RESIDENTIAL ENERGY REBOUND EFFECT ASSESSMENT BY USING SERIOUS GAMES

Oscar Garay García

Master of Science Thesis: Engineering and Policy Analysis

Faculty of Technology, Policy and Management – Delft University of Technology

TU Delft

Residential energy rebound effect assessment by using serious games

By

Oscar Garay García

in partial fulfilment of the requirements for the degree of

Master of Science
in Engineering and Policy Analysis

at the Delft University of Technology,
Faculty of Technology, Policy and Management,
to be defended publicly on Wednesday August 17th, 2016 at 03:00 PM.

Thesis committee:

Chairman: dr.ir. Bert Enserink (Policy Analysis Section)
1st supervisor: dr.ir. Els van Daalen (Policy Analysis Section)
2nd supervisor: dr.ir. Emile Chappin (Energy and Industry Section)
External supervisor: Bas van Nuland (The Barn Games)
External supervisor: Iman Mohammed

TU Delft

Acknowledgments

It is time to thank. And not because I have too, because I fell it. We are not sailing alone in this life, many sailors and captains are sailing with us. If we sink, we all sink, if we reach the shore, we all do it. Therefore, I can’t take credits of this thesis all by myself, that would be too selfish.

First, I would like to thank my thesis committee for all their support throughout these 6 months of hard work. First, thanks to Professor Els for so many meetings, e-mails and revisions. I don’t even remember how many times I was in her office learning from her expertise. I owe her great part of this work. Second, thanks to Professor Bert for his good sense of humor and for giving very accurate feedback on the job that I was carrying out. Third, thanks to Professor Emile for his easy going style of guiding students and for his very precise and high quality feedback. Fourth, thanks to Bas for standing me in my eagerness to start the experiments as soon as possible, and for giving me his time in getting the game ready. Last, but not least, thanks to Iman for her first push in my initiation with serious games, her help at the beginning was very helpful to get the first impression of what serious games were all about.

In addition, thanks to my family who was always on my side. Although they were not physically with me, they always showed me their support in my busiest and most stressful days. Esperanza, Patricio, Diana, Jimena, Kamal, Rodrigo, Alonso, Valentina, Cristobal, Rosario, Sergio, and many cousins, uncles, aunts and grandparents. Yes, we are a very big family and it is great! However, I would like to give special thanks to my mom, Esperanza, who always gave us her unconditional support, no matter what. She always knew where were we, what exams I had to take the following day and what I had for lunch. Despite being 12 thousand kilometers away from the Netherlands, we always felt her at home, spoiling us and loving us.

My pals in the Netherlands also deserve a great recognition for making me smile and laugh. Danilo, Dalton, Rabee and Dimos, we had great moments together, I hope I also made you laugh and enjoy. The Chilean community also deserves some credits. Luzma, Ale, Dori, Nico, Moises and Carmen, thanks for so many good moments. Being away from home was way easier with all of you around here.

My infinite thanks as well to all the people who participated in this research by giving me one hour of their lives to play the game and answer questionnaires. Without them this research would not have been possible.

I would like to thank the love of my life, my wife Macarena. She is the greatest person I have ever met, she left everything she had to be with me, no conditions no questions asked. We had wonderful moments in these 2 years, unforgettable moments that our grandchildren will love to hear. Without her I would not have made this journey. She is the captain of my life, my sailor, my everything. Thanks for always staying right there next to me, even in the hardest moments.

Finally, I want to thank the clock, because it has always been on my side…

Executive summary

Energy reduction has been on the political agenda since the last couple of decades. One of the most common policies to reduce energy consumption has been improving efficiency. However, the phenomenon called the rebound effect may threaten the effectiveness of policies aiming to reduce consumption by improving efficiency. The rebound effect is the process in which energy savings, after energy efficiency improvements, are lower than expected.

The scientific community agrees on the existence of the rebound effect and the possibility to measure it. In spite of this agreement, the rebound effect has been catalogued by many scholars as a highly controversial concept. In fact, the two main controversies around the rebound effect are: its size and its importance for the policy making process. The main causes that have produced the rebound effect to be controversial are: methodological issues of previous attempts to assess the effect, different numeric definitions and fuzzy and different system boundaries.

The main objective of this thesis research was to find a new and innovative methodology to assess the rebound effect in order to improve the methodological issues of previous attempts that have analyzed the rebound effect. As a result, in reducing the causes of controversies, the controversies themselves may be reduced as well. The methodology that seems to improve the mentioned methodological issues is serious games. Serious games have several advantages that seem to fit and solve the shortcomings of the previous attempts that have analyzed the rebound effect in the past. In particular, the *NRG game* is the specific serious game that was used to perform the actual assessment in this research.

The NRG game was used to carry out a new rebound effect assessment. In doing so, 50 people played the game in two different groups in a way to perform a modified before/after analysis: one group of 25 people using a low efficiency house and 25 people using a high efficiency house. The conclusions of this assessment showed that the rebound effect was, indeed, detected to be present when the behavior of the two groups was compared. In fact, two main signs of the rebound effect were detected. First, having a low efficiency made the low efficiency group to reduce their energy consumption more than the high efficiency group. Second, having a high efficiency made the high efficiency group to increase their comfort level (a direct measure of the luxury level of their houses in the game) more than the low efficiency group. As a result, the rebound effect was detected by keeping track of the total energy consumption and comfort level of each player.

In addition, some of the methodological issues of previous rebound effect assessments that have used before/after analysis were improved, for instance, the possibility to perform ex-ante assessments, the inclusion of psychological factors of people in the results without making inaccurate assumptions and the inclusion of more than just one energy services in the experiments, among others. As a result, serious games were proven to be a handy tool to assess and analyze the rebound effect, improving the quality of previous assessments.

Despite having proved the usefulness of serious games in assessing the rebound effect, some limitations of this research were identified: the sample under study was not a good representation of the population, the reliability of the rebound effect size calculation is compromised, the findings of this assessment can’t be applied in a real life context and so forth.

Nevertheless, these set of limitations do not hamper the achievement of this research's objectives. Future research topics can be focused on finding a set of criteria on how to better analyze if the results from the gaming environment can be applied in a real life context, perform new rebound effect assessments in order to find the real influence in the results of differences in the income players receive in the game and find common definitions on how to calculate base case scenarios for rebound effect assessments using before/after analysis.

Table of Contents

1. Introduction

Since the last couple of decades, policy makers and politicians all over the world have started to design policies in order to control the ever-growing energy consumption phenomenon. Not only the depletion and increasing prices of the fossil fuel sources, but also the increasing environmental awareness have made policy makers take this issue seriously (Frondel et al., 2012). In order to tackle this problem, one of the most common policies that has been implemented all over the world is increasing energy efficiency in, for instance, power plants, transmission and distribution grids and heating and electric household appliances. If we focus on those policies aiming to reduce household energy consumption by improving the energy efficiency of household appliances, we encounter a problem: the rebound effect. The rebound effect is the phenomenon in which energy savings, due to improvements in energy efficiency, are lower than expected. When consumers internalize that improving energy efficiency leads to a reduction of the energy price they perceive (i.e. their monthly energy bill), consumers might unconsciously increase their energy consumption, reducing the expected energy savings that policy makers may have forecasted (Berkhout et al., 2000, Greening et al., 2000). If it is assumed that improving energy efficiency results in a lower energy price, and the fact that energy consumption is price-elastic, then the logical conclusion is that improvements in efficiency lead to a higher consumption.

