.

The electricity consumption of the participants is compared to their consumption in the same period last year. Over the whole three-month period and all participants combined, the data shows a decrease in consumption of 3.2%, if the most extreme participant is excluded (+74%). All three months, the neighbors group shows the best result, with an impressive average decrease in electricity consumption of 9.7%. The group without a connection shows a decrease of 5.7%, and the friends group an increase of 6.8%. Of the three different apps, the participants that used the most gamified app, Smappee (used by three participants in each social group), show an impressive average decrease in consumption of 7.9%, compared to a decrease of 3.7% for the Smart Dodos app, and an increase of 13% by the U-meter app users. All data obtained during the experiment was cross-checked with a survey before and after the experiment. These surveys reveal that participants found that the competition, leaderboard, and tips game design elements were most motivating. Also, 28 of the 29 participants think game design elements can cause people to deal more consciously with electricity. An interesting result that shows that energy apps and game design can help in reducing the energy consumption of households. Hopefully, this result inspires researchers to conduct similar larger-scale experiments, because due to the small sample size, it is impossible to draw definite conclusions.

Table of contents

1. Introduction

Climate change has been recognized as a threat to our planet by 195 nations under the Paris agreement of the United Nations Framework Convention on Climate Change (UNFCCC) in 2015. Nations agreed on taking action to combat this threat and limit the rise in global temperature below 2 degrees Celsius (United Nations, 2015). Following up on this agreement, conferences were held in Katowice and Madrid to create implementation guidelines and to operationalize the Paris agreement. The countries that signed the agreement are required to submit new national climate action plans by 2020, and the Madrid conference served to build ambition ahead of 2020 ("About the UN Climate Change Conference - December 2019," n.d.). However, during this Madrid conference the nations failed to come to an explicit agreement.

As a reaction to the meetings of UNFCCC, the European Commission, on its own, created multiple directives on energy policies over the past years and the latest being the ambitious 'Green Deal' to become climate natural by 2050. Next to producing energy from renewable sources, energy efficiency is one of the key drivers to reduce CO_2 emissions (European Parlament, 2012). The EU directive on energy efficiency of 2012 sets out a target to reach 20% energy efficiency by 2020. Another target to reach at least 32,5% energy efficiency by 2030 compared to 2007 was introduced in an amending directive in 2018 (European Comission, 2019). The largest energy consumers in the EU are buildings; they represent 40% of the total energy consumption and 36% of the CO_2 emissions (European Comission, 2014b). Due to this large share, the EU also created a specific directive on the energy performance of buildings. This directive emphasizes on the residential sector (European Parlament, 2010). The reason why this particular directive was created can be explained by the high energy consumption of households. In 2010 households were responsible for 26.7% of the total energy consumed in the EU, making it the largest energy-consuming sector after transport with 31.7% (European Union, 2013). This share is rising over the past decades, and in 2017 households, or the residential sector, accounted for 27.2% of the total final energy consumption in the EU ("Energy consumption in households," 2019).

Taking into account the share of the residential sector in the total energy consumption in the EU, residential energy-saving measures could potentially have a substantial effect on reaching the EU efficiency targets. New building regulations and the retrofitting of existing buildings will cause an improvement in energy efficiency, and appliances also become more efficient due to new regulations and labeling systems (European Commission, 2009). However, these measures can be expensive and without governmental support unreachable to lower-income households. The development of technologies that reduce energy consumption is necessary but insufficient unless widely adopted by consumers. If there are measures that do not require significant investments and are available to all households, these could create a quick win with an enormous potential impact, even if the reduction per household is small. Next to the characteristics of a building and the appliances, the behavior of residents determines the amount of energy that is consumed (van Raaij & Verhallen, 1983).

To promote energy conservation this human side needs to be addressed, which means that more attention needs to be paid to social- psychological aspects of conservation. The complexities of human information processing and the positional characteristics of individual consumers, place energy conservation in the field of social and behavioral sciences (Costanzo et al., 1986). This specifically applies to the residential sector, where human decision making determines the total energy consumption. The environmental behavior of an individual can be influenced by many personal and social factors such as knowledge, education, values, political- and worldviews, religion, norms and social class (Gifford & Nilsson, 2014). Due to this wide range of factors, extensive research can be relevant to the creation of measures that conserve household energy consumption. A review of intervention studies aimed at energy conservation within households concluded that commitment, goal setting, information provision, feedback, and rewards are all intervention types that potentially can change someone's behavior. Residents can lower their energy consumption with the help of these interventions, especially when they are repeated, monthly weekly, or daily (Abrahamse et al., 2005).

A field of research that becomes increasingly popular for interventions and behavioral change is gamification, and a review on applied gaming interventions showed that both game interventions and gamification can foster energy-savings behaviors (Morganti et al., 2017). One of the most often used definitions of gamification is "the use of game design elements in non-game contexts'' (Deterding & Dixon, 2011). Another view on gamification describes it as "a process of adding game design elements and creating gamefull experiences, relying on particular game elements or experiences, instead of the development of a game" (Albertazzi, Ferreira, & Forcellini, 2019). A survey study among 1,500 students on the effectiveness of game design elements in Malaysia predicted a positive effect on energy conservation behavior. There was no field research carried out, so the researchers of this survey were not able to properly examine the effect if these elements were implemented in an energy-saving campaign (Wee & Choong, 2019). Nevertheless, it indicates that energy consumers are open to the idea of using game design interventions to influence their behavior. If these types of game design interventions can be automated and made accessible to a broad audience, it could be a relatively easy way to accomplish the goal of lowering the energy consumption of households.

Energy applications (apps) can be a solution to make these game design elements available to everybody, especially if these apps can be easily accessible on smartphones. In the Netherlands, around 93% of the inhabitants have a smartphone, making it the country with the highest smartphone penetration rate of all countries surveyed yearly by the accountancy and consultancy firm Deloitte. Next to the Netherlands, other surveyed European countries have at least an 80% penetration rate ("Smartphonebezit gegroeid naar 93% van Nederlanders, veelvuldig gebruik storend," 2018). A gamified energy app that includes real information on a households energy consumption should communicate with a metering system that tracks this consumption. The EU requires member states to ensure the implementation of smart metering systems in line with the Energy Efficiency Directive. For electricity, there is a target of rolling out such systems in at least 80% of the households by the end of 2020 in all EU member states (European Comission, 2014a). The Netherlands is among the countries that proceed with the large-scale roll-out, and all households will be equipped with such a smart meter by the end of 2020.

Awareness of the energy that is consumed can be considered as a first step to involve the end-user of energy actively. However, a smart meter by itself does not enable a resident to monitor its energy consumption in a convenient way (AlSkaif et al., 2018). Therefore, Information and Communication Technologies (ICT) take upon an essential role in the development of eco- feedback (Verkade & Höffken, 2017). ICT involves the development of apps that communicate with a smart meter and show the end-user how much energy is consumed. A study that has tested a mobile-based gamified app called Power Advisor, showed that all participants, even the seven households who had little or no interest, had increased their awareness of energy consumption. This was accomplished with the help of elements that provide feedback on energy use and comparing their consumption to others (Kjeldskov et al., 2012).

Another experiment with an energy game to stimulate energy savings in households shows that setting up a game can potentially change the habits of the participants (Geelen et al., 2012). Even though this experiment showed a positive result, the game was only available to a select group. A research executed in the U.S. created a gamified platform that is used by more than six million people and managed by the company OPower. This platform aims at inducing energy conservation with the help of ''Home Energy Reports'' including personalized feedback, energy conservation information and social comparisons. A survey among 6000 users and their data on energy consumption during the reports and two-year pre-experiment, showed an immediate overall energy conservation thanks to these gamified interventions. This conservation decayed relatively quickly after the positive results of the first interventions. However, sending these data-driven personal reports frequently over a more extended period still caused a reduction in energy consumption and shows that users slowly adjust their behavior. One of the most remarkable facts was that a randomly selected group which stopped these reports after two years showed much more persistent energy conservation than they had in between initial reports. They formed a new consumption behavior that was maintained and improved by their perseverance (Allcott & Todd, 2012).

The research of Konstantakopoulos et al. (2019) with student dorms in California presented a social game in a specially designed online environment. It measured the energy consumption of appliances with the help of Internet of Things (IoT) sensors and a deep learning algorithm. The only incentive for the students was a competition and a lottery with gift certificates. The users had no financial gain in the experiment since the energy costs were included in the rent. The results were positive and it can be concluded that this specially designed environment could save much energy when implemented in more student housing buildings. Another study among student-households with a competition incentive in the Netherlands called "The Energy Battle'' showed a positive result with an average energy saving of 24%. The researchers concluded that competition could be a powerful tool to motivate participants intrinsically. The facts that nobody attended the party that was given as a price for the best performing group, and that the reduction in energy consumption diminished after the competition, indicates that it was purely the competition element that motivated the students (Geelen et al., 2012). Two other studies in Sweden with pervasive games called Power Explorer and Power Agent, concluded that social interaction in the form of peer pressure and the cooperation of family members could be very motivating. They also suggested that casual gameplay with less extreme behavior has more effect on the long term (Bang, Svahn, & Gustafsson, 2009; Gustafsson, Katzeff, & Bang, 2009).

A research executed by Rai & Beck (2017) on the use of a trivia-style game in breaking informational barriers in residential solar adoption, shows that games can help in addressing information gaps and facilitating energy-related behavioral change. Suggesting that next to monitoring, data-driven feedback, and competition, providing general information in a gameful way can help in correcting misperceptions around energy use and conservation. The same researchers analyzed energy apps that are available in the United states (U.S). They concluded that energy apps underutilize game design elements, as well as the behavioral constructs that are known to impact decision making and behavior (Beck et al., 2019).

Although the energy apps available in the U.S have been analyzed in previous studies, there is no research on the use of similar energy apps that are available in the EU app stores. The U.S research provides information on the fact that game design elements are implemented, but it does not test the applications with end-users (Beck et al., 2019). The experiments in the Netherlands, Sweden, and the U.S all used specially designed game environments and tested specific game design elements. This reveals a research gap that can be relevant to impact the total energy consumption of households with the help of gamified apps. The Netherlands seems to be an ideal country to fill this gap, considering that energy apps are widely available, the smart meter roll-out, the previously done research with positive results, and the smartphone penetration rate.

1.1 Research goal and relevance

Climate change can be seen as a collective responsibility and the most critical disaster that the world is currently facing. Every individual living on the planet is in a smaller or larger extent an actor that causes this threating situation (Banks, 2013). Of course, most individuals do not have the power to change the energy strategies of industries or countries, they can only change their own behavior. Gamification could help with such a change in behavior. It can reduce the energy demand of residents without the need for significant investments or new policies. It has the potential to help all households lowering their energy consumption and subsequently contribute to mitigating climate change. In this process, there is also a financial incentive on an individual level since lowering energy consumption will automatically lower energy costs. Especially for those who cannot afford investments for retrofitting measures, this can be an opportunity to ease their financial situation and make them aware of the direct connection between personal and environmental gains. Of course, one person that changes its behavior will not have an impact on climate change. Nevertheless, the collective behavior can undoubtedly affect lowering the energy consumption of the residential sector as a whole. This individual awareness, relatively small costs for implementation, and the potential environmental impact, make research into game design elements within energy apps a relevant research topic. Especially now, due to the roll-out of smart metering systems and high smartphone penetration rate in the EU.

As mentioned in the first section, previously done research on the effect of game design elements used specially designed environments (Bang et al., 2009; Geelen et al., 2012; Gustafsson et al., 2009; Konstantakopoulos et al., 2019). The research of Beck, Chitalia, & Rai (2019) evaluates the available energy apps in the U.S but does not use these apps in an experiment to measure the effect of these elements. This study combines the methods and knowledge of this related research to point out elements that affect a user's behavior. The framework of AlSkaif et al. (2018) can be qualified as a guideline to determine what kind of game design elements are implemented in energy apps. With the help of this framework, the available energy apps in the Dutch market can be evaluated. The effect of the evaluated energy apps and the implemented game design elements can be tested with an experiment in real households. Taking a look at the studies which form the knowledge base for this present research; competition, data-driven feedback, social comparison, and information provision seem to be promising elements. Therefore, these game design elements need be included in the experiment of this study, despite whether these are implemented in the evaluated apps. The missing elements can be designed by the researcher and sent to participants during the experiment. In formulating the research question, we need to consider the two main goals of this study:

1. Evaluating energy apps that are available in the Dutch market and app stores on the use of game design elements.
2. Using the apps in an experiment to determine the effect of the game design elements.

The performance of the research tasks in this study follows the order of the above mentioned research goals. The first goal is reached by doing desk research on game design elements, selecting and evaluating the available apps. The second goal is reached by performing field research and using the apps in real households to determine the effect on user's behavior and energy consumption. With this in mind, the research question reads as follows;

To what extent do energy apps in the Dutch market use game design elements, and what is the effect of these elements on the behavior and energy consumption of its users?

2. Theoretical background

To answer the research question, it is essential to consider in what way game design elements can influence a user's behavior. Research of Sailer et al. (2013) describes gamification as an innovative approach to foster motivation, and game design elements can address different motivational mechanisms. Therefore, a better understanding of these motivation mechanisms can be seen as a starting point of this research. The article written by Sailer et al. (2013) describes six different perspectives of motivation that relate to gamification. This can be useful within the experiment to evaluate if game design elements that trigger a certain motivation perspective have more effect than others.

Trait perspective: Individual characteristics that are relatively stable over time and have a strong influence on behavior. These include strong motives such as achievement, which can be triggered by success and progress, power that is triggered by control and competition, and affiliation that is connected to membership.
Behaviorist learning perspective: This includes motivation that grows over time and is based on positive and negative reinforcements in the past. It is the result of previous experiences, which means that users can be affected by rewards for good behavior and direct feedback.
Cognitive perspective: This motivation is based on means-ends analyses relating to internal processes such as expectancies, estimations and assessments. It can be triggered by describing and visualizing the future and setting a clear and achievable goal. It emphasizes the importance of a person's action within a given situation.
Self-determination perspective: The focus of this motivation perspective are the social-contextual conditions. A feeling of competence, autonomy, or social relatedness can have an effect, and can also be addressed by extrinsic motivation.
Interest perspective: This takes individual preferences and content into account. People are motivated when sparked by content that aligns with their interests and skills. Providing personal feedback and adapting the difficulty to a person's skills and competences can, therefore, be motivational triggers.
Emotion perspective: This relates to the interaction of emotions within cognitive and motivational processes. It can be influenced by instructional strategies that cause a decrease in negative emotions like fear, envy, and anger, and increases positive feelings like sympathy and pleasure.

Game design elements can potentially influence all these perspectives and earlier done research shows an overall positive attitude towards the use of gamification within the field of energy conservation. However, if game design elements can trigger such a positive effect, it could possibly also trigger a negative effect on users' motivation. Most studies about gamification only focus on the positive effects, however a review study of Hyrynsalmi, Smed, & Kimppa (2017) looked into the limiting and harmful effects of game design elements. In this study it was concluded that several authors have expressed their worries about end-results that are over-optimized. Game design elements can also be de-motivating due to frustrating simplicity or childishness of the requested tasks. These are limiting effects, but more interesting would be the harmful effects, yet these are only discussed by a few authors. The respective authors mention that gamification could encourage users to perform behaviors only if rewarded or that the game design elements can distract users from their primary purpose. For the experiment in this research, it is important to take into consideration if certain elements that are implemented in the apps or added during the experiment, can have such a negative impact.