The rebound effect has been widely studied and analyzed by several scholars (see Binswanger (2001), Greening et al. (2000), Sorrell et al. (2009), Dimitropoulos (2007), among others). Although the whole scientific community agrees that the rebound effect is true and can be measured, several controversies around the rebound effect concept have been present all along. First, the size of the rebound effect is a controversy since different scholars have calculated and measured different magnitudes of the effect (Binswanger, 2001, Greening et al., 2000, Sorrell et al., 2009). The latter has caused disputes between those scholars about what size is the correct one. Second, the importance of the rebound effect on the policy making process is a controversy too (Sorrell and Dimitropoulos, 2008). One the one hand, some scholars argue that the rebound effect will never offset all the energy savings brought by improvements in energy efficiency, reason why it is not worth to take further actions to reduce the effect. On the other hand, some other scholars claim that even though some energy savings are lost due to the rebound effect, policy actions should be always taken to reduce the effect as much as possible in order to increase the achieved energy savings.

There are three main reasons of why those controversies around the rebound effect are present. First, previous attempts to measure the rebound effect have failed in implementing a reliable methodology, making methodological mistakes while performing the actual measurement (Aydin et al., 2015, Dimitropoulos, 2007, Sorrell and Dimitropoulos, 2008). Second, different numeric definitions have been used by different scholars to measure the rebound effect (Frondel et al., 2012, Madlener and Alcott, 2009, Sorrell, 2007, Sorrell and Dimitropoulos, 2008). Third, previous attempts to analyze the rebound effect have defined fuzzy and different system boundaries (Berkhout et al., 2000, Greening et al., 2000).

Considering the previously mentioned controversies around the rebound effect and the causes of those, any research that helps in reducing those causes will be a contribution in reducing the mentioned controversies. The present thesis research aims to find a new and innovative methodology that helps in reducing the causes of the rebound effect controversies.

The remaining parts of this introduction are as follows. First, the problem context is described. Second, the research objectives are defined. Third, one main research question and three sub research questions will be proposed. Fourth, the methodology to follow throughout this thesis research is also described. Fifth, the remaining chapters of this thesis research are briefly introduced.

1.1. Problem context

As mentioned before, the rebound effect is the phenomenon in which energy savings, due to improvements in energy efficiency, are lower than expected. This undesirable phenomenon jeopardizes the effectiveness of the policies aiming to reduce energy consumption. As a result, the rebound effect has become popular in the scientific community, resulting in many different studies that aim to measure and analyze its possible consequences. However, those different studies have obtained different results, introducing controversies around the rebound effect concept. As previously mentioned, the first controversy is related to the actual size of the effect and the second one is related to how important the rebound effect is for the policy making process. The previous controversies have caused that policy makers don't have a clear answer when they have to deal with the rebound effect. The previous problem is exactly where this thesis research is located. As a result, the problem context can be formulated as follows:

There is no consensus inside the scientific community on what the size of the rebound effect is and how important it is for the policy making process. This has led policy makers to struggle when they forecast and evaluate the possible consequences of the rebound effect.

1.2. Research objectives

In order to contribute to solve the previous research problem, it is important that, first, causes of the mentioned controversies around the rebound are identified. If new studies find a way to reduce the causes of the identified controversies, the problem previously described might be reduced. In that context, we can identify the main research objective of this research as follows:

Propose a new and innovative methodology that helps in reducing the causes of controversies around the rebound effect.

1.3. Research questions

In order to successfully achieve the previous research objective, a set of research questions have been proposed. The following is the main research question:

How can the rebound effect be adequately assessed in order to reduce the causes of controversies?

In order to operationalize the main research question, three sub research questions have been proposed as well:

1. What are the main controversies associated to the rebound effect concept?
2. What new and innovative methodology can be proposed to assess the rebound effect in order to reduce the causes of controversies?
3. What are the results and lessons learned when the chosen methodology is used to perform a new rebound effect assessment?

1.4. Methodology

To be able to answer the research questions, different methods will be applied. In the following those methods will be explained in further detail.

1.4.1. Literature review about the rebound effect

In order to answer sub research question 1, a literature review will be carried out. The literature to analyze will be related to the rebound effect and previous attempts to measure it in the past. The aim is to clearly identify what makes the rebound effect a controversial concept and why this has happened. In doing so, not only theoretical studies but also empirical measurements are going to be analyzed. Chapter 2 of this report will analyze the rebound effect in detail.

1.4.2. Literature review about a new and innovative methodology

Having identified the controversies of the rebound effect and their causes, a new literature review will be carried out in order to answer sub research question 2. This time the literature to analyze will be about this new methodology that may help to reduce the causes of controversies of the rebound effect: serious games. The main advantages of serious games seem to solve and reduce the causes that have produced those controversies. The aim of this second literature review is to justify and explain why serious games constitute a useful methodology to assess the rebound effect. In addition, some drawbacks of serious games will be also analyzed. Chapter 3 of this report will analyze the serious games methodology in detail.

1.4.3. Rebound effect assessment

Having justified why serious games seem to help to reduce the causes of the rebound effect's controversies, a new assessment of the effect will be carried out by using this methodology in order to test its usefulness. In doing so, sub research question 3 will be answered. The aim of this step is to actually prove that serious games are, indeed, a contribution to reduce the rebound effect's controversies. Chapters 4, 5, 6 and 7 will describe in detail how this new assessment of the rebound effect will be carried out and what the main results are.

1.5. Thesis structure

The remaining chapters of this thesis research are organized as follows. Chapter 2 analyzes the literature review about the rebound effect. Chapter 3 analyzes the literature about serious games in order to justify why this methodology helps in reducing the causes of the rebound effect's controversies. Chapter 4 describes the specific serious game that will be used to assess the rebound effect: the NRG game. Chapter 5 analyzes the results of the experiments carried out by using the NRG game. Chapter 6 and 7 deal with the rebound effect assessment based on the results from the experiments carried out by using the chosen game. Chapter 8 shows the main conclusions of this thesis research, showing the main limitations, recommendations and ideas for future researches.

2. The rebound effect

Before the industrial revolution in the 18th century, energy consumption was basically restricted to heating or cooking purposes. People used to burn wood, coal or organic material to heat their houses or cook. Any other use of energy was in its infancy. However, after the industrial revolution this energy consumption pattern changed forever, energy started to be used for industrial purposes as a way to boost the economy. First, burning coal was used by trains as a way to transport goods and by manufacturing machineries. After some years, electricity appeared onto the energy stage too in order to provide people with a reliable and safe lighting system for their cities and houses. When people realized the huge potential of electricity, they started to produce it in larger quantities in order to support industrial applications as well, such as transport, communications, military and infrastructure purposes (Ashton, 1966). New ways to massively produce energy had to be found to supply the increasing energy demand. Nowadays, coal, oil, gas and renewable sources provide the energy people need to support their life styles. However, countries have become so dependent on energy that some problems have arisen (OECD/IEA, 2014). First, energy sources have become very expensive due to the high demand and relatively low supply (Shafiee and Topal, 2010). Second, depletion of fossil fuels, such as oil, gas and coal, are expected to occur sooner or later reason why new energy sources must be found to compensate it (Shafiee and Topal, 2009). Third, extreme reliance on fossil fuels has produced serious environmental problems, such as global warming (Cox et al., 2000). As a result, nowadays humanity is facing a very important dilemma on how to ensure that people keep improving their life styles having cheap and reliable energy sources without compromising the environment.

The previous dilemma has, so far, been tackled by basically two different policies: green energy sources and energy consumption reduction (Frondel et al., 2012). By implementing both strategies people may keep having relatively cheap energy sources to ensure their life quality without risking the environment. On the one hand, green energy sources (such as, bioenergy, wind energy, solar energy and so forth) have become very popular in many countries, reducing their reliance on fossil fuels. On the other hand, energy consumption reduction is a more complex strategy considering that each year the total population in the planet rises by 86 million people (WorldBank, 2015). Instead, slowing the energy consumption down seems to be a more feasible target. Either way, if the problem is not tackled by both strategies, the dilemma might not be solved, jeopardizing not only our own life's quality, but also the future generations' life quality.

If we further analyze energy consumption reduction policies, we can see that policy makers have tackled this ever-growing energy demand phenomenon by implementing energy efficiency improvement policies. Improving energy efficiency has actively been on the political agendas of several countries and it is nowadays one of the pillars of greenhouse gases reductions (Grepperud and Rasmussen, 2004, Herring and Roy, 2007, Frondel et al., 2012). In the Netherlands, for instance, the goal is to improve energy efficiency by 2% per year over the period 2011-2020, and complement it with a reduction of greenhouse gas emissions of 30% by 2020 (Hieminga, 2013). However, empirical and theoretical evidence have shown that improving energy efficiency has not reached the expected energy savings that policymakers may have forecasted due to behavioral changes of energy users (Greening et al., 2000, Berkhout et al., 2000). This phenomenon of not achieving the expected energy savings when energy efficiency has improved is called the *rebound effect*. The first author who reported this phenomenon was William Stanley Jevons in 1865 (Sorrell, 2009, Wang et al., 2012, Wei, 2010, Herring and Roy, 2007), reason why the rebound effect is also called the *Jevons Paradox*. From that year on, the number of authors who have studied the rebound effect are countless.

The remaining parts of this chapter are organized as follows. First, the concept of the rebound effect is defined and different types are described. Second, a numeric definition is given. Third, previous attempts to measure the rebound effect are analyzed. Fourth, controversies around the rebound effect concept are introduced. Fifth, the causes of why the previous controversies have shown up are discussed. Sixth, implications of the previous controversies for policymakers are discussed and analyzed. Finally, a short conclusion of this chapter is given.

2.1. Conceptual definition and rebound effect types

After energy efficiency improvements it is natural to think that energy consumption is reduced by the same percentage in which the efficiency was improved. If energy efficiency improves by 10%, then energy consumption should be reduced by 10%. However, due to the rebound effect energy consumption is not reduced by the same percentage energy efficiency was improved, and some of the potential energy savings are spent back by the consumers. Although we can find several different definitions of the rebound effect in the literature (see chapter 2.5.2), the definition that Herring and Roy (2007) provide is explanatory and clear:

"The rebound effect is the extent of the energy saving produced by an efficiency investment that is taken back by consumers in the form of higher consumption". (Herring and Roy, 2007, p. 2)

In the previous definition, the higher consumption that Herring and Roy (2007) refer can take the form of more hours of use of the energy service or a higher quality of it. When technological progress occurs, the equipment is made more efficient resulting in a reduction of the input needed (energy in this case) to produce the same amount of output. Due to the necessity of less energy to produce the same output, the cost per unit of output falls. Berkhout et al. (2000) exemplifies this phenomenon by comparing it with the amount of kilometers a car can drive with the same amount of gasoline. They explain that if a car can drive more kilometers with the same amount of gasoline, then the fuel costs per kilometer fall, and so does the total cost per kilometer. After a reduction of the cost per unit of output, what follows is an increase in consumption. The driver tends to increase his/her energy consumption, i.e. more kilometers will be driven. The amount in which the energy consumption increases depends on the energy's own price elasticity. On the one hand, if the mentioned elasticity is low then the consumption should not increase so much. On the other hand, if the mentioned elasticity is high then the energy consumption should increase accordingly. Authors like Berkhout et al. (2000) and Sorrell and Dimitropoulos (2008) have called this effect the *first order rebound effect* or just the *direct rebound effect*. This type of rebound effect predicts that energy consumption of the same energy service in which the efficiency was improved increases.

But what if we now consider possible extra consequences of the direct rebound effect? The changes in energy consumption associated to, for instance, car usage might impact on other products that use energy as well. Imagine that, as a consequence of the reduction of the driving costs per kilometer, the car driver decides to use the air conditioning at a lower temperature while driving as a way to unconsciously compensate the initial cost reduction. In doing so, the driver will be using extra energy to keep the car cooler during the summer. This constitutes another type of rebound effect, in which the energy consumption of energy services different than the one in which efficiency was improved increases. The amount in which the energy consumption of this second energy service changes is defined by the cross price elasticity between the two services. Again, authors like Berkhout et al. (2000) and Sorrell and Dimitropoulos (2008) have called this effect the *second order rebound effect* or just the *indirect rebound effect*.

In addition to the direct and indirect rebound effect, several scholars have also defined the *third order rebound effect* or *economy wide effects*, which is the absolute sum of the direct and all the possible indirect effects (Greening et al., 2000). When the energy efficiency of one energy service increases, not only the service itself (direct rebound effect) or some other services (indirect rebound effect) change consumption, but also the whole economy might be affected. When the car driver decides to decrease the temperature of the air conditioning or drive more kilometers than he/she used to do, then the entire market might change. Some production curves of car or air conditioning companies (just to mention some) might change in reaction of this new consumption pattern. In addition, it may happen as well that oil companies change the way they produce gasoline trying to adjust their production to the new reality. This domino effect touches different markets of the whole economy, producing new equilibriums after demand and supply have adjusted their output to maximize their utility functions.

Ouyang et al. (2010) provided a diagram to illustrate how the rebound effect works in households' energy consumption (Figure 2.1). First, technical improvements produce a more comfortable indoor environment and reduce the energy use. Second, this energy use reduction increases the disposable income which, altogether with a more comfortable indoor environment, produces an increment of the household lifestyle. Finally, a better lifestyle with more disposable income produces extra energy consumption (either in the same energy service producing a direct rebound effect, or in other energy services producing an indirect rebound effect). The latter partially offsets the initial energy reduction. This process of offsetting initial energy savings is called the rebound effect.

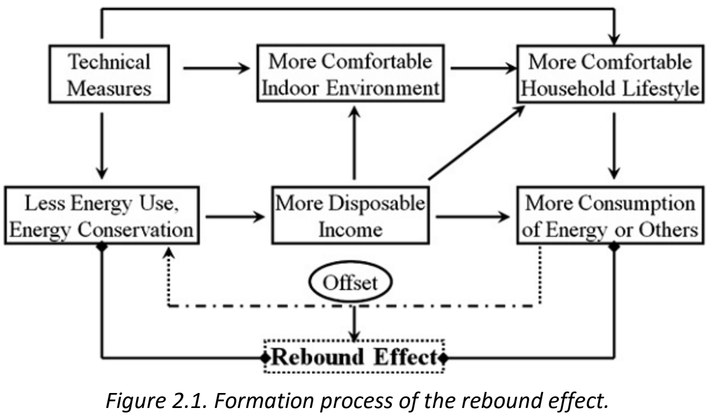


Figure 2.1. Formation process of the rebound effect.
Source: Ouyang et al. (2010)

2.2. Numeric definition

Having given a conceptual definition and different types of the rebound effect, a numeric definition can also be given. The following formula reflects in an easy way the portion of energy savings that are spent back by the consumers due to behavioral changes. It doesn't distinguish if the direct, indirect or economy wide effects are calculated. Due to its simplicity it could be used in any situation as long as expected and actual energy savings are clearly identified (Jin, 2007).

$$Rebound\ effect = \frac{expected\ savings - actual\ savings}{expected\ savings} \times 100 \quad (1)$$

We have developed a simple example in which some calculations are made using the previous formula. Imagine if the monthly consumption in lighting a house is 50 KWh and all the lightbulbs are changed for efficient ones which consume 75% less energy. We would expect that the new monthly consumption in lighting was 12.5 kWh (25%×50kWh). In other words, to get the same amount of lighting less energy is needed. Thus, the total cost of lighting per unit of output drops. As a result, the users may start to use the lights more hours per month than before or buy extra ones to install in other places, as a reaction of the lower costs. For instance users can consume in the new situation a total of 20 kWh per month, not the expected 12.5 kWh. In this case, the rebound effect can be seen as a reduction of the expected energy savings, from 37.5 kWh (expected savings) to 30 kWh (actual savings). If we apply the previous formula we will get that the rebound effect is equal to 20%, which means that 20% of the expected energy savings are spent back by the consumers by using the lights more hours than before or by installing extra lamps in other places.