In Switzerland, a study was conducted to determine the long-term effect of a mobile app game intervention. The researchers concluded that while the positive effect was significant during the research, it diminished a year later. This disappointing result was surprisingly not how the participants experienced it themselves. They reported their behavior as more efficient compared to their behavior prior to the intervention (Wemyss et al., 2019). This on itself can also be a limiting or even harmful effect. This false sense of 'doing good' could cause users to be less motivated to take more actions to lower their impact on the environment. Taking this into account, it is essential to see if the experience of users in the experiment differs from the data on their consumption.

Before a positive or negative effect of game design elements within energy conservation apps can be researched, it first needs to be clear what kind of elements are useful in the field of energy conservation. The framework of AlSkaif et al. (2018) clearly describes the possible game design elements that can be implemented in energy apps. In total the framework consists of five different game design types, and a total of 18 game design elements. For this research, the elements are described in more detail than in the framework to make sure there is no overlap between elements.

1. Information provision
Information provision elements can give residential customers an overview of their energy consumption and behavior. It allows them to understand what the impact is on energy consumption if they take specific actions (Geelen et al., 2012; Seaborn & Fels, 2015). For information provision there are three game design elements;

Statistics: The app provides statistics on the energy that is consumed. These statistics can be real-time or based on historical data (for example, from the previous day, month or year). Information should be presented using clear and compelling visuals and graphics.
Messages: The app sends messages that provide information on energy consumption. These messages are based on energy consumption data (data-driven reports). The messages should be communicated at intervals that consumers accept, either using an inbox in their web or mobile apps or through email or text messages.
Tips: The app provides tips on how to lower energy consumption. These tips are not restricted to consumption and can also provide information related to energy produced by Renewable Energy Resources (RESs), or measurements that can be taken to improve the overall energy consumption of a building, such as insulation measures. These can be both general and personal.

2. Rewarding system
A reward system could potentially motivate a user to take specific actions and to alter their consumption behavior. This reward system can make use of credits that are proportional to the effort of a user (Seaborn & Fels, 2015). For a rewarding system there are three game design elements;

Discount on electricity bill: The app provides information on how consumers can lower their energy costs by changing their behavior, effort, and impact. For example, giving discounts on energy prices during peak electricity consumption hours.
Virtual currency: The app has a virtual currency system that enables users to earn virtual currency when they take specific actions, for example, using their self-produced energy. This currency can be used to get discounts on products, energy, or any other type of reward.
Prizes, offers, and coupons: The app gives users the incentive to actively use the app by offering prizes, special offers, or coupons. These rewards can also be given just for using the application without taking any further action, for example, finding a cheaper energy supplier based on the energy consumption of a user.

3. Social connection
To make energy apps more fun and appealing to residential consumers, elements that create a social connection can be implemented. This social connection can be both with acquaintances of users, such as neighbors, family, or friends, as well as with unknown users, like similar households or app users in general. This social connection can be created with a competition, energy community or collaboration element (Lampropoulos et al., 2019).

Competition: The app enables users to compare their energy consumption with other users; friends, family, neighbors, or an average household. There is no need for a reward in this element since competition itself can create a normative social influence and bring social aspects into the system.
Energy community: Users of the app can become part of an energy community within the app itself or through social media, a web-based forum, or a private server. Within this community, users can share their experiences and ask questions to each other and the energy community administrator. It is not necessary to link this platform directly to the app. However, the supplier of the app needs to provide information on how to become a member of this community.
Collaboration: Within the app or the energy community, shared goals are set or discussed to change from solely an individual effort, score, or achievement to additional collective efforts. Collaboration can be on a local level, for instance, to reduce the consumption of a neighborhood, or to use electricity produced by neighbors who have RESs installations. It can also be on a national level for all application users, for instance, shifting demand to peak RESs production hours.

4. User interface
Users can be influenced and motivated by an interactive interface of an app. Next to usefulness and ease of use, the interface design could be made enjoyable and exciting (AlSkaif et al., 2018). These

types of elements could have an overlap with other types of elements, and therefore for this research, a clear distinction is created.

Dashboard: Most apps will make use of a dashboard, as this term is widely used for the home screen of an app. It is expected that a considerable number of apps will use statistics and graphs on energy consumption within their dashboard. Due to this overlap, showing only statistics on the amount of energy that is consumed is not a gamified dashboard. Adding elements that trigger action from a user, such as challenges, goals, stories, or comparisons to other users, or the energy consumption of another period is conceived a gamified dashboard.
Leaderboard: Leaderboards can be part of a competition to show how a user is doing in comparison to other users. To distinguish the leaderboard and the competition element, a leaderboard needs to show specific users and ranks according to their performance.
Progress bar: A progress bar can have different goals, for example, to show a user's progress in comparison to another time interval (last month or year), or to show the progress to reach a particular goal (e.g., saving a certain percentage of energy). A progress bar needs to be clearly visualized as an open space that can be filled showing progress or a bar that is laid over a measuring rod.
Notifications: Notifications could overlap the messages and tips elements from the category information provision. To ensure an overlap is prevented as much as possible, notifications within the user interface are seen as push notifications that notify users on things that stand out in their account. For example, a high consumption of certain appliances. These notifications aim at the user to take immediate action or to provide a good feeling by showing positive results of particular actions.
Message box: A message box is seen as a chat or contact form to enable users to connect the app administrator directly.
Degree of control: If an app enables users to create pre-set programs or if the user can control devices from the app, such as the thermostat, this gives the user a degree of control.

5. Performance status
The behavior of users can be intrinsically changed if users get rewarded for completing specific challenges or tasks. This can create a performance status and will show users that they are doing good This can include three game design elements (AlSkaif et al., 2018).

Points: Points that can be earned by taking specific actions within the application or by completing challenges. All actions give the user a certain amount of points.
Badges: Badges are visual icons that show achievements and can be obtained by taking specific actions within the app.
Levels: This can refer to ranks that the user can earn by taking actions like lowering energy consumption. Each level can be connected to a particular title that virtually describes customer engagement in the applications.

Next to evaluating available apps in the Dutch market based on the described types and elements, this study will focus on experimenting with end-users. As described in the introduction, the most promising elements need to be included in the experiment. The whole experiment, including these elements, needs to be designed. Morschheuser et al. (2017) created a framework after reviewing available studies on how to successfully design gamification. This framework consists of 7 phases, all with actions that need to be considered before moving to the next step.

Project preparation: All gamification experts recommend starting with the identification of problems that should be addressed. Clear goals need to be set for the use of gamification to measure the success of the project.
Analysis: User analysis is an essential element to define the user needs, motivations, and behavior in the current system. This can include interviews, observations, measurements of actual user behavior, analyzing behavior chains, surveys, and diaries.
Ideation: The phase for explorative brainstorming followed by consolidating the ideas to create a list of ideas for the design phase.
Prototype design: To successfully develop game design elements it is recommended to implement an integrative design process, which includes frequently testing and improving until the elements are efficient and reach the target for that specific element.
Implementation: Many studies that are reviewed for this framework describe implementation as continuous prototyping and recommend an iterative procedure in development cycles.

Evaluation: This phase is crucial to investigate if the developed solution meets the defined objective. This can include interviews, surveys, impact studies, A/B-testing, or live playtesting.
Monitoring: Most experts recommend a monitoring phase with regular intervals in which the system usage is investigated. Gamification is an ongoing process and should not become part of the regular work routine within an organization. Targets can be adjusted and the goal is to continuously improve the elements with the help of launch and post-launch monitoring.

For the experiment in this study, the overall goal of the elements is clear; create game design elements that could lower users electricity consumption, or uncover a potential negative effect of an element. The continuous process of evaluating and improving the game design elements will not be possible due to the temporary nature of the experiment. Nevertheless, for the creation of game design elements in this research, the framework of Morschheuser et al. (2017) is beneficial due to the smooth and well-described steps that can be followed to create useful extra game design elements within the experiment.

A survey among 1.500 students in Malaysia showed that nine game design elements; personal profile, non-fixed structure, challenge, feedback, short cycle time theme, competition, cooperation, chat-based social network, potentially have the most effect (Wee & Choong, 2019). As mentioned before in the first section of the introduction, many other studies also show that social connection or pressure, competition, cooperation, and feedback are successful game design elements to lower energy consumption (Bang et al., 2009; Geelen et al., 2012; Gustafsson et al., 2009; Kjeldskov et al., 2012). For the experiment in this research, it is important to take into account the findings of previous done studies, and to include these most promising game design elements and test them with the app users.

3. Methodology

For this research, a total of 860 hours is available spread over a period of nine months. The first ten weeks and about 160 hours is spend on writing a research proposal, learning from other researchers on the subject of game design elements and energy apps. The evaluation of the apps needs to be completed in 9 weeks, with 180 hours available. This includes finding, contacting, installing, and evaluating the apps. The evaluation needs to be completed before November 2019, since it is expected that at least four weeks and 80 hours are needed to get all the participants ready for the experiment. The experiment itself will cost three months starting from December 2019 with 240 hours available for monitoring and the creation of game design elements. During the whole period and especially during the last month, 200 hours is spent on processing the obtained data and writing this report. As an introduction to the methodology of this research, a planning schedule is presented in figure 1.

Figure 1. Timetable research

3.1 Evaluation apps

The evaluation off the apps is done in three simple steps: Selecting energy apps, evaluating the available apps, and selecting apps for the experiment.

3.1.1 Selecting energy apps

To select which apps will be evaluated, a similar method as used in the research of Beck et al. (2019) is applied in this study but for the Dutch instead of the U.S market. In that research, the evaluation was restricted to the U.S because they found that “identification of energy apps is challenging without awareness of existing energy programs and difficult to analyze without fluency in the language and socio-cultural context of the energy system”(p. 33). This research is restricted to the Dutch market, however, it can be expected that there are English apps available in the Netherlands. These apps are excluded because all evaluated apps need to be usable in all Dutch households, even the households where nobody speaks English. Another criterion is that only apps that are considered suitable for the experiment are evaluated, where the research of Beck et al. (2019) was aimed at all apps that target direct energy use behavior. This included energy efficiency and conservation behavior, solar or renewable energy, home energy efficiency upgrades, and efficient transportation. The selection of apps for this study is restricted to the ones that target energy efficiency and conservation behavior within households. This can include web apps that can be found with Google search engine, and mobile apps that are available through the Apple app store and Google play store.

As in the research of Beck et al. (2019), the first step to identify suitable apps is a general keyword search with Google search engine for 'apps' plus one of the following Dutch keywords; energy, energy efficiency, energy savings, energy monitoring, energy consumption, house monitoring system, smart home, smart meter (energie, energie efficiëntie, energie besparen, energie verbruik, energie monitoren, huis monitor systeem, slim huis, slimme meter). This keyword search resulted in 30 unique apps that were found on websites of suppliers, blog sites about energy apps, or websites that compare apps like https://www.energieverbruiksmanagers.nl/overzicht-energieverbruiksmanagers/. From these 30 apps, 13 apps are available in the Apple App store and Google play store. To prevent keyword bias, the 13 app store apps were used to reverse engineer the keywords that are used by developers. A tool in app analytics named AppAnnie (“AppAnnie Intelligence,” n.d.) shows for which keywords an app is ranked. With the help of this tool and the previously found apps, an extended list of in total 23 keywords is created. These keywords are used in the App stores and Google search engine to find other suitable apps. For all the keywords, the first 50 hits were examined on the criteria that they are applicable in households and target energy consumption. This extended search resulted in a sample of 48 apps.

The evaluation of the apps is based on a framework developed by AlSkaif et al. (2018), this research describes that the apps where residential customers can be engaged and play an active role can be classified into three categories; energy efficiency, self-consumption, and demand response. Self-consumption is mainly directed at households that implement solar Photovoltaics (PV) installations, and these apps can help to give insights into the electricity that is locally produced and how much of this electricity is consumed directly, stored into batteries or fed back to the grid. Apps that are developed for this purpose cannot be used in the experiment since most of the selected households do not have a PV-system. Demand response programs aim at altering short term electricity use patterns for market optimization, balancing supply and demand, increased integrations of RESs, and ensuring system security (Lampropoulos, Kling & Ribeiro, 2013). If the selected apps have the goal of altering short term decisions of users, they need intensive monitoring. In this research, intensive monitoring is not feasible due to the limited time to monitor each participating household. These constraints narrow the evaluation of apps that have the primary goal of energy efficiency. From the 48 apps that were selected, seven apps are mainly used for PV systems, tracking the electricity production and self-consumption. None of the apps are used for demand response programs, leaving 41 apps suitable for the evaluation of the use of game design elements.

The final sample of applications
The 41 apps that are suitable for evaluation can be divided into four categories that all require a different approach to get access to the app. These categories and the approach used to get access are described below;
1. Free and accessible to all: A total of nine apps are accessible to everybody that wants to make use of these apps and can be easily installed by creating an account and connecting the app to a smart meter. Some of the apps can also be used without a smart meter by filling in analog meter readings. From the nine apps in this category, eight can be evaluated because one is only accessible if the municipality where a user lives provides an access code.
2. Only accessible to customers of a particular energy supplier: In the Netherlands, there is a wide range of energy suppliers. Part of these suppliers created an energy app that can only be used by customers that have a contract with that particular supplier. From the 41 apps, nine apps belong to this category. To obtain access to these apps, the network of the researcher of this study is used to find customers of all suppliers. They were asked to provide access with the use of their credentials. Thanks to this method, it was possible to get access to eight of the nine apps in this category.
3. Paid and accessible to everybody: Some apps have a paid subscription model. From the 41 apps, five of them belong to this category. The suppliers of these apps were contacted and asked to take part in the evaluation of this research by giving temporary access to the apps. All reacted positively and indicated that they wanted to cooperate.
4. Paid extra device needed and accessible to everybody: Most apps of the initial sample belong to this category, 18 in total. Most of these extra devices need to be connected to a smart meter and enable real-time monitoring instead of the 24-hours delayed monitoring that is provided by the apps in the other three categories. Six of the app providers did not react at all, even after trying to get in touch several times by phone and email. Three of the suppliers indicated that the apps are no longer delivered to new customers because the apps have already been stopped or will be stopped shortly. Nine of the apps can be evaluated thanks to the suppliers who send a device or gave access to a demo account.

In total, access to 30 of the 41 apps is obtained, and these 30 make up the final sample of this research. In the following step, it is explained how the apps are being evaluated on the use of game design elements.

3.1.2 Evaluating energy apps

For the reasons of the clearly described goals and requirements for gamified energy apps, and well-researched game elements and categories, the framework of Alskaif (2018) is leading in creating an evaluation method. The aim is to draw up an easy to understand formula that evaluates the use of different game design types and elements, and provides a score that enables a comparison between the different apps. Within this conceptual framework, the overall goal is described as engaging residential customers with the help of energy apps. As mentioned in section 3.1.1, there are three distinct categories, energy efficiency, PV self-consumption, and demand response, and this research is focused on energy efficiency within households. The framework describes the requirements;

awareness, knowledge, privacy, incentives (economic, environmental, and social), ease of use, self-control and fun, to enable this engagement. The necessary components to meet these requirements are the technical system (smart meter, energy management system, network, software and the app), game design elements, and a value proposition. The technical component is used in all the apps of the final sample, and the value proposition differs between the categories, for example, monthly paid subscriptions or providing customers with an app to give insights in their consumption and invoices (category two). For this research, it is evaluated to what extent the selected apps make use of the third component; game design elements. For this evaluation, the game design elements described in the theoretical framework are used.

The evaluation formula

With the elements and types to which these elements belong described, the apps of the final sample can be evaluated. For this evaluation, the apps are thoroughly examined. An excel sheet is created with all the apps in the most left column and the five-game design types in the following five columns. All the elements that are implemented in a certain app are noted. This provides an overview of the number of elements and types that are used within an app. To evaluate how gamified an app is, a formula is created that includes the game design types and game design elements. The essential factor in this formula is the use of the elements. Therefore, every element that is used provides a score of 1 point (nE). This makes that the end score is mostly based on the number of elements implemented.

Next to the elements, it is assumed that using multiple types of elements creates a more gamified app. For example, if an app uses all information provision elements but none of the other types of elements, it still receives the full points for the elements. However, the elements used are so narrow focused, that an app with the same number of elements but spread across different game design types is more gamified. Therefore, the score for the elements is multiplied with 1+0.1 per type of element (T).

There are quite some differences among the same elements, like a competition with only a comparison to an average household, or a competition with scoreboards and extensive overviews of the comparison to particular users. Evaluating the elements on itself can be tricky because it is easily biased by the researcher that carries out the evaluation, it can be based on an opinion. It would be possible to create scorecards for all elements with specifics to distinguish to what extent a particular element has been implemented. However, due to the limited amount of time available for the evaluation, such detailed scorecards are not created for this evaluation. An app simply does or does not make use of an element, although there is made one exception for the statistics element.

Statistics can be seen as the most crucial element in an energy app. Without statistics, a user cannot get the necessary insights on energy consumption. However, these statistics can be very basic, only showing numbers on energy consumption, or more enhanced by using these statistics to compare the consumption to a previous time period, or giving insights into, for example, costs and CO_2 emissions. Showing these extra insights could positively impact users to lower their consumption. Therefore, an extra score of 0.1 is added if the statistics show more than purely numbers and graphs on the amount of energy that is consumed (S). If the app is evaluated the following formula is used to grant the app a game design score;

$$G = nE * (1 + T + S)$$

Where;
G = Game design score
nE = Number of elements implemented
nT = Types of elements implemented (0.1 – 0.5)
S = Score for statistics (0 or 0.1)

With five types and 18 elements that the apps are evaluated on, the game design score can differ between 0 and 28.8 points. For example, if an app makes use of basic statistics (information provision), messages (information provision), and a competition (social connection) element, the game design score for this app is; 3*(1+0.2+0) = 3.6.

3.1.3 Selecting energy apps for the experiment

To select the energy apps that are used in the experiment, it is decided that the apps need to make use of as many game design elements as possible. Therefore, the apps need to score at least five points in the evaluation. Next to this game design score, a few other requirements are listed to ensure that the most suitable apps are used.

- Participants can use the app, regardless of their energy supplier. This excludes the second category of the evaluation (see section 3.1.1).
- Participants can use the app without having to pay during the research period.
- The app needs to show statistics on energy consumption without the need to enter meter readings manually.

Considering these requirements and the results of the evaluation as showed in appendix A, six apps are suitable for the experiment, two from the first category, one of the third category and three of the fourth category. To enable a comparison between the categories, the final selection of apps will consist of one app from all three remaining categories. Since the apps in the first category are free to use, the app with the highest score is selected. For the other categories, the suppliers are contacted to ask if they want to sponsor this research by supplying the app for free during the experiment. An app provider from the second category responded positively and was willing to sponsor the use of the app for four months. For the last category, the provider of one app responded positively and told the researcher that they were looking into the possibility of adding more game design elements. Therefore, this experiment is a perfect match. They sponsored nine devices for the experiment, and by doing so, they finalized the selection of apps that are used in the experiment.

3.2 The experiment

As for the evaluation of the apps, the experiment is executed in three steps; selecting housing, conducting the experiment, and collecting and processing the obtained data. These steps together will take about four months. Since not all households make use of natural gas for heating (some use district heating), it is decided that during the experiment only the electricity consumption is monitored, and this electricity consumption data is compared to electricity consumption data of the same period last year.

3.2.1 Selecting housing

Other studies show that a social connection or pressure, competition and cooperation are successful game design elements to lower energy consumption (Bang et al., 2009; Geelen et al., 2012). This suggests that the social connection between participants can have an effect on the outcome of this experiment, and that only researching one type of social connection can create biased results. To overcome this potential bias, and to be able to compare the social aspect in lowering energy consumption, three groups with different social connections are created; a group of neighbors, a group of friends of the researcher, and a group of people with none or a negligible social connection. The first group was formed with the help of a housing cooperation. They selected an apartment building from which the residents received an email asking them to participate in this research voluntarily. To find participants for the second group, friends of the researcher were contacted to see if they wanted to cooperate. The third and last group was formed by asking around in the network of the researcher to see if they could help in finding people that were willing to participate.

To be sure that the potential participants and their houses are suitable for the experiment, a few initial requirements were created.

- Smart meter: Most apps in the evaluation need to be connected with a smart meter to track energy consumption. Therefore, the participants must have a smart meter installed.
- Same house for a year: It is essential that the electricity consumption during the experiment is compared to the electricity consumption of last year. Due to this comparison, the participants need to live in the same house for at least a year.
- Historical data: For the same reason as the previous requirement, the participants need to have access to their historical data on electricity consumption for the months December 2018, January 2019, and February 2019. This data can be obtained through monthly reports sent by their suppliers.

After emailing and calling potential participants, a total amount of 45 people replied positively to the request to participate. Nineteen people from the first group, 15 from the second, and 11 from the third group. As described in the second step of the research of Morschheuser et al. (2017) in the theoretical background, it is crucial to analyze the participants to define the user needs, motivations, and behavior in the current system. Therefore, a survey is created to get more information on the socio-economic of the participants, and to see if they have an affinity with sustainability and lowering their energy consumption. Also, they were asked to provide their historical electricity consumption data and a picture of their smart meter. Unfortunately, this survey resulted in a substantial drop out due to no response at all, not obtaining the needed data, or in the end not having a smart meter. This resulted in 25 participants in total, eleven participants in the first group, seven in the second group, and seven in the third group.

As mentioned in section 3.1.3, three apps are used in the experiment. It is desirable that at least three participants in all groups use the same app. With this in mind, the surveys are looked over one more time to find the most suitable candidates that did not completely fit the mentioned criteria. Two participants of the second and one of the third group were selected that only had historical data for the months January 2019 and February 2019. For the third group, also one participant was selected without historical data, but who was really enthusiastic about the research. Thanks to these extra participants, the final sample size for the experiment is 29. eleven in the first group, nine in the second group and nine in the third group.

3.2.2 Setup of the experiment

For the experiment, a period of three months is available starting from December 2019. Within this timeframe, the first step is to test whether the apps on themselves can cause participants to lower their electricity consumption, and what elements are most useful to accomplish this goal. Secondly, extra game design elements are added to the experiment to see what effects particular elements have.

As described in section 3.2.1, before the start of the experiment, a first survey is used to collect data on the willingness of participants to lower their electricity consumption. This information is of high importance due to the fact that if the experiment shows significant differences between participants, this can possibly be explained by the data provided in the first survey. For instance, they can be more intrinsically motivated if they want to play an active role to lower their impact on the environment, or when they can improve their living standard if they lower their energy costs.

In all groups, at least three participants are selected to use a particular app. The selection is done randomly to make sure the selection is not biased. To start the experiment an account is created for all participants, and a user manual is created to show them what features the three different apps have to offer. A few days before the experiment, all participants received an email with this user manual and their login credentials. On the first of December 2019, a kick-off email was sent to let them know the experiment officially started, and in short, why this research is important concerning energy and sustainability, as mentioned in the introduction of this research, to see if energy apps and game design elements can contribute to lower the electricity consumption of households.

The first month is used to test if the app on itself affects the electricity consumption of participants. Of course, they are free to ask any questions about the apps or electricity in general (by phone or email). However, no active interventions are planned so a conclusion can be drawn on purely using the app. Starting from January 2020, interventions are planned to test extra game design elements. Taking into account the game design framework of Morschheuser, et al. (2017), the goal of the interventions is to lower electricity consumption. With the evaluation of the apps, the project preparation phase is completed. The first survey is used for the user analysis phase, and for the ideation phase, previously done experiments were helpful to come up with ideas for these interventions. A difference is made between the elements added in the second month, and the ones added in the third month. The second month will include elements that will likely trigger intrinsic motivation, and in the third month elements with more extrinsic motivation triggers, such as prizes that participants can win, are added to the experiment .

Based on the literature review in the theoretical background, and taking into account the findings of those previous experiments, a schedule of interventions with game design elements that trigger different kinds of motivation perspectives is created. This schedule is shown in table 3.

Date	Intervention type	Intervention explained	Motivation persp.
01-12-2019	Information provision	Kick-off email that explains why saving energy is important.	Emotion (intrinsic)
23-12-2019	Information provision	Happy holidays e-card explaining the effect of saving energy on the environment and motivating them to take control of their energy consumption.	Emotion + Self-determination (intrinsic)
05-01-2020	Information provision - Messages - Feedback	Data driven report on their electricity consumption in December; how much electricity is consumed in comparison to last year, best day and worst day of the month and most important insights.	Behaviorist (intrinsic)
06-01-2020	Social connection + information provision	Starting a WhatsApp group for all three groups that they can use to ask question.	Trait (membership) (Intrinsic)
06-01-2020	Social connection - Collaboration	An email that the goal of this research is to lower the consumption with at least 7% per group. Mentioning that if this goal is reached the whole group receives a city challenge app gift card (city games).	Cognitive (Collaboration) (intrinsic + extrinsic)
06-01-2020	Information provision	Send a weekly tip in the WhatsApp group on how to save electricity.	Trait + Behaviorist (intrinsic)
02-02-2020	Information provision - Messages - Feedback	Data driven report on their electricity consumption in January. Adding badges for completed challenges and a comparison of energy consumption to other participants.	Emotion + Self-determination (intrinsic)
02-02-2020	Social connection - Competition	Introducing the competition for the February and announcing that the three best performing participants receive a gift card for a movie theater.	Trait (extrinsic + intrinsic)
14-02-2020	Social connection - Competition	Sending a leaderboard that shows how participants are doing in the competition so far	Trait (extrinsic)
26-02-2020	Information provision + social connection	Last message to motivate why lowering energy consumption is important, and a link to the final survey of the research.	Emotion + self-determination (intrinsic + extrinsic)
02-03-2020	Information provision + rewarding system	Thank you message with the results of the competition and the link to the final survey.	Behaviorist (intrinsic)

Table 1. Game design intervention planning

At the end of the experiment, a final survey is sent to all participants. They are questioned about their experience using the app and the effect of the added game design elements. This final survey is valuable because a lower or higher electricity consumption does not necessarily mean that the app or the added game design elements are responsible for that particular result. The experience, perception and the adjustments in behavior of participants are essential to draw conclusions. Also, in the first survey the participants were asked about major changes in their household situation compared to the same period last year, that could possibly explain an increase or decrease in electricity consumption. If during the experiment the apps show striking consumption data (like no consumption, or extreme high consumption), the particular participant is asked about the reason behind these extreme values. With this data, it is more reasonable to draw conclusions on the effect of the app and added game design elements

3.2.3 Collecting and processing data

Because the researcher created all the app accounts for the participants, it is relatively easy to collect the data during the experiment. This data is added to an excel file that also shows the electricity consumption data of the previous year. This creates the opportunity to see if anything changes in their electricity consumption per month, and over the whole three-month period. Also, the electricity consumption per group and app can be calculated to see if there are notable differences between the three social connection groups; neighbors, friends, and no connection, or the different apps.

This data is cross-checked with the data provided by the participants in the first and final survey. These surveys are created with Google forms, and all participants received links to get access. Thanks to this easy to use system, the researcher has access to the data that is provided by every participant individually, and a summary of all participants of a certain group. It is investigated if results that stick out are explainable and what this explanation could be. With the combination of the surveys and data on electricity consumption, qualitative conclusions can be drawn on the effect of the experiment, and questions can be asked if this same effect can be expected when implemented in larger-scale experiments.

4. Results

The results of this research consist of two parts. First, the apps evaluation shows if available energy apps in the Netherlands make use of game design elements and to what extent. Second, the experiment elaborates on the fact if the use of these apps and implemented game design elements have an effect on the energy-related behavior of their users, and if adding extra game design elements changes anything to this.

4.1 Evaluation apps

In appendix A, the full evaluation of all available energy apps is shown. As can be seen, almost all apps implemented at least one game design element, except one, this app only shows the electricity consumption on a particular moment in time. The results of the evaluation are presented in three parts. First, the game design types that are used in the apps. Second, the game design elements that the apps implemented. Third and last, the game design score of the apps.

4.1.1 Game design types

Taking into account that presenting statistics in compelling graphs or tables is a game design element, 'information provision' is by far the most used game design type. However, statistics is the fundamental element of an energy app, without this element, the app can be considered worthless to most people. Only 12 of the 30 evaluated apps used another element from this game design type. After information provision, 'user interface' and 'social connection' are the most applied game design types. They are used in respectively 16 and 14 apps. The 'rewarding system' and 'performance status' game design types are only utilized in four and three apps, respectively. These game design types can trigger the extrinsic motivation of users, that is why the available apps on the Dutch market seem to aim mainly at the intrinsic motivation of end-users. Figure 2 shows how many apps use a particular game design type and how many game design types are used in the apps.

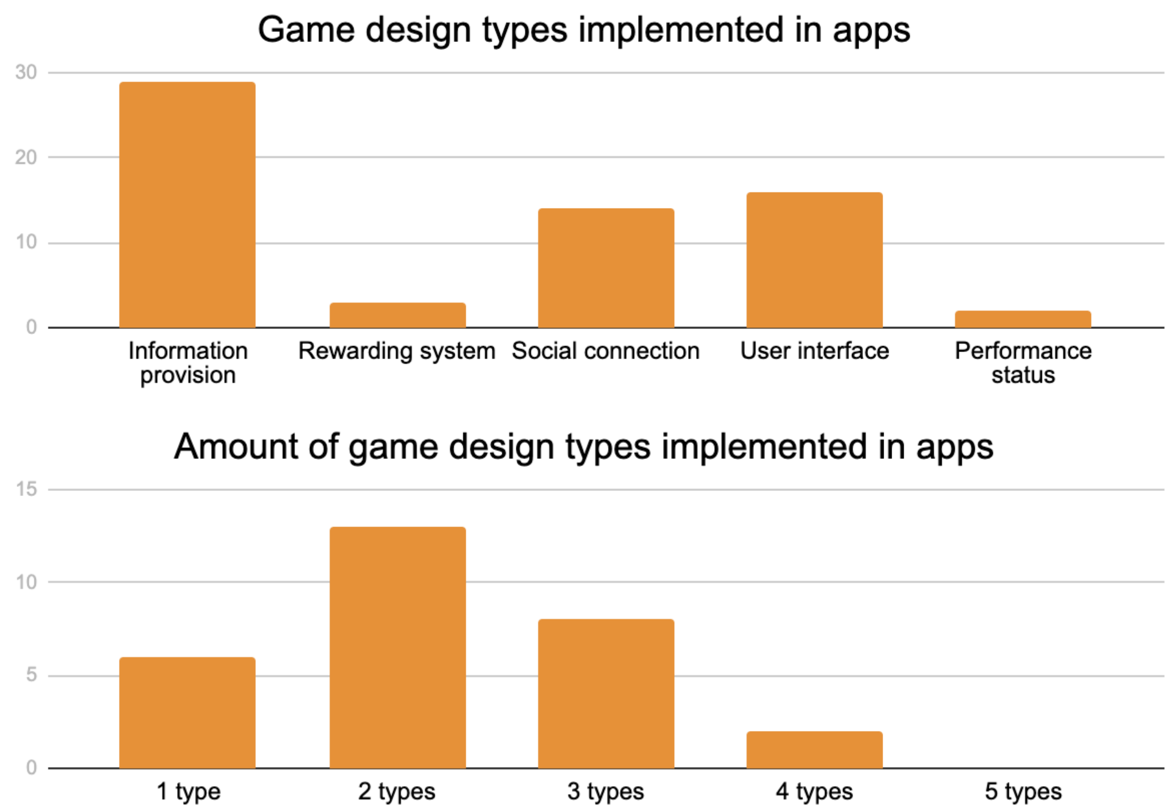


Figure 2. Game design types used in evaluated apps the Dutch market

As can be seen, 67%, 20 out of the 30 apps evaluated, use two or fewer game design types. None of the apps use more than four types, making it evident that game design is not yet fully integrated into the energy apps available in the Dutch market.