As we argue in chapter 2.5.2, several numeric definitions of the rebound effect can be identified in the literature. Those different numeric definitions are based on different micro and macro-economic assumptions about how elasticities are analyzed and by what factors they are influenced. Those assumptions are specific to each particular study, reason why it turns out very difficult to apply those different definitions to different studies with different data and different methodologies, constituting one of the sources of controversies around the rebound effect (see chapter 2.4). However, since formula (1) just translates the conceptual definition of the rebound effect into numbers and it does not make any assumption about elasticities or micro and macro-economic concepts, it can be applied to any case as long as expected and actual savings are identified. Due to the latter reason, formula (1) will be used hereinafter when calculations about the rebound effect are made (chapter 6).

2.3. Previous attempts to analyze and measure the rebound effect

The number of previous studies that have measured or studied the rebound effect are many. Just by searching in google scholar for the term "rebound effect", a total of 249 thousand results are given (results given the 15th of March of 2016). If the same term is searched just in the title of the article, this number is reduced to 739 articles. If we break down the latter number, we can see that almost 50% of those articles were reported after the year 2010 (Figure 2.2).

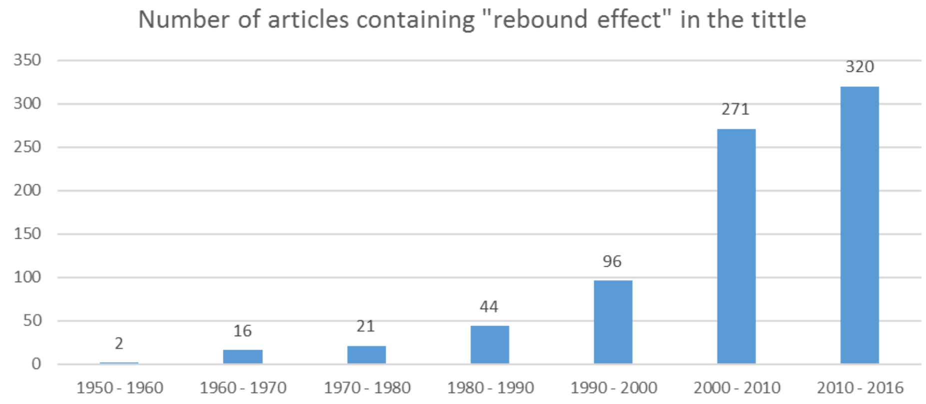


Figure 2.2. Number of studies about the rebound effect by year (results given the 15th of March of 2016).

The previous behavior is not a coincidence. The rebound effect has started to be widely analyzed in the past 20 years as a way to understand why policies aiming to improve energy efficiency have not completely succeeded in reducing energy consumption. Figure 2.2 also gives us a feeling of how important the rebound effect concept has become in the last 20 years.

When we analyze the studies that have tried to measure the rebound effect we can find out that authors have mainly focused in measuring it in specific energy services. For example, measuring the rebound effect in transportation (Frondel et al., 2012, Greene, 1997, Roy, 2000, Small and Van Dender, 2005), households' electricity consumption (Jin, 2007, Khazzoom, 1986, Grepperud and Rasmussen, 2004, Ouyang et al., 2010), heating services (Aydin et al., 2015, Dubin et al., 1986, Haas and Biermayr, 2000), and so forth. However, as previously mentioned, the existence of the second order rebound effect makes difficult to separate different energy services due to the fact that they are all somehow intertwined. Due to the latter reason is why we can find authors like Sorrell et al. (2009), Greening et al. (2000) and Binswanger (2001) who tried to summarize and give some kind of coherence to the huge set of previous studies that have measured the direct rebound effect in different energy services and countries. We have grouped together the findings of those three authors in Table 2.1. It shows the rebound effect size range in different energy services measured by several studies. It also shows the main methodologies used and the countries analyzed.

Table 2.1. Comparison table of previous studies aiming to measure the direct rebound effect.

Direct rebound effect in different energy services	**Sorrell et al. (2009)**	**Greening et al. (2000)**	**Binswanger (2001)**
Automotive transport	10% to 30%	10% to 30%	<10% to 50%
Space heating	5% to 40%	10% to 30%	5% to 65%
Space cooling	1% to 26%	0% to 50%	-
Water heating	<20%	<10% to 40%	-
Residential lighting	-	5% to 12%	-
Number of studies analyzed	43	68	14

Direct rebound effect in different energy services	Sorrell et al. (2009)	Greening et al. (2000)	Binswanger (2001)
Methodologies used[1]	- ES (31) - QE (12)	- ES (N.A) - QE (N.A)	- ES (11) - QE (3)
Countries analyzed	OECD countries	Mainly USA	Mainly USA

In addition, some other relevant studies were found in the available literature that have also measured the direct rebound effect in the same energy services as Table 2.1, but were not included in the studies Sorrell et al. (2009), Greening et al. (2000), Binswanger (2001) did. Those new studies have carried out new measurements of the rebound effect. Table 2.2 shows a summary of those.

Table 2.2. Studies focussed in the direct rebound effect in different countries.

Direct rebound effect in different energy services	Source	Rebound effect size	Country	Methodology[1]
Automotive transport	Wang et al. (2012) Small and Van Dender (2005) Frondel et al. (2012)	96% 3% to 12% 57% to 62%	China USA Germany	ES ES ES
Residential energy use	Ouyang et al. (2010)	30% to 50%	China	B
Space cooling	Jin (2007)	57% to 60%	South Korea	QE
Space heating	Haas and Biermayr (2000) Aydin et al. (2015) Aydin et al. (2015) Herring and Roy (2007)	20% to 30% 27% to 41% 55% Not a final number	Austria Netherlands Netherlands UK	ES-QE ES QE S

If we take a closer look at Table 2.1 and Table 2.2, we can see that great differences exist on the rebound effect size. For instance, space cooling energy services show rebound effects from 0% to 60%. In addition, automotive transport shows a rebound effect size from 10% to almost 100%. The large rebound effect size differences that can be found in the literature will be analyzed in chapter 2.4.1 when controversies about the rebound effect are discussed.

In contrast to what happens with the direct rebound effect, the available literature that has measured the economy wide rebound effect is not no numerous (Sorrell, 2007). However, Dimitropoulos (2007) has summarized the main studies focused in the macro level rebound effect (9 studies in total). In addition, Lin and Liu (2012) and Jin (2007) have also studied the macro level rebound effect by doing new measurements in Asia. Table 2.3 shows the findings of those studies. This time, the *methodology* column is not shown due to the fact that all the studies have used econometric studies to measure the effect. This result is not a

[1] ES: Econometric studies / QE: Quasi-experimental analysis / B: Benchmarking with other countries / S: Direct surveys.

surprise since to measure the economy wide rebound effect, a variety of macroeconomic factors must be measured (like employment, inflation, GDP, money supply, interest rates and so forth), leaving no other option to measure it than using econometric studies.

Table 2.3. Studies focussed in economy wide rebound effect.