4.1.2 Game design elements

The amount of game design elements implemented in an app differs between zero and seven per app. As mentioned before, the statistics element is implemented in 97% of the apps, making it the most implemented element. The 30 apps in total implemented 88 game design elements, with an average of 2.9 elements per app. Of the 18 available elements in the framework of AlSkaif, et all. (2018), 12 are used in at least one of the apps. After the statistics element, the competition (12 apps), tips (nine apps), and progress bar (nine apps) are the most used elements. The elements, discount on the electricity bill, virtual currency, collaboration, leaderboard, virtual currency, points, and levels, are not used at all. In figure 3, the number of apps using a particular element, and the number of elements used in the apps are shown.

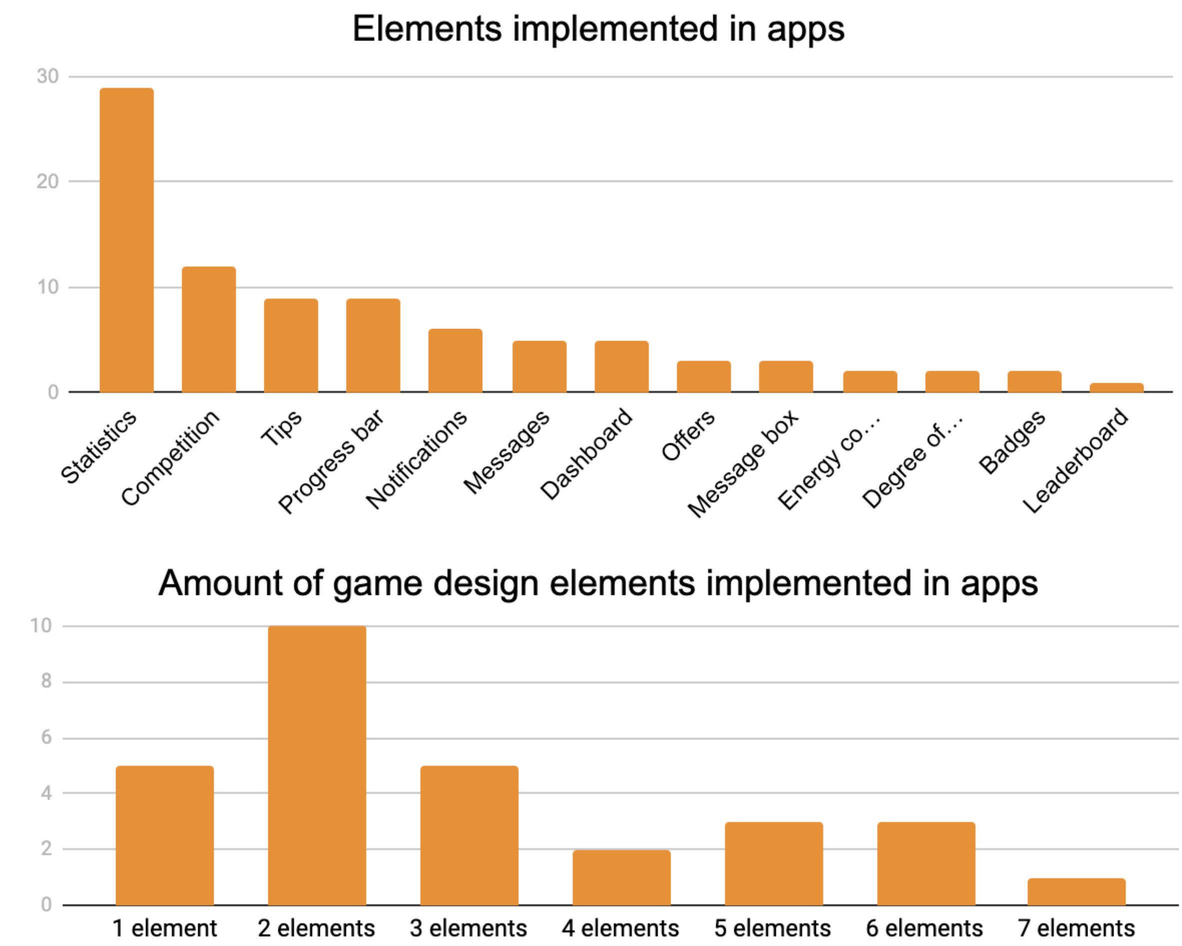


Figure 3. Game design elements implemented in apps

Fifteen of the 30 evaluated apps implemented two or fewer game design elements. As mentioned in section 3.1.1. the apps can belong to four different categories; Free and accessible to all, only available to customers of a certain supplier, paid and available to all, and only available with a paid extra device. The average amount of elements used in an app differs among these categories with respectively; 2.6, 2.9, 3, and 3.2 elements per app. Three of the four apps that use more than five elements belong to the category 'only available with a paid extra device'. Overall, it can be concluded that the paid apps implemented more game design elements than free apps.

Nonetheless, the difference between the categories is minimal, and with the average of 3.2 elements in the category 'only available with a paid extra device' , it may be concluded that implementing game design elements has probably not been a top priority for app creators. Therefore, adding more game design elements in an experiment could be interesting to both science and app creators.

4.1.3 Game design score

The highest score obtained in the evaluation is 9.8 out of 28.8 points. This score is obtained by the app Toon, developed by the energy company Eneco (Eneco, n.d.). This app implemented seven game design elements belonging to three different game design types. Due to the extensive statistics element, it also earned a 0.1S score. Toon belongs to the 'only available with a paid extra device' category, and also the second and third best scoring app, Enelogic with 9.0 points (Enelogic, n.d.), and Smappee with 8.4 points (Smappee, n.d.) belong to this category. As mentioned in section 3.1.2, the evaluation formula does not take into account to what extent an element is used and this is especially noticeable with the competition element. Most apps that use this element only show a simple comparison to an average household. If the formula took into account to what extent an element was implemented, it is expected that the apps from the 'only available with a paid extra device' category would score even higher.

The average app uses only two out of five available game design types, and two out of 18 available game design elements. The average score differs with respectively 4.0 points (Free and accessible to everybody), 3.7 points (only available for customers of a certain supplier), 3.9 points (paid and available to everybody), and 4.3 points (Only available with a paid extra device). It shows that after the apps that are only available with a paid extra device, the free apps have the highest game design score. This is interesting because 'free and accessible to everybody' apps on average implemented the least game elements. The fact that they spread the elements over more game design types causes this higher game design score. Overall, the difference is minimal and therefore it may be said that all apps score pretty low when it comes to adding game design.

The overall average score is 3.8 points, while the highest possible score is 28.8 points. As in the evaluation research in the U.S (Beck et al., 2019), it can be concluded that this could be considered as a first sign of implementing game design in energy apps. However, the amount of elements used is pretty low. Therefore, it would be interesting to see what additional elements have the most effect on users in the experiment to conclude how game design could help to reduce the electricity consumption of households.

4.2 Experiment

All participants that were selected for the experiment finished both the survey before the experiment on their personal characteristics, and after the experiment on their experience during the experiment. However, one participant could not make use of an app, even after trying multiple apps (U-meter and Smart dodos). This is probably since this participant used the option not to share data with third parties when his smart meter was installed. Luckily, this was the only participant that could not provide historical data. This participant did not try the Smappee because all sponsored devices were already in use by other participants.

In the experiment results section of this report, first, a summary of the participants demographics obtained from the first survey is presented. After these insights, the data on electricity consumption, and the most valuable insights of the surveys are presented per month of the experiment. Lastly, the overall results are compared to the experience of the participants (last survey) and their initial demographics.

4.2.1 Demographics participants

Before the start of the experiment, all participants received the first survey on their household specifications, socio-economic status, and their affinity with sustainability and energy. Based on the most noteworthy insights of each group, the following demographic summaries are created.

Neighbors

seven out of the eleven participants in this group live on their own, and the other four live with their partners, one of these couples did not live together a year earlier. They are all well-educated, with all having at least a bachelor's degree. All of the participants, and their partners, work at least 32 hours a week. They spend between 20% and 50% of their household income on rent. Almost half of these participants choose a particular energy supplier solely based on price, and nine out of eleven participants did not make use of an energy app before this experiment, and so far, not even considered using one. About half of them think sustainability is an important issue. Nevertheless, less than 20% is actively working on lowering their impact on the environment or advocating this importance to others.

Friends
Eight out of nine participants have a two-person household, and four of them have a three-person household. In three of these households, the number of people living in the household changed during the last year. Eight out of nine participants are well educated with at least a bachelor's degree. All participants work a minimum of 25 hours a week, and most of their partners also work a minimum of eight hours a week. Seven of the participants spend between 10% and 30% of their household income on rent or mortgage and the other two between 30 and 40%. The reason to choose a particular energy supplier differs among all of them, from the lowest price, the offer of green energy, to the gift they received as a new customer. Already five out of nine participants used an energy app before this experiment. Three used the free energy app from their supplier and two the app Toon. Six of them said that they are aware of their energy consumption in comparison to an average household. Within this group, eight people think sustainability and energy is an important topic, and seven of them consider it an important theme during political elections. 66% of the group mentioned they are actively working on reducing their impact, and five of them even advocates the importance of the subject to others.

No social connection
From the nine participants, seven live in a two-person household and the others live alone. All of them are well educated with at least a bachelor's degree. Seven participants of this group have a job, and the remaining two are retired, this is also the reason why in those households more time is spent at home compared to last year. Another reason is illnesses of two of the participants. Four out of nine spend less than 10% of their household income on rent or mortgage and the rest spend between 10% and 40%. The reason for choosing a particular energy supplier differs among them. Five out of nine participants already use an energy app and four of them use the app that was made available by their energy supplier. None of the participants ever paid for using an energy app, however, three participants are willing to pay a one-time amount for the right app. Five participants would lower their consumption if they knew they use more than an average household and think energy consumption is important when buying new appliances. All participants in this group think sustainability is an important issue and seven of them think it is an important theme during elections. However, only five are working on reducing their impact and only one person also advocates this to others.

4.2.2 First month

During this month, only 25 participants are included in the calculations due to missing data on electricity consumption in December 2018 of some participants. The most interesting results are the average savings. It determines if a consumer can decrease its consumption with the help of an energy app. These results are presented later on in this section. First, the data on the total consumption of the participants combined is shown in table 2. An overview is presented for the different groups and app users in this research. This table shows the number of participants that showed a decrease in electricity consumption, the total difference in kWh compared to the period last year, the increase or decrease in consumption per group, and the best and worst-performing participant.

Group	**Participants that lowered consumption**	**Total consumption of each group Dec 2018 (kWh)**	**Total consumption of each group Dec 2019 (kWh)**	**Difference in consumption of each group**	**Best participant**	**Worst participant**
Neighbors	7/10	1.076	949	-11.8%	-30,9%	+25%
Friends	2/8	1.637	1.877	+14.7%	-12,4%	+25,6%
No connection	2/7	1.789	1.635	-8.6%	-25,5%	+11,4%
Total	11/25	4.502	4.461	-0.9%	-30,9%	+25,6%
App	**Participants that lowered consumption**	**Total consumption of each group Dec 2018 (kWh)**	**Total consumption of each group Dec 2019 (kWh)**	**Difference in consumption of each group**	**Best participant**	**Worst participant**
Smappee	5/9	1.331	1.236	-7.1%	-20,9%	+25%
Smart Dodos	4/10	1.935	1.921	-0.7%	-25,5%	+25,2%
U-meter	2/6	1.236	1.304	+5.5%	-30,9%	+25,6%
Total	11/25	4.502	4.461	-0.9%	-30,9%	+25,6%

Table 2. Results of total electricity consumption in the first month of the experiment per group

In the first month, 44% of the participants experienced a decrease in their consumption, and the total consumption of all 25 participants together was 4.461 kWh, compared to 4.502 kWh in December 2019. This results in a decrease in total electricity consumption of 0.9%. The neighbors group clearly showed the most favorable results with a total decrease in consumption of 11.8%. The friends and no connection group had a total increase of 14.7% and a decrease of 8.6% respectively. For the different app users, the result of Smappee is most favorable with a total decrease in consumption of 7.1%. Smart dodos users showed a decrease of 0.7% and for U-meter the result is least favorable with an increase of 5.5%. As mentioned before in this section, the most interesting result is the average increase or decrease of the participants, these results are presented in figure 4.

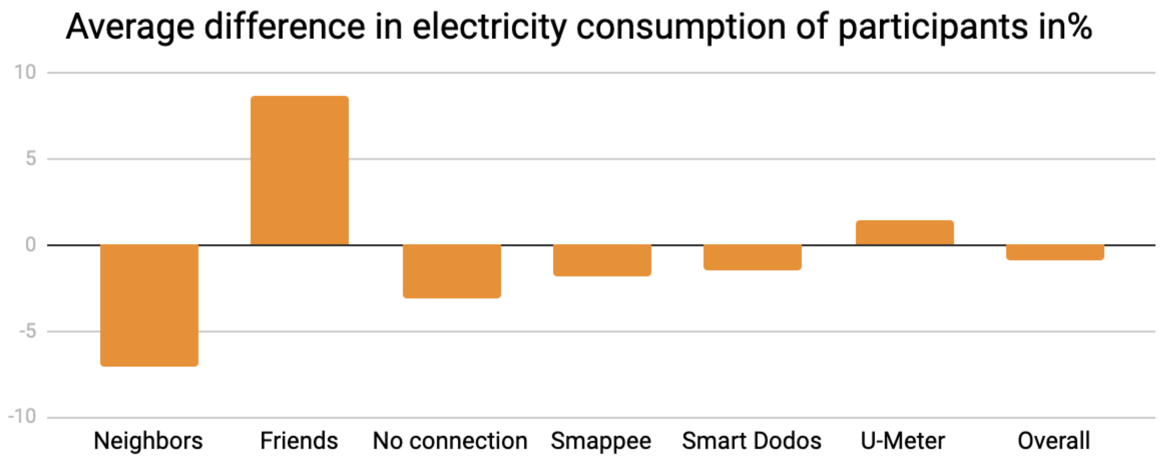


Figure 4. Average increase or decrease in consumption for each group during the first month of the experiment.

The overall bar shows an average decrease in consumption of 0.9% for all 25 participants combined. Nevertheless, as mentioned before, most participants used more electricity than they did in December 2018. Looking at the different groups and apps, the clear winner is the neighbors group, with an average decrease of 7.1%, and Smappee users with a decrease of 1.8%. The friends group increased their consumption with 8.7%, and U-meter users with 1.4%. The group without a social connection showed an average decrease of 3.2%, and Smart dodos users also showed a decrease of 1.5%. Solely based on these results of the different groups, no conclusion on the use off an energy app can be drawn. Therefore, the survey at the end of the experiment can provide more insights.