Source	**Rebound effect size**	**Country**
Dimitropoulos (2007)	>100%	China
	15%	Holland
	35% to 70%	Japan
	170% to 350%	Kenya
	<100%	Norway
	120%	Scotland
	54% to 59%	Sudan
	60%	Sweden
	30% to 50%	UK
Jin (2007)	30% to 38%	South Korea
Lin and Liu (2012)	53,2%	China

The economy wide rebound effect size also shows great differences throughout countries and different authors, just like the direct rebound effect did. Moreover, in some cases the size interval can be very wide, especially in the case of Kenya in which the measured rebound effect goes up to 350% (a phenomenon that is often called the "backfire rebound effect" in which the net energy consumption increases compared to the situation without efficiency improvements (Jin, 2007, Ouyang et al., 2010)).

If we now take a look at the main methodologies used by previous attempts to measure to rebound effect, we can see that the most common one is to use econometric theories applied to analyze historical data. The second most common methodology is to use quasi-experimental analysis, or most commonly referred as before/after analysis. Finally, direct surveys and benchmarking techniques have also been used, but just in a very small number compared to econometric studies and quasi-experimental analysis. In chapter 2.5.1 those methodologies are further discussed.

When we see that the attempts to measure the rebound effect are so numerous, one could think that this has helped to agree on the rebound concept and consequences. However, the reality of the rebound effect is far away from being agreed. In fact, several scholars have labeled the rebound effect as a controversial phenomenon. The following chapter shows the main controversies associated to the rebound effect.

2.4. Controversies

Although the whole scientific community agrees that the rebound effect exists and is present when the energy efficiency increases (Frondel et al., 2012, Sorrell and Dimitropoulos, 2008), it has some controversies associated to its concept. In fact, authors like Binswanger (2001), Dimitropoulos (2007), Greening et al. (2000), Grepperud and Rasmussen (2004), Hertwich (2008), Lin and Liu (2012) and Sorrell (2007) have referred to the rebound effect as a highly controversial concept. In addition, it is not unusual to find articles, like the ones written by Gillingham et al. (2013) and Friedrichsmeier and Matthies (2015), in which the rebound effect is discussed just by the debates it generates and its controversial characteristics. In fact, Friedrichsmeier and Matthies (2015) say that *"the debate is getting out of hand in a way that is no longer helpful to our understanding of energy consumption behavior"* (p.80).

In this thesis research, we analyze two different types of controversies that are associated to the rebound effect concept: its size and its importance. In fact, Sorrell (2007) states that *"rebound effects are very difficult to quantify, and their **size** and **importance** under different circumstances is hotly disputed"* (p. 5). In addition, Saunders (2000) also says that *"rebound has a sound theoretical basis; however, its **magnitude** and **importance** are an empirical question"* (p.1). Those two types of controversies have been separated in this research to give a more coherent and sounder flow of ideas, although they are intertwined and it is difficult to totally isolate them.

2.4.1. Size

The size of the rebound effect is the first controversy that is present. As shown in Table 2.1 and Table 2.2, the measured rebound effect can go from 0% (there are no energy savings lost due to behavioral changes) up to almost 100% (all the predicted energy savings are spent back by the consumers). The size of the rebound effect depends on the country in which the measurement is carried out, the energy service under study (space heating, automotive transport, residential lighting and so forth), the circumstances involved in the measurement and the time period analyzed (Sorrell et al., 2009). In addition, Aydin et al. (2015) also state that the rebound effect may vary across different socio-economic segments of the society, making the rebound effect size vary between one individual to another. As a matter of fact, Haas and Biermayr (2000) presented a study in which 11 dwellings in Austria showed different rebound effects after thermal efficiency improvements. The lowest rebound effect identified was 0% and the highest one was 61%. The latter shows that even statistically similar people might show a completely different rebound effect, somehow confirming to what Aydin et al. (2015) said.

Despite all the latter evidence that shows the rebound effect can't be analyzed as if it was a unique number with similar characteristics across different situations and/or realities, we can still find authors like Santarius (2014) and Gillingham et al. (2013) who try to find a unique rebound effect size. Santarius (2014) argues that the size of all rebound effects is very likely to exceed 50%, while Gillingham et al. (2013) say that behavioral responses after energy efficiency improvements are in between 5% to 30%. To those scholars who try to find a unique rebound effect size, Friedrichsmeier and Matthies (2015) respond that consensus on the absolute magnitude of the rebound effect distracts the attention of the scientific community on important matters, like evaluating and planning strategies that minimize the rebound effects in the many contexts and realities where they occur.

To sum up, it becomes clear that the size of the rebound effect is an important controversy inside the scientific community. Some scholars argue that there might be a single rebound effect while some others neglect that idea claiming that every case should be analyzed independently. Either way, the controversy associated to the rebound effect size still remains.

2.4.2. Importance for the policy making process

The second controversy that is present in the rebound effect concept is its importance for the policy making process. Some scholars argue that the rebound effect can seriously hamper the policy making process, while others claim that the rebound effect does not jeopardize the effectiveness of energy policies aiming to improve efficiency (Sorrell and Dimitropoulos, 2008). This controversy becomes important for policy makers when they face the decision of whether to ex-ante analyze the rebound effect consequences before improvements in efficiency or not (Sorrell, 2007, Roy, 2000). Two different reasons of why scholars defend one opinion or the other are identified: 1) because of the size of the rebound effect and 2) independent of the size.

First, some scholars are influenced by size of the rebound effect to take one of the two sides: the rebound effect as an important concept for the policy making process or the rebound effect as a not important concept for the policy making process. On the one hand, scholars who have calculated or estimated a small rebound effect tend to catalog it as a not important concept for the policy making process. On the other hand, scholars who have calculated or estimated a large rebound effect tend to catalog it as an important concept the policy making process.

Second, some other scholars take one of the two sides independently of how big or small the rebound effect is. On the one hand, scholars who think the rebound effect is not important for the policy making process claim that no matter how big the rebound effect is, energy savings would be achieved anyway, reason why no further analysis of the effect should be carried out by policy makers. On the other hand, scholars who think the rebound effect is important for the policy making process claim that no matter how small the rebound effect is, policy makers should always try to reduce it in order to increase energy savings as much as possible.

As a consequence, we can conclude that scholars defend one of the two different positions (rebound effect is important for policy making and rebound effect is not important for policy making) by using one of the two different reasons (considering the size of the rebound effect and not considering the size of it). In order map the authors who take one of the two sides by taking one of the two arguments, we have constructed Table 2.4 and Table 2.5. Table 2.4 maps authors who hold the idea that the rebound effect is indeed an important concept for the policy making process, while Table 2.5 maps authors who hold the idea that the rebound effect is not an important concept for the policy making process. Both tables separate the authors according to the two reasons we have analyzed.

Table 2.4. Author's opinions to claim the rebound effect is an important concept in the policy making process.

Opinion: Important concept in the policy making process	
Reason: Dependent on the size (big)	**Reason: Independent of the size**
Frondel et al. (2012): *"From a policy perspective, the fact that the estimated rebound is* ***relatively high*** *irrespective of driving intensity calls into question the effectiveness of efficiency standards as a pollution control instrument"* (p. 6). Jin (2007): *"In conclusion, these suggest that rebound effect is* ***an important factor*** *that the government of South Korea must consider when planning its energy efficiency improvement policy."* (p. 1). Lin and Liu (2012): *"The results show that, over 1981 - 2009, energy rebound effect amounts averagely to 53.2%, implying that* ***China cannot simply rely on technical means to reduce energy consumption and emission****."* (p. 1). Ouyang et al. (2010): *"The existence of a* ***high rebound****, at least 30% and up to 50%, has been presumed in the household energy efficiency of China"* (p. 6). Wang et al. (2012): *"A* ***majority of the expected reduction*** *in transport energy consumption from efficiency improvement* ***could be offset due to the existence of rebound effect.****"* (p.1).	Dimitropoulos (2007): *"…importance of the macro-economic rebound effect should* ***not be underestimated****."* (p. 1). Binswanger (2001): *"…the rebound effect with respect to households is* ***indeed a relevant phenomenon*** *that is too often neglected in the discussion of sustainable development."* (p. 12). Sorrell (2009): *"A prerequisite … is a recognition that* ***rebound effects matter and need to be taken seriously.****"* (p. 13).