As mentioned in sectioned 4.2.1, within the neighbors group only two participants used an energy app before the experiment. In the final survey, six of the eleven participants in this group mentioned that they only opened the app once at the start of the experiment. All these participants had the U-meter or Smart Dodos app. All three Smappee users in the neighbors group opened the app every two weeks. In the other groups, most participants used the app at least once every month. As in the neighbors group, the Smappee users in the friends and no connection group opened the app every two weeks, two participants even opened Smappee more than once a week. This is a clear sign that Smappee is more appealing to the participants and this app also shows the best result.

In all groups, about 70% of the participants did not change the personal settings in the app (price for electricity and household type). In the email that they received at the beginning of the experiment, including a manual of the app, it was clearly explained how they could change the settings and why this is recommended. They also received feedback that the settings were still not changed. These settings enable the apps to compare the household with similar households (competition element), and give the user insights into the costs for electricity use. That these participants did not change the settings could mean that these features were not interesting to them, or that they did not want to spend too much time in the app. Nevertheless, they were still able to see the statistics on their electricity consumption.

One of the Smart dodos users mentioned that the app was not mobile friendly and that it was difficult to log in from a mobile phone. She said she would use the app more often if she could easily access the app on her smartphone. A U-meter user mentioned that energy apps needs to be kept simple so that they are easy to use. Overall the participants evaluated Smappee with 7.9 out of 10 points, Smart dodos with 6.6, and U-meter with 6.3 points. Most participants gave 7 and 8 points to all apps. The lowest scores were given by two participants that already used the Toon app (app with the highest game design score) before they started using Smart Dodos and U-meter. Both Smappee and Toon are very easy to use and have a mobile app. Both apps provide real-time insights on electricity consumption and these extra insights could be the reason why users look at the app more often.

Examining the game design elements implemented by the apps, as expected, the statistics element was most interesting to users since this is the fundamental element of an energy app. Most participants did not even make use of the other implemented elements. Almost none of the app users received messages or notifications because users needed to adjust the settings to receive these messages and notifications. The participants that did receive messages or notifications thought these were not valuable to them. The comparison to similar households was available in all three apps, a simple type of a competition element. This competition element was only used by six participants and only one of them thought this was a valuable feature. Finally, when asked if the participants were more aware of their energy consumption and if they adjusted their behavior due to the app, about half of the participants reacted positively. Quite remarkable when the other answers about the implemented game design elements are taken into account. This result makes it even more interesting to see what the effect of adding extra game design elements is.

4.2.3 Second month

During the second month of the experiment (January 2020), all the 28 participants that handed in data on their electricity consumption are included in the calculations. In this second month, extra game design elements were added to trigger the participants intrinsic motivation (i.e., messages, feedback, tips, and collaboration). Just before Christmas, a holiday e-card was sent, including an explanation of the positive impact on the environment if participants lower their electricity consumption (see Appendix B). At the beginning of the month, participants received a first data-driven energy report and an overall target was set to save 7% electricity in each group (i.e., a shared goal game design element). To help the participants reach this target, a WhatsApp group was created to enable them to ask questions related to electricity consumption. These groups were also used to send weekly tips on how to conserve electricity. These extra intrinsic motivating elements are presented in Appendix B.

As in section 4.2.2, first, the data on the total consumption of the participants combined are shown in table 3. The results of the average increase or decrease in consumption of the participants in the second month are presented later on in this section.

Group	**Participants that lowered consumption**	**Total consumption of each group Jan 2019 (kWh)**	**Total consumption of each group Jan 2020 (kWh)**	**Difference in consumption of each group**	**Best participant**	**Worst participant**
Neighbors	9/11	1.188	971	-18.3%	-51.3%	+4.6%
Friends	2/9	1.979	2.168	+9.6%	-35.6%	+27.4%
No connection	4/8	1.848	1.887	+2.1%	-28.6%	+88.9%
Total	15/28	5.015	5.026	+0.2%	-51.3%	+88.9%
App	**Participants that lowered consumption**	**Total consumption of each group Jan 2019 (kWh)**	**Total consumption of each group Jan 2020 (kWh)**	**Difference in consumption of each group**	**Best participant**	**Worst participant**
Smappee	7/9	1.349	1.176	-12.8%	-35.6%	+13%
Smart Dodos	6/10	2.003	2.017	+0.1%	-51.3%	+27.4%
U-meter	2/9	1.663	1.833	+10.4%	-28.6%	+88.9%
Total	15/28	5.015	5.026	+0.2%	-28.6%	+88.9%

Table 3. Results of total electricity consumption in the second month of the experiment per group

The results show that 54% of the participants lowered their consumption, and again the neighbors are the best performing group. In total, they decreased their total electricity consumption with 18.3%. An impressive result when compared to the increase of 9.6% in the friends group, or the increase of 2.1% in the group without a connection. Another remarkable result is that the extremes of the best and worst-performing participants increased compared to the first month of the experiment. This result is partly explainable by looking at consumption graphs in the apps. The participant that consumed 51.5% less electricity was not home for two full weeks. The participant with the 88.9% increase told the researcher that he spent more time at home due to illness, and a new housemate that works from home moved into the household. For the different app users, Smappee again shows the most positive result with a decrease in total consumption of 12.8%, compared to the 0.1% increase of Smart Dodos, and 10.4% increase of U-meter users. Figure 5 presents the average increase or decrease in consumption of a participant in the second month of the experiment.

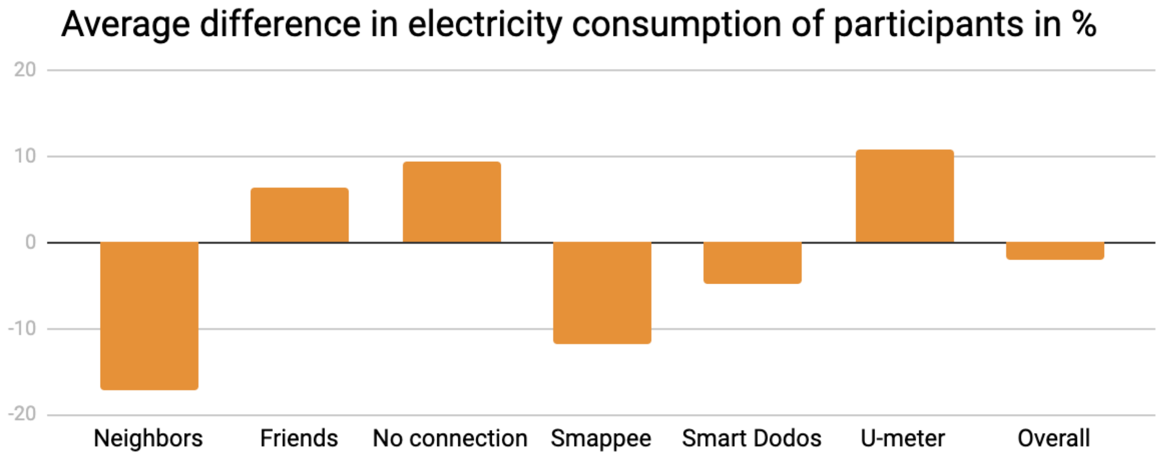


Figure 5. Average increase or decrease in consumption for each group during the second month of the experiment.

As can be seen, while the total consumption increased by 0.2%, participants decreased their consumption on average by 2.1%. Again the neighbors group shows the best result with an average decrease of 17.2%. Even if the best performing participant is excluded, the decrease in consumption is still 13.8%. The friends group shows an increase of 6.4%, and the no connection group an increase of 9.4%. However, if the worst performing participant is excluded from the no connection group, the average is a 2% decrease in consumption. Smappee users decreased their consumption with 11.9% on average, compared to a decrease of 4.8% for Smart Dodos, and an increase of 10.9% for U-meter users (decrease of 1.1% if the same worst performing participant was excluded). If the best and worst-performing participants are excluded from the overall result, the average decrease in consumption is 3.6%. Next to the data on the participants electricity consumption, the final survey included questions that can give more insights about the added game design elements.

The participants in all groups thought the Christmas e-card was a nice gesture. In both the friends and no connection group, four participants in each group said it was a trigger to think about their electricity consumption. For the neighbors group, this was only the case for two participants. All except two participants opened the first data-driven report, and in general, about half of them fully read the report. In the neighbors group, six out of eleven participants thought the report gave them valuable insights on their consumption, but only three thought the feedback was interesting for them. In the friends group, four out of nine participants mentioned that the report gave them valuable insights, and only two told that the feedback was valuable. In the group without a connection, this was the case for respectively seven and four out of nine participants. More than half of the participants mentioned that the feedback would be more interesting if it was more based on their personal situation. Due to the limited data available in the apps, this was unfortunately not possible.

Only one person asked a question in the WhatsApp group during the whole second month. In the neighbors group three, the friends group six, and no connection group five participants thought having a WhatsApp group to ask questions was valuable. A few of them also mentioned that the fact that there were more people in the group prevented them from asking a question. The results of the tips sent in the WhatsApp group were most promising in the second month. All participants answered questions in the final survey about which tips they found interesting and which tips they used to lower their electricity consumption. All participants found at least one of the four tips interesting and 83% of them also used at least one of the tips to lower their electricity consumption. Twenty of the 29 participants thought that a weekly tip about how to save electricity was helpful, seven were neutral about the usefulness, and only two participants thought sending tips is not useful to change their behavior. It means that 69% of this sample group thinks that receiving tips about energy conservation could be a significant game design element.

Only 31% of the participants were positively impacted by the shared goal to save 7% electricity. In general, the final survey reveals that the neighbors group is most doubtful about this added game design element. This is striking because this group showed the most favorable result. Another striking fact is that the shared goal element more influenced the participants in the other groups, whose electricity consumption increased compared to the same period last year. This could mean that they have a false sense of 'doing good' and possibly reveals a limiting or even harmful effect (Hyrynsalmi et al., 2017).

4.2.4 Third month

During the third month of the experiment (February 2020), data of 27 participants was included in the calculations. One of the participants that handed in her data had an electricity consumption of zero during February last year. Therefore, she was excluded in the calculations of the third month. In this third month, the extrinsic motivation was triggered with the help of a competition and prizes the participants could win by conquering the first, second, or third place of each group in terms of lowering electricity consumption compared to the same period last year. The competition game design element was introduced with an email and a WhatsApp message that stated that the participants could win movie theater vouchers if they would win the competition. Halfway the competition, the participants received a message in the WhatsApp group with a leaderboard game design element (see Appendix C). This leaderboard game design element is fully connected to the competition game design element and together they form a more enhanced competition than the already implemented competition element in the apps. As mentioned in section 4.3.2, the simple comparison with other similar households, also considered a competition game design element in the evaluation of this research, was not used at all by the participants.

As in the previous two sections, first, the data on the total consumption of the participants combined is shown in table 4. The results of the average increase or decrease in consumption of the participants in the third month are presented later on in this section.

Group	**Participants that lowered consumption**	**Total consumption of each group Feb 2019 (kWh)**	**Total consumption of each group Feb 2020 (kWh)**	**Difference in consumption of each group**	**Best participant**	**Worst participant**
Neighbors	7/10	964	930	-3.5%	-15%	+14.1%
Friends	5/9	1.675	1.842	+10%	-26.8%	+40.5%
No connection	4/8	1.682	1.546	-8.1%	-44.7%	+60.2%
Total	16/27	4.321	4.318	-0.1%	-44.7%	+60.2%
App	**Participants that lowered consumption**	**Total consumption of each group Feb 2019 (kWh)**	**Total consumption of each group Feb 2020 (kWh)**	**Difference in consumption of each group**	**Best participant**	**Worst participant**
Smappee	6/9	1.181	1.094	-7.4%	-26.8%	+13%
Smart Dodos	8/10	1.852	1.731	-6.5%	-44.7%	+37.5%
U-meter	4/8	1.288	1.493	+15.9%	-15%	+60.2%
Total	16/27	4.321	4.318	-0.1%	-44.7%	+60.2%

Table 4. Results of total electricity consumption in the third month of the experiment per group

In the third and last month of the experiment, 59% of the participants reduced their electricity consumption, this is the highest percentage of all three months. For the first month in the experiment, the group without a connection showed the best result with a decrease in total consumption of 8.1%. The neighbors group showed for the third month in a row a decrease in total consumption, this time a decrease of 3.5%. For the friends group, the total consumption increased with 10%. As in the second month, the extremes are explainable by a holiday and the household where a new housemate moved in. Both these participants belong to the group without a connection. For the different app users, Smappee again shows the most positive result with a decrease in total consumption of 7.4%. The Smart Dodos users showed a decrease of 6.5%, and the U-meter users showed an increase in total consumption of 15.9%. Figure 6 presents the average increase or decrease in electricity consumption of the participants in the third month of the experiment, February 2020.

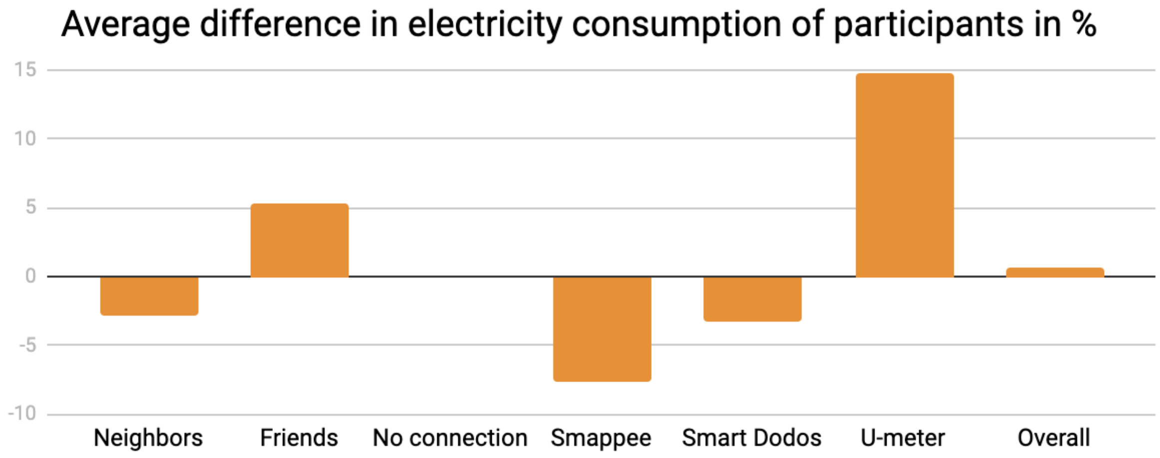


Figure 6. Average difference in consumption third month

As can be seen, the overall result shows an increase in consumption of 0.7% (compared to a decrease of 0.9% in the first month, and a decrease of 2.1% in the second month). Nevertheless, the neighbors group again shows the best result with an average decrease in consumption of 2.9%. The friends group shows an increase of 5.3% (the best result compared to an increase of 8.7% and 6.4%), and for the group without a connection there is no difference compared to last year. If the worst and best participants are left out (they both belong to this group), the no connection group shows a decrease in consumption of 2.6%. Another interesting fact is that the different app users showed a similar pattern as in the previous two months. Again, Smappee is the best performing app with an average decrease of 7.6%, followed by Smart dodos with a decrease of 3.1%, and in last place U-meter with an increase of 14.8% (8.3% without the worst performing participant included). This result makes sense since Smappee users opened the app the most off all participants, at least once every two weeks, and some even more than once a week. The Smart Dodos, and U-meter users did not open the app more than once a month, and as mentioned in section 4.2.2, six participants only opened the app once at the beginning of the experiment.