Table 2.5. Author's opinions to claim the rebound effect is not an important concept in the policy making process.

Opinion: Not important concept in the policy making process	
Reason: Dependent on the size (small)	**Reason: Independent of the size**
Berkhout et al. (2000): *"…the share of energy use in total energy use is* ***too small for the policy maker to bother*** *that energy conservation policy will rebound."* (p. 7). **Greening et al. (2000):** *"This leads us to the conclusion that* ***the rebound is not high enough*** *to mitigate the importance of energy efficiency as a way of reducing carbon emissions."* (p. 11). **Small and Van Dender (2005):** *"For now, our conclusion is that policy analyses of regulations affecting fuel efficiency should assume that* ***the rebound effect is considerably smaller today*** *than has been measured in the past, and is likely to become smaller still as time goes on."* (p. 25).	**Gillingham et al. (2013):** *"Increasing energy efficiency brings emissions savings… People may drive fuel efficient cars more and they may buy other goods,* ***but on balance more-efficient cars will save energy****."* (p. 476).

If we keep analyzing Table 2.4 and Table 2.5 we can realize that the controversy over the importance of the rebound effect relies on whether policy makers should spend time and resources forecasting and minimizing the rebound effect or not. On the one hand, scholars who hold the opinion that the rebound effect is an important concept in the policy making process (Table 2.4) defend the idea that those resources and time must be spent to increase the effectiveness of policies aiming to reduce energy consumption as much as possible. On the other hand, scholars who hold the opinion that the rebound effect is not an important concept in the policy making process (Table 2.5) claim that policy makers should not spend those resources and time considering that energy savings will be achieved anyhow. Either way, we can identify controversial opinions over the importance of the rebound effect for the policy making process.

It is important to highlight once again that the controversy over the importance of the rebound effect refers uniquely to how the policy making process is affected by the presence of the rebound effect, and not whether it is real or not. The existence of the rebound effect is agreed to be true by the scientific community and it doesn’t constitute a controversy (Frondel et al., 2012, Sorrell and Dimitropoulos, 2008).

Having identified the two main controversies associated to the rebound effect concept (size and importance), the main causes of those controversies are analyzed in the following chapter. If causes of controversies are identified, then strategies to reduce them might be designed.

2.5. Causes of controversies

In the previous chapter two main controversies associated to the rebound effect were analyzed: its size and its importance. After reviewing different literature about what causes those controversies, we have identified three main reasons: methodological issues, different numeric definitions and fuzzy system boundaries. There are no specific studies that explicitly deal with all of those causes of controversies. Therefore, different studies had to be reviewed to discover them. In the case of methodological issues, the analyzed literature was Aydin et al. (2015), Dimitropoulos (2007), Sorrell and Dimitropoulos (2008), Sorrell et al. (2009), Herring and Roy (2007) and Ouyang et al. (2010). In the case of different numeric definitions, the analyzed literature was Frondel et al. (2012), Madlener and Alcott (2009), Sorrell and Dimitropoulos (2008) and Sorrell (2007). Finally, in the case of fuzzy system boundaries, the analyzed literature was Greening et al. (2000), Berkhout et al. (2000), Turner (2009) and Sorrell (2007). In the following, each of those three causes is analyzed in detail.

2.5.1. Methodological issues

One of the reasons that has caused controversies around the rebound effect to show up is the presence of methodologies issues in the previous assessments and measurements (Aydin et al., 2015, Dimitropoulos, 2007, Sorrell and Dimitropoulos, 2008). As previously mentioned in this chapter, a thorough analysis of the literature that has analyzed and measured the rebound effect in the past shows four main methodologies used: econometric studies of historical data, quasi-experimental analysis, direct surveys and benchmarking techniques. This chapter gives the most important shortcomings of the studies that have used the previous methodologies to measure the rebound effect.

2.5.1.1. *Econometric studies of historical data*

Dougherty (2008) defined econometrics as *"the application of statistical methods to the quantification and critical assessment of hypothetical economic relationships using data"* (p. 1). In other words, econometrics use statistics to analyze data in order to determine economic phenomena or relationships. The great majority of studies that have used econometrics to measure the rebound effect have focused on measuring elasticities (energy efficiency elasticity and price elasticities), meaning the percentage of change in one variable when the other variable has changed, keeping all the rest of the involved variables constant (Sorrell et al., 2009, Lin and Liu, 2012).

Sorrell et al. (2009) have analyzed with critical lenses the econometric studies and they have come to the conclusion that, even though the results are more methodologically correct than other methodologies, some shortcomings have made those studies overestimate the magnitude of the effect. First, the lack of reliable and enough data has caused measurements to show inaccuracy. Most of the times, data about consumption is easily accessible, but data about energy efficiency is not available. The latter reason has made scholars to estimate historic energy efficiency by using third variables and relations that may not be fully correlated with it, increasing the error and reducing the significance of the final results. Second, isolating different effects in order to calculate a pure rebound effect due to consumers' behavioral changes is very difficult. When scholars have analyzed historical data from the past 30 years, for instance, countless factors may have changed over time. Most of the times it is very difficult to identify those changes and take them all into account. For example, when analyzing the energy rebound effect in vehicle transportation in the past 30 years, the number of passengers per car, highways quality, car size, car age, average distance driven and even weather conditions, may have had an influence in the calculations. If those effects are not isolated the rebound effect would be overestimated and, as a result, lower statistical significance would be achieved. Third, neoclassical assumptions on the way the elasticities are calculated increase the bias of the result. The most crucial neoclassical assumption that affects the quality of the results is the rationality principle. The rationality principle predicts that the consumers will react in the same way to a decrease of prices and to improvements in efficiency, since the latter should have an identical effect on the cost of energy services. However, consumers do not act in a rational way and they may react differently when the price of energy services is reduced and when energy efficiency is improved (Berkhout et al., 2000, Friedrichsmeier and Matthies, 2015). Fourth, the econometric models may have a great complexity, creating some difficulties in their interpretation. Most of the times econometric models use a very large set of assumptions and data, making them difficult to understand and trace errors. Last, but not least, since the data used by econometric models comes from the past, just ex-post rebound effect calculations are obtained. Unfortunately, this doesn't give policy makers a tool or methodology to analyze future developments of the effect or to ex-ante evaluate the phenomenon.

2.5.1.2. *Quasi-experimental analysis*

The second most common methodology to assess and measure the rebound effect is by using quasi-experimental analysis, or most commonly referred as before/after analysis. In this kind of analysis two groups are analyzed in terms of a stimulus that one group is exposed to and the other one is not. By analyzing how those two groups behave over time, the researcher draws conclusions on how the stimulus influenced the exposed group (Meyer, 1995, Verschuren and Doorewaard, 2010).

Applied to the rebound effect measurement, a typical quasi-experimental analysis considers the comparison of one group in which energy efficiency has been improved against a second group which the energy efficiency was held constant. By analyzing the energy consumption of those two groups the rebound effect may be calculated. Despite the methodological simplicity it offers, quasi-experimental analysis has not been used as much as econometrics measurements (Sorrell et al., 2009).