One of the most promising game design elements in previously done research was the competition. The final survey of the experiment reveals that more than half of all participants feel more motivated to reduce their electricity consumption thanks to the competition element. For the neighbors, six out of the eleven participants responded positively to the added competition element. Within this group, five of the participants also experienced a positive effect when they received the leaderboard after two weeks. Four of the participants even responded in the WhatsApp group while they never interacted in the group before this leaderboard message. This was also the case in the friends group, after they received the leaderboard, four participants responded in the WhatsApp group. Even more interesting is the fact that this group has a stronger social connection, and five out of nine participants mentioned that they were motivated to reduce their electricity consumption due to the competition element. Three of them reacted neutral to it, and only one participant said that it did not influence him at all. Seven participants of the friends group said that the leaderboard also had a positive effect on them. In contradiction to the answers in the survey, three of these participants increased their consumption. However, two of these participants welcomed a baby last year and were already using the app Toon before this experiment. Therefore, this contradiction is explainable, and the fact that the friends group was motivated by the competition and leaderboard elements is an exciting result. It shows that a competition can be extra motivating to people with a strong social connection.

When we take into account the group without a connection, this signal becomes even more substantial. Within this group, only three participants mentioned that they became more motivated to reduce their consumption thanks to the competition element. However, five participants experienced a positive effect thanks to the leaderboard element. It could possibly reveal that seeing the results and names of other participants creates some sort of a social connection between participants. Nevertheless, nobody responded to the leaderboard message in the WhatsApp group, while this was the case in the other groups. Overall 62% of all participants reacted positively on the competition and/or the leaderboard after two weeks. The prizes they could win were motivating to 37% of all participants, showing that winning a prize can trigger the extrinsic motivation of a substantial share of users but not all of them.

Part the participants were intrinsically motivated solely based on competition itself, as in the research of Geelen et al. (2012).

4.2.5 Overall result

In general, the data during the three months shows a recurring image of the different groups and apps. The results in all months reveal that the neighbors group is the best performing group, and Smappee is the best performing app. In figure 7, the average results of the participants over the whole three-months experiment are presented. For the data of all participants, see Appendix D.

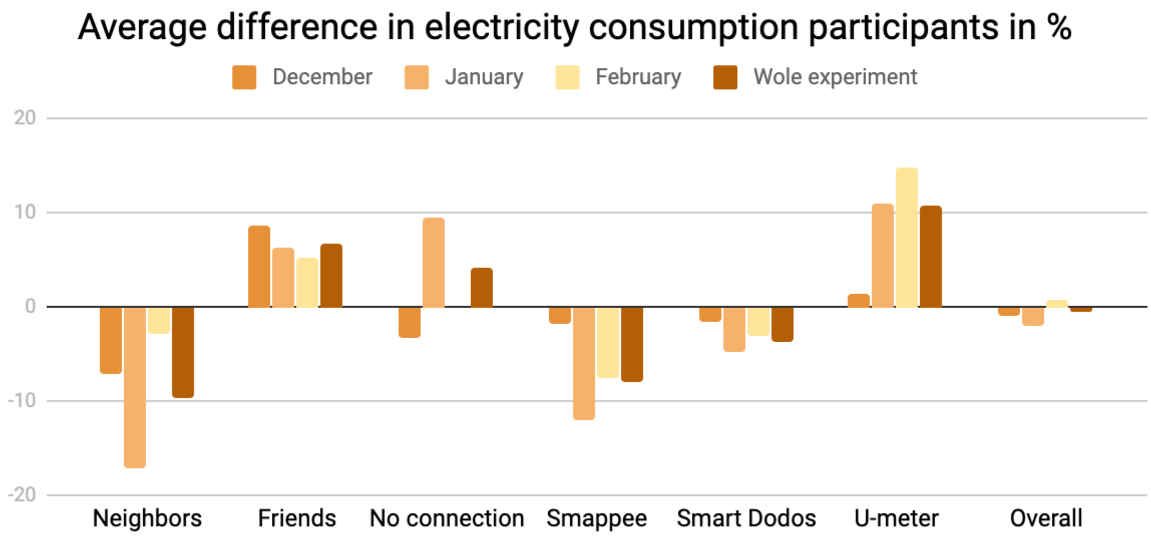


Figure 7. Average difference in consumption for all months separate and combined over the whole experiment

The overall result of the whole experiment, including all participants, shows a decrease in consumption of 0.4%. This is such a small difference that it is impossible to say if the experiment caused participants to consume less electricity. The neighbors group shows an average decrease in consumption of 9.7% and everybody in this group, except one person, showed a decrease in consumption over the whole period of the experiment. It is a motivating result, especially considering that nine of the eleven participants within the group did not use an app before the experiment. However, six participants mentioned that they only looked at the app once at the beginning of the experiment. Therefore, it is impossible to say if the change in consumption can be attributed to the use of an energy app. Unfortunately, both other groups show an increase in consumption. For the friends group the average increase in consumption is 6.8%, and for the no connection group 4.2%. A change in household situations can partly explain this increase. In the friends group, two participants with an increase of 33.2% and 17.3% both welcomed a baby last year. The highest increase in consumption of 73.6% in the no connection group was of a household where a new housemate that works from home moved in. If this person in the no connection group was excluded, this group would show an average decrease in consumption of 5.7%, and this would be a very positive result. This same household also has quite an impact on the overall result, without this participant the overall result would show a decrease in consumption of 3.2%.

In the final survey, 17 out of 29 participants mentioned that overall they have been more aware of their electricity consumption during the research. A sign that energy apps and game design elements can have a positive effect. Looking at the different apps, it is clear that the Smappee users show the most favorable result with an average decrease of 7.9%. The Smappee users opened the app at least once every two weeks, compared to only once at the beginning and once every month for the other two apps. Another sign that using an app with game design elements can help with conserving energy. Only three participants provided feedback on the use of Smappee in the final survey. Smappee has a function that recognizes appliances and one participant mentioned that it would be nice if this feature had more options. Another participant, that was a fan of the added competition element, said that she was missing such a competition in Smappee. The last participant commented that the app is great. These are positive comments and this was not the case for the U-meter app. The U-meter users gave the app the lowest evaluation and this app shows an apparent increase in energy consumption of 10.8% (3% if the participant with an increase of 73.6% was excluded). Unfortunately, only three U-meter users provided feedback in the final survey on the app. They mentioned that an app should be kept simple, that the app did not have many features, and that a real-time monitoring feature was missing. This last comment was made by a participant that already used the Toon app before the experiment, the app with the highest game design score in the evaluation in this research. It is interesting when these comments are compared to the features of the Smappee app. Smappee is really easy to access with a mobile phone app and immediately shows the users how much electricity is consumed in real-time. The app has a simple though attractive design and is really easy to use. This is not the case with the Smart Dodos app that can only be accessed on a web browser. Two participants mentioned this in their feedback on the app, they would use the app more often if it was incorporated into a mobile app. Two other commented that the app could use a more lively and appealing design. As for U-meter and Smappee, the rest of the Smart Dodos users did not give any feedback on the app.

Taking these comments into account, the fact that the Smappee users opened the app the most, that they evaluated Smappee with the highest score, and that the app shows the best results in the comments, makes it clear that Smappee is the best app used in the experiment. Another interesting fact is seen in the friends group, all participants increased their consumption, except Smappee users. Two of the three Smappee users in this group lowered their consumption, and the other one shows the same consumption as last year. Nevertheless, looking at the final survey, this is only partly accountable to the already implemented game design elements. The ease of use and real-time monitoring seem to play an important role. This makes it more interesting to see what participants thought of the added game design elements. When asked if game design elements could help in becoming more aware of their energy consumption, 28 of the 29 participants reacted positively. Another clear sign that users can appreciate game design when it is implemented in the right way.

When asked what added game design elements were most effective, there are a few elements that score the best according to the participants. They had the option to choose a maximum of 3 elements each, and figure 8 shows what they answered.

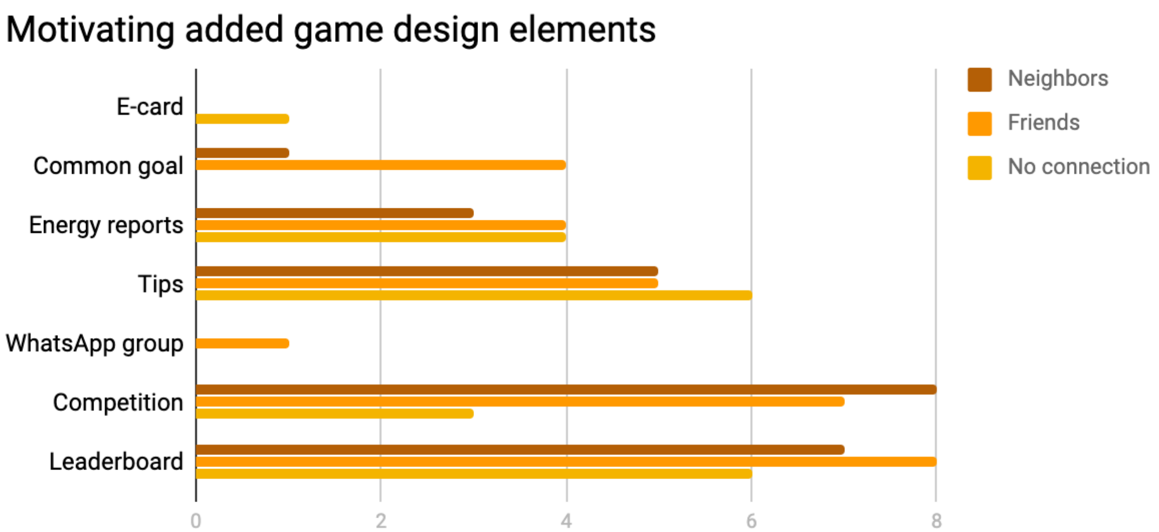


Figure 8. Participants that voted on the most effective game design elements, with a maximum of three per participant.

In general, the competition and leaderboard elements are the most effective, according to the participants. Taking into account that they never used the competition element in the apps, it is noticeable that a leaderboard and a real battle between users are features that make a competition attractive. Even a competition without prizes would motivate part of the participants. However, the prizes they could win motivated eleven of the participants, so a competition with prizes is probably more successful. Next to the competition and leaderboard, the tips element is most useful according to the participants. That participants were positive about this element makes sense because it helps them to better understand their consumption behavior, and the available options to conserve electricity (AlSkaif et al., 2018). Therefore, the 'information provision' elements tips, and messages (feedback), are next to the statistics element essential for an energy app with the goal to conserve electricity. The data-driven reports motivated ten of the participants, still a pretty high score of 35%, and these reports included feedback. Participants mentioned that more personalized feedback would be valuable to them.

Such a data-driven report element, without personalized feedback, is already implemented by many energy suppliers without the use of an energy app. They send monthly energy reports to inform customers about their energy consumption statistics. Maybe the data of these reports, and a bigger sample group could reveal if such elements cause a reduction in electricity consumption. Lastly, the e-card and WhatsApp group are elements that did not get many votes. The common goal was only appreciated in the friends group, again a possible sign that a social connection matters in implementing game design elements.

5. Discussion

The goal of this research is to evaluate energy apps in the Dutch market that promote energy efficiency within households, on the use of game design, and to see what effect game design has on end-users. The evaluation is based on the framework of AlSkaif et al. (2018), and this framework includes five game design types and 18 game design elements. To enable a comparison of how gamified an app is, a simple formula was created that incorporated an evaluation of these types, and elements. This formula provides a basic overview of how many of these energy apps in the Dutch market implemented game design. However, the granted scores are, in some cases, questionable due to the simplicity of the formula. As declared in section 3.1.2, for some game design elements, scorecards could be created to get a more honest comparison between apps, and to evaluate to what extent an element is implemented. Especially the assigned score for a competition game design element is doubtful, while it is the second most used element according to the evaluation. The definition of this element in this research allows a simple comparison to similar households to be seen as a competition. However, end-users probably do not experience such a comparison as a competition. The experiment in this research shows that 18 of the 29 participants appreciated the added competition game design element. At the same time, almost none of them made use of the comparisons to other households that were already implemented in all three apps. This reveals a limitation of the method used to evaluate apps. For future research, it is recommended to revise the definition of the elements, and to create a formula that takes into account the extent to which an element is implemented. Nevertheless, for this research the evaluation is more used to generally conclude if game design is applied in household energy apps in the Dutch market. As concluded in a study that evaluated energy apps in the U.S. (Beck et al., 2019), at this moment, both the energy apps in the U.S and Dutch market are not so gameful. The use of game design in household energy apps seems to be in an early stage, and there is still much to be gained if such elements prove to be effective. For this reason, the evaluation of household energy apps in this research could be outdated relatively soon. It is a snapshot of the current situation, and it is advised to do it again in the future with a revised definition of the elements, and formula to draw a more accurate conclusion on the use of game design.

The data on electricity consumption in the experiment shows a mixed effect on using an energy app, and adding game design elements. Overall, the total consumption decreased with 0.2%, and the average consumption of participants decreased by 0.4% (3.2% if the participant with the highest increase in consumption is left out of the calculations). This minimal difference makes it hard to draw conclusions purely based on this data. Research of Anderson, et al. (2020) on the statistical significance and statistical power in the design, implementation, and analysis of energy studies, deliberates on the fact when a study can draw quantitative conclusions. The experiment in this study is conducted with a sample size of 29 participants, and taking into account the method of Anderson et al. (2020), the minimum sample size would need to be around 2200 participants to draw any conclusions. With a sample size of 2200 participants, already a quite mild error margin is taken into account. For this research, a sample size of 2200 participants was not feasible due to limited available time to conduct the experiment. Finding so many participants would have been too time-consuming, and next to the leading researcher, a team of several extra researchers would be required for monitoring all participating households in the experiment. Despite this flaw in the experiment of this study on energy apps, it is still interesting to look at the data to inspire larger-scale experiments.

In the experiment, three different apps were used, Smappee, Smart Dodos, and U-meter. These three apps were selected since they had a relatively high game design score in the evaluation compared to the other apps. The most exciting result is that the different app users show a recurring pattern over the time of the experiment. The Smappee and Smart Dodos users show an average decrease in consumption in all three months, while U-meter users show an increase in consumption every month. Even when the worst performing participant (an overall increase of 73.6%) is left out of the calculations, the U-meter users increased their consumption on average with 3%, compared to a decrease of 7.9% for Smappee and 3.7% for Smart Dodos. The final survey provides some clarification. However, more in-depth interviews would have been helpful to see what participants experienced with the app. Of the nine participants that used the U-meter app, three used the app only once a month, three only twice over the whole period of the experiment, and three participants only looked at the app once at the beginning of the experiment. That these participants used the app so little was revealed at the end of the experiment, thanks to the final survey. Besides, most of the respondents that used U-meter did not reply to the question to provide feedback on the app. Therefore, in future experiments, it is advised to ask more often how many times participants used the app to enable the possibility to ask questions on their motivation to use, or not use, a particular app. The only U-meter users that gave feedback mentioned that the app needs to be kept simple, that it does not have many features, and that a feature of real-time monitoring was missing (Toon user). The two participants that used the Toon app before the experiment (during the experiment one used Smart Dodos, and the other used U-Meter), both gave the lowest ratings to the apps. This is rather interesting since Toon has the highest game design score in the evaluation of the energy apps performed in this research. With the help of in-depth interviews, more info would be available on their opinion.

Another goal within this study was to research if a social connection had any effect on the use of game design elements. Previously done research suggested that such a social connection could be helpful (Geelen et al., 2012; Konstantakopoulos et al., 2019). They observed that collaboration within families and households, or competition between students would increase the motivation of participants to conserve energy. During the experiment in this study, a shared goal and competition element were introduced. The responses in the survey hint that a strong social connection, as in the friends group, has a positive effect. Nevertheless, the data shows something else. The friends group increased their consumption the most. Consequently, these contradictory results are not reliable to draw definite conclusions. Even the participants that show an increase in electricity consumption mentioned that the competition and leaderboard elements motivated them. Overall, all participants in all groups, except one, mentioned that they believe game design elements can make people more conscious about their electricity consumption behavior. In contrast, ten of these participants increased their overall consumption. As described in the theoretical background, this can be a limiting or even harmful effect of game design (Hyrynsalmi et al., 2017). Due to time constraints, it was not possible for the researcher to interview all these participants, and the general information from the survey cannot fully explain if this indeed is a limiting or harmful effect.

The question to answer with the experiment was if participants would reduce their electricity consumption thanks to game design elements, and due to the small sample size and results, the answers to this question are inconclusive. Nevertheless, changing human behavior is always dependent on the willingness of users. Therefore, it cannot be predicted how much energy could be saved if this experiment was implemented on a larger scale. Thanks to automated smart systems, like occupancy monitoring with motion sensors, and other technologies, the electricity that is used for lighting could be saved without the need for changing the behavior of users. Research into these ICT systems in office buildings shows that occupancy can be predicted, and energy consumption can be adjusted to these predictions. Besides, real-time occupancy monitoring could help in lowering energy consumption (Salimi & Hammad, 2019). Integrating these automated building monitoring systems could potentially also be beneficial for the household sector. Some apps, like Toon (Eneco, n.d.), already implemented features that enable users to operate appliances, for example the thermostat or lighting, without the need to be home. If such features would be automated with occupancy sensors, it will potentially result in energy conservation without the need for human decision making.

Probably these automated systems would work better at office buildings due to the more predictable use of energy. Research with 26 interviews with different types of office employees shows that overall, people are willing to save energy and that they are aware of energy-wasting scenarios. They also believe that a gamified app could help if it involved actual energy consumption and tips about what they can do to lower their consumption (Lounis et al., 2017) . This again shows that personal feedback could be helpful by adjusting users behavior. A similar response can be seen in the survey after the experiment in this study. Participants mentioned that more personalized feedback could help in changing their behavior, and they were positive about the tips given in the WhatsApp group. In the data-driven reports that were created, a simple form of personal feedback was added. However, due to the limited features that the apps implemented, it still can be seen as general feedback that applies to all households. The only app that shows personal feedback, based on appliances and snooze electricity consumption, is Smappee, and this app is the clear winner of the experiment. Despite this fact, it is impossible to conclude that feedback has a positive effect based on this research. In future research, more attention could be paid to incorporating personalized feedback into an experiment.

Overall, this study gives some insights into the effect of using energy apps and adding game design elements. The sample size makes it difficult to draw any quantitative conclusion. However, it can function as an inspiration for future research. A possible future experiment should incorporate a larger sample size, preferably at least 2200 participants, and enough researchers to enable the monitoring of all these participants. Also, in-depth interviews with the participants could reveal more about the experience of users and the app they are using. It is valuable information to improve the apps and added game design elements. The game design elements created for this study could be copied but improved. Examples could be more personalized feedback in the data-driven reports and sending leaderboards more often during the competition. It is interesting to see if the extra attention for these elements translates into real reductions in electricity consumption.

For the suppliers of these energy apps, it is recommended to invest time in testing the already implemented features and designs with customers. This research reveals that while most apps incorporated the same elements, the experience of users differs completely. Keep an eye on research in the field of gamification, it can potentially cause reductions in electricity consumption, and even small reductions in consumption can have a significant impact when used by many. If such energy apps can be made accessible to all European households, these apps can help in mitigating the effect of climate change. At the same time these apps can provide users with a sense of control, and even save them electricity costs. It is expected by the researcher of this study, that if users experience significant reductions in energy costs, they would be willing to pay a small fee for an app.

6. Conclusion

This study had two goals, evaluating the energy apps on the Dutch market on the use of game design, and using three apps in an experiment to determine the effect of game design elements. These two goals were combined in the research question; *To what extent do energy apps in the Dutch market use game design elements, and what is the effect of these elements on the behavior and energy consumption of their users?*

In total, 30 apps were evaluated, and 50% of the apps use two or fewer game design elements, while 18 elements are available according to the framework of AlSkaif et al. (2018). Taking into account all 30 evaluated apps, an app uses 3.2 elements on average, and all except one app used the statistics element. To determine how gamified an app is, a formula was specially created for this research. The average game design score of all apps is 3.8 out of 28.8 points, and the best scoring app is Toon, created by the energy supplier Eneco, which scored 9.8 points. These scores show that apps in the Dutch market did not incorporate game design to a high extent, a similar conclusion as the research of Beck et al. (2019) that evaluated energy apps in the U.S.

The evaluated apps were divided into four categories: Free and accessible to all, free but only accessible to customers of a certain supplier, paid but accessible to everybody, and paid and only accessible with an extra device. The top three best scoring apps all belong to the category where a paid extra device is needed. This indicates that paid apps use more game design elements to attract customers. Nevertheless, the final survey of the experiment shows that the way the statistics element is implemented has the most effect on users. The app with the best result in the experiment, Smappee (scores 8.4 points in the evaluation), was easy to use, and its statistics were compelling to users. Smappee incorporated real-time monitoring, and the participants that already used the app with the highest game design before this experiment (Toon) mentioned that they were missing the real-time monitoring feature in Smart-dodos (7 points in the evaluation), and U-meter (7 points in the evaluation). These were the other two apps used in the experiment. The fact that participants with Smappee made most use of the app, and that according to the final survey participants only used the statistics element in the apps, concludes that while statistics is the most basic element of an app, it is also the most crucial element. Statistics are needed to provide users with insights on their electricity consumption. This awareness is the starting point to improve their behavior. Therefore, it is recommended for app creators to spend special attention in developing an attractive design to show the statistics, and to test this element with potential users. Next to the design, users need to have easy access to the app, and a mobile app is more favorable over a web app. That the participants did not make use of the other game design elements that were already incorporated in the apps, does not mean that adding extra elements is not beneficial to conserve electricity. The opposite seems to be true when looking at the experiment in this research.

Twenty nine participants, in three groups with a different social connection (neighbors, friends, and no connection), signed up for an experiment of three months in this research. In the first month, they only used the app without adding game design elements. In the second month, game design elements that trigger their intrinsic motivation were added, and in the third month, elements were added that trigger (mostly) the extrinsic motivation. The data on the electricity consumption shows that all participants together decreased their consumption on average with 0.4%, and leaving out one extreme participant with an increase of 73.6%, shows an average decrease in consumption of 3.2%. This is a positive result since five participants spend more time at home (baby at home, illness, and pension), and only one participant spend less time at home, compared to last year (all these participants belong to the friends and no connection group). Overall, the results per month show a recurring pattern, with the neighbors decreasing their consumption with an average of 9.7% over the whole experiment. The friends group increasing their consumption with an average of 6.8%, and the no connection group decreasing their consumption with an average of 5.7% (if the participant with an increase of 73.6% is excluded). In total, 16 of the 28 participants (with data on their electricity consumption of last year) experienced a decrease in consumption. Next to the fact that for quantitative conclusions the sample size is too small, this data is inconclusive on the effect of the different game design elements, and motivational perspectives that are triggered by these elements. Therefore the survey at the beginning and the end of the survey were very valuable.

According to the survey, the participants appreciated the added competition and leaderboard elements the most of all added elements. The movie theater gift certificates they could win during the competition motivated eleven of the participants, but not all of them. They seem to be motivated from a trait perspective, meaning from individual characteristics that are relatively stable over time and have a strong influence on behavior. The participants that responded positively to the competition mentioned that they are competitive in general. The participants also appreciated the weekly tips about using washing machines and dryers more efficiently, snooze electricity consumption, and replacing lightbulbs with led lights (See appendix B). These tips also trigger the trait perspective, and by using them, participants can get control over their consumption, it can give them a sense of achievement. Within this research, it is impossible to say if this affects the long term electricity consumption. It could help if these people are reinforced that they are 'doing good' by other elements. The data-driven report can potentially take care of this, and according to the participants, it is the third most compelling game design element. A combination of tips, and reinforcements through data-driven reports could trigger the behaviorist motivational perspective. This motivational perspective is based on good behavior in the past and rewarding this good behavior. Looking at the results in the survey, the elements that trigger this trait- and behaviorist motivational perspective are most effective. Elements that trigger an emotional or cognitive perspective, like the WhatsApp group (message box), the holiday e-card, and the common goal (collaboration), seem to be less effective. For the common goal element, there is an exception for the friends group. Their social connection seems to trigger half of the participants in this group for a common goal to conserve electricity.

Looking at the consumption data again, it was evidenced that the friends group did not collaborate at all to lower their consumption. Every month they consumed more electricity than they did last year, while most group members mentioned that they changed their behavior thanks to the added game design elements. Eight out of nine participants said they have an affinity with energy and sustainability, and even six of them are actively working on reducing their impact on the environment. This could be a signal of a limiting or harmful effect of game design (Hyrynsalmi et al., 2017), but can also be caused by external factors. The data shows an increase in consumption of 6.8%, while the participants had the feeling they were 'doing good.' However, some of the participants that consumed more electricity experienced a change in their household situation. Besides, the increase in consumption was the highest in the first month and the lowest during the last month, this still shows improvement. Another possible explanation that they did not show a decrease in consumption could be their affinity with energy. Possibly they already tried to lower their electricity consumption before this experiment. In that case, the experiment has less effect than with a group that partly misses this affinity with sustainability, like the neighbors group.

The neighbors group shows the most positive consumption data, and overall they decreased their consumption with 9.7%. Within this group, only five out of eleven group members mentioned that they think energy and sustainability is an important topic, and only three people actively used an app during the experiment. A possible reason why they consumed less electricity is that they have a more economic incentive than the other groups. More than half of these participants spend between 30%-50% of their household income on rent. This is much more than the group without a social connection, where eight out of nine participants spent less than 30% on rent or mortgage.

The group without a social connection shows a total increase in consumption of 4.2%. However, as mentioned before, this is most accountable to one household with an increase of 73.6%. Without this household, the group shows a decrease in consumption of 5.7%. The fact that most participants in this group and the neighbors group show a decrease, while most participants in the friends group show an increase in consumption, makes it hard to conclude that a stronger social connection has a positive effect. Nevertheless, the friends group shows the best result during the month that included the competition element. This observation in combination with the final survey during the experiment and previously done research (Bang et al., 2009; Gustafsson et al., 2009) suggest that a strong social connection definitely can have a positive impact on using game design in energy apps. The most critical recommendation of this research would be to conduct similar research with a larger sample size. Such an experiment could draw definite conclusions on the overall and social effect of game design elements in energy apps.

With that in mind, the last part of the research question cannot be answered with high certainty since a larger sample group is necessary. Nevertheless, thanks to this research, it is clear that energy apps on the Dutch market do not make use of game design elements to a great extent. This can be seen as a missed opportunity because this, and previously done experiments, show that game design elements could positively impact users behavior. It can contribute to the conservation of household energy consumption, the second-largest energy-consuming sector of the EU.

7. Bibliography

About the UN Climate Change Conference - December 2019. (n.d.). Retrieved December 9, 2019, from https://unfccc.int/process-and-meetings/conferences/un-climate-change-conference-december-2019/about-the-un-climate-change-conference-december-2019

Abrahamse, W., Steg, L., Vlek, C., & Rothengatter, T. (2005). A review of intervention studies aimed at household energy conservation. *Journal of Environmental Psychology*, *25*(3), 273–291. https://doi.org/10.1016/j.jenvp.2005.08.002

Albertazzi, D., Ferreira, M. G. G., & Forcellini, F. A. (2019). A Wide View on Gamification. *Technology, Knowledge and Learning*, *24*(2), 191–202. https://doi.org/10.1007/s10758-018-9374-z

Allcott, H., & Todd, R. (2012). *THE SHORT-RUN AND LONG-RUN EFFECTS OF BEHAVIORAL INTERVENTIONS: EXPERIMENTAL EVIDENCE FROM ENERGY CONSERVATION*.

AlSkaif, T., Lampropoulos, I., van den Broek, M., & van Sark, W. (2018). Gamification-based framework for engagement of residential customers in energy applications. *Energy Research and Social Science*, *44*(May), 187–195. https://doi.org/10.1016/j.erss.2018.04.043

Anderson, B., Rushby, T., Bahaj, A., & James, P. (2020). Ensuring statistics have power: Guidance for designing, reporting and acting on electricity demand reduction and behaviour change programs. *Energy Research & Social Science*, *59*(August), 101260. https://doi.org/10.1016/j.erss.2019.101260

AppAnnie Intelligence. (n.d.). Retrieved September 12, 2019, from https://www.appannie.com/top

Bang, M., Svahn, M., & Gustafsson, A. (2009). *Breaking New Ground: Innovation in Games, Play, Practice and Theory. Proceedings of DiGRA 2009 Persuasive design of a mobile energy conservation game with direct feedback and social cues*.

Banks, M. (2013). Individual responsibility for climate change. *Southern Journal of Philosophy*, *51*(1), 42–66. https://doi.org/10.1111/sjp.12008

Beck, A. L., Chitalia, S., & Rai, V. (2019). Not so gameful: A critical review of gamification in mobile energy applications. *Energy Research and Social Science*, *51*(January), 32–39. https://doi.org/10.1016/j.erss.2019.01.006

Costanzo, M., Archer, D., Aronson, E., & Pettigrew, T. (1986). The Difficult Path From Information to Action. *Energy Conservation Behavior*, *41*(June), 521–528. https://doi.org/10.1037/0003-066X.41.5.521

Deterding, S., & Dixon, D. (2011). *From Game Design Elements to Gamefulness: Defining “Gamification.”* 9–15.

Eneco. (n.d.). Toon Eneco. Retrieved March 19, 2020, from https://www.eneco.nl/energieproducten/toon-thermostaat/

Enelogic. (n.d.). Enelogic app. Retrieved March 19, 2020, from https://enelogic.com/nl/mogelijkheden-slimme-meter

Energy consumption in households. (2019). Retrieved July 18, 2019, from eurostat website: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Energy_consumption_in_households

European Comission. (2014a). *Cost-benefit analyses & state of play of smart metering deployment in the EU-27*.

European Comission. (2014b). Energy Performance of Buildings. Retrieved December 9, 2019, from https://ec.europa.eu/energy/en/topics/energy-efficiency/energy-performance-of-buildings/overview

European Comission. (2019). Energy efficiency directive. Retrieved December 9, 2019, from https://ec.europa.eu/energy/en/topics/energy-efficiency/targets-directive-and-rules/energy-efficiency-directive

European Commission. (2009). *Ecolabels DIrective*. 10–35.

European Parlament. (2010). *Directive on the energy performance of buildings*.

European Parlament. (2012). *Directive on energy efficiency 2012 EU*. (October), 1–56.