Like in the case of econometrics studies, Sorrell et al. (2009) are very critical with the measurements of the rebound effect that have used quasi-experimental analysis. As a matter of fact, they say that *“the methodological quality of most of these studies is relatively poor”* (p.3). First, they say that most of these studies do not have a clear control group (i.e. the group that has not been exposed to the stimulus). In some studies, it does not become clear for which group the energy efficiency was improved because of different stimuli applied at the same time. Second, to be able to statistically compare the two groups in terms of how energy efficiency influenced the exposed group, the one and only variable that must change is the energy efficiency, otherwise the final calculation of rebound effect would be biased and influenced by some other hidden differences in both groups. In other words, the ceteris paribus principle should be strictly followed. In order to do so, several demographic characteristics of both groups should be measured to check how similar they are. Most probably the previous shortcoming has been caused because of the selection bias previous studies have shown, in which the participants of the experiments decide to participate rather than being randomly chosen from the population, producing significant demographic differences between the two groups. Third, and continuing on the same reasoning as the previous point, previous studies have shown that it is very challenging to totally isolate the rebound effect from other effects, somehow breaking the ceteris paribus principle. For example, when energy consumption of two groups are analyzed over time, different stimuli may have occurred to just one group and not the other one. For instance, house composition, energy prices, external or internal temperatures, income and house ownership may have changed in just one group, introducing bias to the analysis. Most of the times it is very difficult to isolate and identify all the stimuli that may have changed the initial conditions of the experiment. Fourth, and like in the case of the econometric studies, studies that have used quasi-experimental analysis mainly constitute ex-post measurements of the rebound effect. The latter doesn’t give policy makers the possibility to evaluate beforehand future developments of the rebound effect. Fifth, people’s psychological aspects have not been considered on quasi-experimental rebound effect measurements due to the overlooking of the consumer’ behavior theory. So far, all the previous attempts to measure the rebound effect have just tried to focus on economic parameters and have neglected the non-economic factors, like consumer’s irrational behavior. This is a critical shortcoming considering that consumer’ behavior is crucial with respect to households’ energy consumption. However, even if consumer' behavior theory had been considered, not all the possible different behaviors of different people would have been considered due to its large number (Haas et al., 1998, Jin, 2007, Roy, 2000, Madlener and Alcott, 2009). Last, but not least, almost all the rebound effect measurements that have used quasi-experimental analysis have just been focused on the household heating energy service. This has restricted the analysis to just the direct rebound effect, neglecting the indirect rebound effect in which reallocation of energy from one service to another occurs (from instance, savings in gas consumption may increase the electricity consumption).

Despite all the shortcomings previously described, using quasi-experimental analysis to measure the rebound effect provides a huge advantage: simplicity (Campbell and Stanley, 2015). When we see that one of the shortcomings of econometric studies is its inherent complexity, one could immediately think that a good candidate to propose a new assessment of the rebound effect would be by using quasi-experimental analysis. If the researcher is able to control the sources of errors and shortcomings previously described, then the results of a before/after analysis would be valuable and it would contribute to improve the methodological issues that have caused the rebound effect to be so controversial.

2.5.1.3. *Direct surveys*

Just a very small set of measurements of the rebound effect have used direct surveys. This result is not surprising due to its intrinsic limitations. Herring and Roy (2007) describe two studies carried out by the UK Open University in which internet and telephonic surveys were used to measure the rebound effect. The results do not show a rebound effect size, just an indication of how many people might change their behavior towards energy consumption after efficiency improvements. The latter result is limited and it doesn't give policy makers any new information about the rebound effect. The limited usefulness of this kind of study to measure the rebound effect is rooted on its two most important methodological limitations: nonresponse bias and lying bias (Verschuren and Doorewaard, 2010, Heijnen, 2008). First, depending on the length of the survey people might stop responding it in an accurate way at some point in the middle of it. This source of bias is called the nonresponse bias. Second, it is very difficult to check if people answer the questions honestly and not lying. The most obvious case of lies in a survey is the so called "socially accepted" answers. Some respondents may lie answering controversial questions by giving socially accepted answers. For example, people might try to answer they have a green energy attitude while they do not, just because it is more socially accepted.

2.5.1.4. *Benchmarking techniques*

Benchmark, as defined by Elmuti and Kathawala (1997), is *"the process of identifying the highest standards of excellence for products, services, or processes, and then making the improvements necessary to reach those standards."* (p.1). Although benchmarking is commonly applied to business and management techniques, in the rebound effect context we refer to benchmarking techniques as the process in which the rebound effect is calculated by making parallelisms and comparisons to other situations similar to the one under analysis. The benchmark concept, in its broad sense, it is not totally applied to the rebound effect measurement.

Just the study made by Ouyang et al. (2010) was identified of having used benchmarking techniques to measure the rebound effect. They indicate that most probably the rebound effect in households' energy consumption in China should be between 30% and 50%. They came to this conclusion by looking at previous measurements of the rebound effect carried out in other developing countries, like India and South Africa. They defend this comparison by saying that India and South Africa, just like China, have very low incomes and living standards, making them comparable in terms of the rebound effect.

The methodology used by Ouyang et al. (2010) shows important shortcomings. First, using benchmarking techniques to measure the rebound effect neglects what Sorrell (2009) and Aydin et al. (2015) said about the uniqueness characteristic of the rebound effect. They stated that the rebound effect depends on several socio-demographics factors, like income, house ownership, age, country and so forth, making each situation unique. Therefore, they claim that each measurement should be analyzed in its own context, and not by taking for granted previous studies results. Furthermore, they also rejected the fact that benchmarking should be used just as a guide and not for statistical precision (Elmuti and Kathawala, 1997). Second, the study made by Ouyang et al. (2010) does not show any indication of where the data of the previous studies in India and South Africa came from. It may have happened that the previous studies obtained the data from totally different sources making them incomparable to one another. Third, no indication is given of what assumptions were used by the previous measurements that were analyzed. It may also have happened that those studies took different assumptions for the way the calculations are carried out, making them, once again, incomparable to one another. Finally, when benchmarking techniques are used to analyze the rebound effect, the research always depends on other studies to collect data. Therefore, it is not possible to generate original data to analyze the rebound effect in each specific case.

2.5.1.5. *Summary*

Several methodological shortcomings of the previous attempts to measure the rebound effect have been identified. Some of the analyzed shortcomings are inherent to the methodology itself, in which some are not possible to solve and some others are possible to solve. In addition, some other shortcomings refer uniquely to specific studies that failed to address them in a correct way, and they are not an inherent shortcoming of the methodology itself. As a result, we have constructed Table 2.6 in order to sum up all the previously analyzed shortcoming by separating them into three categories: (1) shortcomings that are inherent to the methodology and are not possible to solve, (2) shortcomings that are inherent to the methodology and are possible to solve and (3) shortcomings that refer uniquely to the way each specific study was carried out.

Table 2.6. Methodological shortcomings of pevious attempts to measure the rebound effect.