European Union. (2013). Manual for statistics on energy consumption in households. In *Manuals and guidelines - Environment and energy*. https://doi.org/10.2785/45686

Geelen, D., Keyson, D., Boess, S., & Brezet, H. (2012). Exploring the use of a game to stimulate energy saving in households. *J. of Design Research*, *10*(1/2), 102. https://doi.org/10.1504/jdr.2012.046096

Gifford, R., & Nilsson, A. (2014). Personal and social factors that influence pro-environmental concern and behaviour: A review. *International Journal of Psychology*, *49*(3), 141–157.

https://doi.org/10.1002/ijop.12034
Gustafsson, A., Katzeff, C., & Bang, M. (2009). Evaluation of a pervasive game for domestic energy engagement among teenagers. *Computers in Entertainment*, *7*(4), 1. https://doi.org/10.1145/1658866.1658873
Hyrynsalmi, S., Smed, J., & Kimppa, K. K. (2017). The dark side of gamification: How we should stop worrying and study also the negative impacts of bringing game design elements to everywhere. *CEUR Workshop Proceedings*, *1857*(May), 96–104.
I. Lampropoulos, W. L. Kling, P. F. Ribeiro, J. van den B. (2013). History of demand side management and classification of demand response control schemes. *IEEE*.
Kjeldskov, J., Skov, M. B., Paay, J., & Pathmanathan, R. (2012). *Using mobile phones to support sustainability*. 2347. https://doi.org/10.1145/2207676.2208395
Konstantakopoulos, I. C., Barkan, A. R., He, S., Veeravalli, T., Liu, H., & Spanos, C. (2019). A deep learning and gamification approach to improving human-building interaction and energy efficiency in smart infrastructure. *Applied Energy*, *237*(January), 810–821. https://doi.org/10.1016/j.apenergy.2018.12.065
Lampropoulos, I., Alskaif, T., van den Broek, M., van Sark, W., & van Oostendorp, H. (2019). *A Method for Developing a Game-Enhanced Tool Targeting Consumer Engagement in Demand Response Mechanisms*. 213–235. https://doi.org/10.1007/978-3-319-99444-4_9
Lounis, S., Kotsopoulos, D., Bardaki, C., Papaioannou, T. G., & Pramatari, K. (2017). Waste no more: Gamification for energy efficient behaviour at the workplace. *CEUR Workshop Proceedings*, *1857*(May), 129–134.
Morganti, L., Pallavicini, F., Cadel, E., Candelieri, A., Archetti, F., & Mantovani, F. (2017). Gaming for Earth: Serious games and gamification to engage consumers in pro-environmental behaviours for energy efficiency. *Energy Research and Social Science*, *29*(November 2016), 95–102. https://doi.org/10.1016/j.erss.2017.05.001
Morschheuser, B., Hamari, J., Werder, K., & Abe, J. (2017). How to Gamify? A Method For Designing Gamification. *Proceedings of the 50th Hawaii International Conference on System Sciences (2017)*, (Hicss), 1298–1307. https://doi.org/10.24251/hicss.2017.155
Rai, V., & Beck, A. L. (2017). Energy Research & Social Science Play and learn : Serious games in breaking informational barriers in residential solar energy adoption in the United States. *Chemical Physics Letters*, *27*, 70–77. https://doi.org/10.1016/j.erss.2017.03.001
Sailer, M., Hense, J., Mandl, H., & Klevers, M. (2013). Gamification as an Innovative Approach to Foster Motivation. *Interaction Design and Architecture(s) Journal - IxD&A*, (19), 28–37.
Salimi, S., & Hammad, A. (2019). Critical review and research roadmap of office building energy management based on occupancy monitoring. *Energy and Buildings*, *182*, 214–241. https://doi.org/10.1016/j.enbuild.2018.10.007
Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: A survey. *International Journal of Human Computer Studies*, *74*, 14–31. https://doi.org/10.1016/j.ijhcs.2014.09.006
Smappee. (n.d.). Smappee app. Retrieved March 19, 2020, from https://www.smappee.com/be_en/homepage
Smartphonebezit gegroeid naar 93% van Nederlanders, veelvuldig gebruik storend. (2018). Retrieved July 22, 2019, from consultancy.nl website: https://www.consultancy.nl/nieuws/15292/smartphonebezit-gegroeid-naar-93-van-nederlanders-veelvuldig-gebruik-storend
United Nations. (2015). Paris agreement. In *Paris Agreement* (Vol. 45).
van Raaij, W. F., & Verhallen, T. M. M. (1983). Patterns of residential energy behavior. *Journal of Economic Psychology*, *4*(1–2), 85–106. https://doi.org/10.1016/0167-4870(83)90047-8
Verkade, N., & Höffken, J. (2017). Energy Research & Social Science Is the Resource Man coming home ? Engaging with an energy monitoring platform to foster flexible energy consumption in the Netherlands. *Chemical Physics Letters*, *27*, 36–44. https://doi.org/10.1016/j.erss.2017.02.015
Wee, S. C., & Choong, W. W. (2019). Gamification: Predicting the effectiveness of variety game design elements to intrinsically motivate users' energy conservation behaviour. *Journal of Environmental Management*, *233*(November 2018), 97–106. https://doi.org/10.1016/j.jenvman.2018.11.127
Wemyss, D., Cellina, F., Lobsiger-Kägi, E., de Luca, V., & Castri, R. (2019). Does it last? Long-term impacts of an app-based behavior change intervention on household electricity savings in Switzerland. *Energy Research and Social Science*, *47*(September 2018), 16–27. https://doi.org/10.1016/j.erss.2018.08.018

Appendix A, evaluation Apps

	App Catergory	Information provision	Rewarding system	Social Connection	User Interface	Performance Status	Categories	Elements	Statistics	Total score
Bliq	1	Statistics, Tips	-	-	Notifications	-	2	3	0	3.6
MijnEnergieInzicht	1	Statistics, Tips	-	-	-	-	1	2	0	2.2
Oxxio	1	Statistics, Tips	-	Competition	Progress bar, Message box	-	3	5	0	6.5
Slimmemeter Portal Basic	1	Satistics, Tips	-	Competition	-	-	2	3	0	3.6
Umeter	1	Satistics, Messages	Offers	Competition	Dashboard	-	4	5	0	7
Energy Zero free	1	Statistics	-	-	-	-	1	1	0	1.1
Energy Tracker	1	Statistics	-	-	-	-	1	1	0	1.1
Mijnwoning.nl	1	Statistics	-	-	-	-	1	1	0	1.1
Greenchoice (boks slim)	2	Statistics, Tips (0.1)	-	Competition		Badges	3	4	0.1	5.6
Van de Bron	2	Statistics	-		Progress bar		2	2	0	2.4
Engie app	2	Statistics	-	Competiton	-	-	2	2	0	2.4
Regelneef (Energie direct)	2	Satistics	-	-	Message box	-	2	2	0	2.4
Oxxio pro	2	Statistics (0.1), Tips	-	Competition	Progress bar, Message box,	-	3	6	0.1	8.4
Eneco	2	Statistics	-	-	Progress bar	-	2	2	0	2.4
Nina van Nuon	2	Satistics,	-	-	Notifications	Badges	3	3	0	3.9
Essent	2	Satistics	-	-	Notifications	-	2	2	0	2.4
Energiemanageronline.nl	3	Statistics	-	Competition	Progress bar	-	3	3	0	3.9
Slimmemeteruitlezen.nl	3	Statistics (0.1)	-	Competition	-	-	2	2	0.1	2.6
Smart Dodos	3	Statistics (0.1), Messages (data	-	Competition	Dashboard, Progress bar	-	3	5	0.1	7
Energiemeter (DBO energie)	3	Statistics	-	-	-	-	1	1	0	1.1
Slimmemeter Portal Plus	3	Satistics, Tips, Messages	-	Competition	-	-	2	4	0	4.8
Toon	4	Statistics, Tips (0.1), Messages	-	Energy community	Dashboard, Degree of contr	-	3	7	0.1	9.8
Huisbaasje	4	Statistics, Tips	Offers	-	-	-	2	3	0	3.6
Aurum	4	Satistics	-	-	Progress bar	-	2	2	0	2.4
IUNGO	4	Statistics	-	Energy community	-	-	2	2	0	2.4
Smappee	4	Statistics (0.1), messages	-	Competition	Dashboard, Degree of contr	-	3	6	0.1	8.4
Youless	4	-	-	-	-	-	0	0	0	0
Enelogic	4	Statistics (0.1)	Offers	Competition	Dashboard, Leaderboard, Pr	-	4	6	0.1	9
Beeclear	4	Statistics	-	-	-	-	1	1	0	1.1
CEMM	4	Satistics	-	-	Notifications	-	2	2	0	2.4

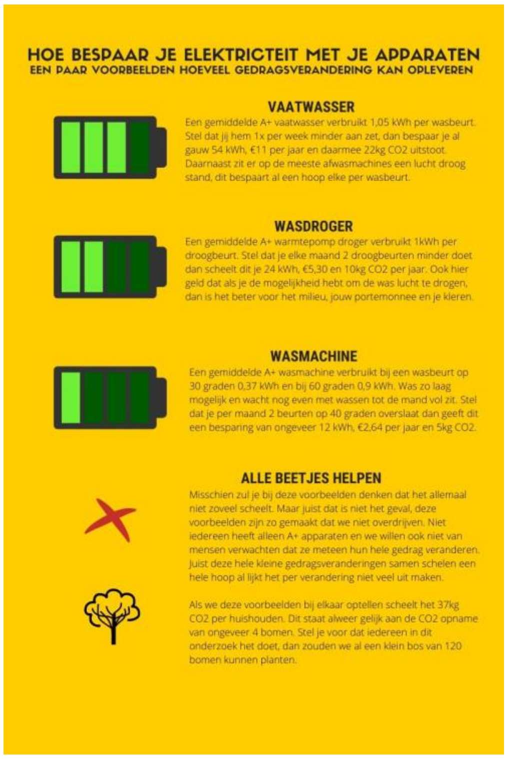
For the full evaluation in excel go to https://docs.google.com/spreadsheets/d/1AveIzfESivPd6gm-1P4b8LJpRolplx7uXJ9GD2VPGcA/edit?usp=sharing

Appendix B, added game design elements January

Happy Holidays e-card

Data driven report December

Tips



Appendix C, Added game design elements February

Appendix D, Data on energy consumption participants

Name	App experiment		Dec 2018	Jan 2019	Feb 2019	Total electricity	Extra persons	Time at home	Dec 2019		Jan 2020		Feb 2020		Total consumption	
Jeroen Buunk	Smappee		155	160	124	439	-		140	9.68%	130	18.75%	132	-6.45%	402	8.43%
Joelle Lufueni	Smappee		153	156	138	447	-		132	13.73%	145	7.05%	125	9.42%	402	10.07%
Tijmen Bruggema	Smappee		52	62	58	172	-		65	-25.00%	55	11.29%	51	12.07%	171	0.58%
Femke Greefhors	Smart Dodos		71	58	61	190	-		57	19.72%	43	25.86%	59	3.28%	159	16.32%
Carlijn Hudepohl	Smart Dodos		106	118	92	316	-		117	-10.38%	110	6.78%	105	-14.13%	332	-5.06%
Martijn Vlasblom	Smart Dodos		116	120	132	368	-		94	18.97%	87	27.50%	121	8.33%	302	17.93%
Manouk Mocking	Smart Dodos		105	117	105	327	-		88	16.19%	57	51.28%	108	-2.86%	253	22.63%
Stephanie Gumb	Umeter		35	53	42	130	-	1day extra at hor	40	-14.29%	44	16.98%	44	-4.76%	128	1.54%
Daan Veeneman	Umeter		178	168	133	479	-		123	30.90%	120	28.57%	113	15.04%	356	25.68%
Thijs Kosterman	Umeter		-	88	79	167	-		88		92	-4.55%	72	8.86%	164	1.80%
Marleen Klinkent	Umeter		105	88	-	193	-		93	11.43%	88	0.00%	76	-	181	6.22%
			1076	1188	964	3228			949	7.09%	971	17.23%	930	2.88%	2850	9.65%
									11.80%		18.27%		3.53%		11.71%	
Boy van der Laal	Smappee		161	178	168	507	-		141	12.42%	145	18.54%	123	26.79%	409	19.33%
Ilse Baijens	Smappee		159	191	142	492	1	Extra roommate	143	10.06%	123	35.60%	113	20.42%	379	22.97%
Michiel Ton	Smappee		146	123	152	421	-		150	-2.74%	139	-13.01%	132	13.16%	421	0.00%
Sanne van meete	Smart dodos		104	106	98	308	-		115	-10.58%	113	-6.60%	86	12.24%	314	-1.95%
Tijs Koedam	Smart dodos		209	213	195	617	1	Baby	241	-15.31%	251	-17.84%	228	-16.92%	720	-16.69%
Jessie de Booij	Smart Dodos		365	390	301	1056	-		457	-25.21%	497	-27.44%	414	-37.54%	1368	-29.55%
Femke Verheuve	U-meter		259	253	195	707	-		291	-12.36%	290	-14.62%	274	-40.51%	855	-20.93%
Sanne de Graaf	U-meter		-	255	230	485	-		224		268	-5.10%	223	3.04%	491	-1.24%
Rogier van Hees	U-meter		234	270	194	698	1	Baby	339	-25.56%	342	-26.67%	249	-28.35%	930	-33.24%
			1637	1979	1675	5291			1877	-8.66%	2168	-6.35%	1842	-5.30%	5887	-6.81%
									-14.66%		-9.55%		-9.97%		-11.26%	
Baue Weemaens	Smappee		99	115	100	314	-		110	-11.11%	82	28.70%	86	14.00%	278	11.46%
Thomas van Eer	Smappee		105	116	96	317	-		117	-11.43%	121	-4.31%	105	-9.38%	343	-8.20%
Andre Hofman	Smappee		301	248	203	752	-	Pension	238	20.93%	236	4.84%	227	-11.82%	701	6.78%
Paul koedam	Smart Dodos		448	446	335	1229	-	Pensioen	334	25.45%	405	9.19%	277	17.31%	1016	17.33%
Jasper Pot	Smart Dodos		165	137	137	439	-	Illnes	168	-1.82%	174	-27.01%	114	16.79%	456	-3.87%
Frank Brandsma	Smart dodos		246	298	396	940	-		250	-1.63%	280	6.04%	219	44.70%	749	20.32%
Michel van Son	U meter		425	389	302	1116	-		418	1.65%	402	-3.34%	337	-11.59%	1157	-3.67%
Stan Koedam	U-meter		-	99	113	212	-	Illnes + other roo	115		187	-88.89%	181	-60.18%	368	-73.58%
Dick Schoenmak	Umeter/Smart dodos		-	-	-	-	-		-		-					
			1789	1848	1682	5319			1635	3.15%	1887	-9.35%	1546	-0.02%	5068	-4.18%
									8.61%		-2.11%		8.09%		4.72%	
		Smappee	1331	1349	1181	3861		Smappee	1236	1.84%	1176	11.94%	1094	7.58%	3506	7.94%
		Smart dodos	1935	2003	1852	5790		Smart dodos	1921	1.54%	2017	4.78%	1731	3.12%	5669	3.74%
		U meter	1236	1663	1288	4187		U meter	1304	-1.37%	1833	-10.85%	1493	-14.81%	4630	-10.83%
		Total	4502	5015	4321	13838		Total	4461	0.85%	5026	2.06%	4318	-0.71%	13624	0.41%
											U-Meter - Stan	-1.09%	U-Meter - Stan	-8.32%	U-Meter - Stan	-2.98%
											No connection -	2.02%	No connection -	8.57%	No connection -	5.74%
											Total - stan	5.43%	Total - stan	1.58%	Total - stan	3.15%
											Smart Dodos - M	-0.35%	Nieghbors- stan	2.55%		
											Neighbors -Manc	13.82%				
											Total - Manouk	0.23%				
											Total - stan & ma	3.66%				

For the full evaluation in excel go to https://docs.google.com/spreadsheets/d/1AveIzfESivPd6gm-1P4b8LJpRolplx7uXJ9GD2VPGcA/edit?usp=sharing