Methodology	Shortcomings that are inherent to the methodology and are not possible to solve	Shortcomings that are inherent to the methodology and are possible to solve	Shortcomings that refer uniquely to the way each specific study was carried out
Econometric studies of historical data	- Neoclassical assumptions produce bias - Complexity - Ex-post rebound effect calculation	- Lack of reliable and enough data - Difficult ot isolate effects different from the rebound effect	
Quasi-experimental analysis (before/after analysis)		- Difficult to strictly follow ceteris paribus - Difficult ot isolate effects different from the rebound effect - Ex-post rebound effect calculation	- Absence of a clear control group - Just single energy service analysis - Neglection of people's psychological aspects
Direct surveys		- Nonresponse bias - Lying bias	
Benchmarking techniques	- Always depends on other studies to collect data (no original data can be collected)		- Neglection of the uniqueness characteristics of the reboud effect - No indication of where the data came from - No indication of what assumptions previous studies used

2.5.2. Different numeric definitions

As briefly discussed in chapter 2.2, several different numeric definitions of the rebound effect can be found in the literature. The latter reason constitutes the second cause of why the rebound effect is such a controversial concept inside the scientific community (Frondel et al., 2012, Madlener and Alcott, 2009, Sorrell and Dimitropoulos, 2008, Sorrell, 2007). As a matter of fact, Frondel et al. (2012) provide four different numeric definitions that scholars have used to measure the rebound effect which may lead to different measurements and conclusions about the effect. Those four definitions take different microeconomic and macroeconomic assumptions about the way the effect is measured. The first definition they provide is rooted in the relative change of energy services demand due to increments in efficiency. The second definition takes for granted that improvements in energy efficiency can be totally mirrored in a reduction of the energy service price. The third definition considers that energy efficiency is held constant throughout the analysis. The fourth definition Frondel et al. (2012) give considers that elasticity of energy demand under changes in energy price is negative. The result of having many different definitions is the fact that different calculations are carried out in which different conclusions about the size and importance of the rebound effect are obtained. In fact, Greening et al. (2000) after analyzing 68 previous attempts to assess and measure the rebound effect (see Table 2.1) came to the conclusion that *"the* *magnitude of the rebound effect varies substantially because of the definition of the activity measure and the methods used"* (p.393).

As mentioned in chapter 2.2, hereinafter when we talk about the numeric definition of the rebound effect we would be referring to formula (1). This formula does not make any assumption on micro or macro-economic concepts like the formulas Frondel et al. (2012) described. As a consequence, it is free of controversies around what assumptions are valid or what assumptions are not.

2.5.3. Fuzzy and different system boundaries

The third cause of controversies around the rebound effect concept inside the scientific community is the fact that previous attempts to measure the rebound effect have defined different and fuzzy system boundaries. First, if different scholars use different system boundaries, the final size of the rebound effect measurement will vary, just like in the previous chapter in which different definitions are used (Sorrell, 2009, Greening et al., 2000). If household energy consumption is analyzed, then the possible system boundaries to be chosen are numerous. One could include inside the borders of the system the influence of weather conditions into the gas consumption, or include the influence of daylight hours across the year into the lighting energy demand. The possible system boundaries are numerous and they must be analyzed carefully in order to include the effects that are most significant into the rebound effect measurement (Turner, 2009). Second, the used system boundaries have shown that sometimes they are fuzzy and blurred (Sorrell, 2007, Berkhout et al., 2000). In some of the previous attempts to measure the rebound effect it is not very clear what it is included into the system boundaries and what it is not. This has caused not only the interpretation of the results to be difficult, but also the applicability of it to future measurements complicated.

2.6. Implications for policymakers

Throughout the present chapter we have been discussing what the rebound effect is and its controversial characteristics, but we also have to analyze how this affects policymakers and the policy-making process. Analyzing Table 2.4 and Table 2.5 we discovered that the opinions of different scholars about how important the rebound effect is for the policy making process are contested. On the one hand, some scholars argue that the rebound effect is indeed an important phenomenon to consider in the policy making process and policy makers should try to reduce it as much as possible. On the other hand, some other scholars claim that the rebound effect is not an important phenomenon for the policy making process and policy makers should not further analyze it because energy savings will be achieved anyhow. Either way, policy makers are in the middle of this dispute, producing confusion and dubiety.

If we take a closer look at the ideas of those scholars who catalog the rebound effect as a not important concept for the policy making process, we could infer that they might be neglecting the *"backfire rebound effect"* (Ouyang et al., 2010, Jin, 2007). The backfire rebound effect is the process in which the size of the rebound effect is greater than one, meaning that energy consumption increases instead of decreasing after energy efficiency improvements. Scholars like Gillingham et al. (2013) say that *"increasing energy efficiency brings emissions savings. Claims that it backfires are a distraction."* (p.1). They strongly reject the possibility of the backfire rebound effect; instead they argue that improvements in energy efficiency will always bring energy savings. However, Sorrell (2007) says that there is strong evidence that backfire rebound effects have been identified in the past. The controversies continue and once again policy makers are in the middle of the dispute.

Policies aiming to increase energy efficiency in order to reduce energy consumption (and consequently greenhouse gases) are cost intensive and long term oriented (Brookes, 2000). Thus, failures in not meeting the policies' expectations might produce huge money losses and extra policy challenges for policymakers. As a result, it is worth to invest extra resources to investigate further the possible consequences of the rebound effect, even more if the backfire effect threatens the whole policy. Taking for granted that energy savings will be achieved anyhow is too simplistic and at least ex-ante evaluations should be carried out in order to decide if the predicted rebound effect is big enough to jeopardize the effectiveness of energy efficiency improvement policies (Sorrell, 2007, Roy, 2000).

We believe that the losses of not considering the rebound effect into the policy making process are much higher than the possible gains it brings. Failing in effectively reducing energy consumption by not considering the rebound effect jeopardizes not only energy consumption reduction policies, but also greenhouse gases reduction policies. The latter rejects all sustainability principles that most of the countries have recently agreed to follow on the COP21 conference held in Paris in 2015 (UnitedNations, 2015). We hold the same opinion as Sorrell (2007), who says that *"rebound effects should be taken into account when developing and targeting energy efficiency policy"* (p.92).

In conclusion, if we assume that it is worth to invest extra resources to investigate the possible consequences of the rebound effect, then new studies that tackle the sources of controversies around the effect will not only make a contribution to reduce the problem, but also to make policies more effective. As a result, studies that improve methodological issues, clarify the different numeric definitions of the effect and help to improve system boundaries, will be a contribution for the scientific community and policy makers. In this research we aim to propose a new and innovative methodology that helps to tackle some of the previous attempts' shortcomings that have caused controversies around the rebound effect.

2.7. Chapter's conclusions

Analyzing the existing literature about the rebound effect, two main controversies were identified: its size and its importance. In addition, three main causes of those controversies were identified as well: methodological issues, different numeric definitions and fuzzy system boundaries. We hold the idea that the rebound effect can seriously jeopardize the effectiveness of policies aiming to improve energy efficiency, and taking for granted that improvements in efficiency will always bring energy savings is just too simplistic. Therefore, it is worth to analyze and assess how the rebound effect affects those policies. As consequence, we believe that any new research that aims to reduce what has caused the controversies around the rebound effect will be a contribution to not only reduce the problem but also to allow policymakers to better tackle the phenomenon. In conclusion, this thesis research will try to propose a new methodology to assess the rebound effect in a way that the causes of controversies previously described are reduced as much as possible. In doing so, if the causes of controversies are reduced, the controversies themselves may be reduced as well.

3. Serious games

As previously mentioned in Chapter 1, the new methodology that will be tested in this research to assess the rebound effect is serious games. We believe that serious games are a feasible methodology to implement in order to tackle the causes of the rebound effect's controversies. This chapter is meant to justify why serious games seem to help in solving the problem. The remaining parts of this chapter are as follows. First, a definition of serious games is given. Second, two different types of serious games are discussed. Third, strengths and weaknesses of those two types of serious games are analyzed. Fourth, reasons of why serious games can help to reduce rebound effect's controversies are discussed. Fifth, disadvantages of using serious games to measure the rebound effect are given. In addition, ways to reduce those disadvantages are analyzed as well. Sixth, a final conclusion is given of why using serious games to measure the rebound effect may be a contribution to reduce controversies around this concept.

3.1. Definition

As Duke and Geurts (2004) said in their book called *Policy games for strategic management*, *"games are as old as humankind"* (p.31). The authors mention that most probably military games for entertainment and training were some of the first games ever known. In fact, chess dates back from the 6th century and represents military confrontations between two kingdoms, in which not only entertainment was involved, but also strategic behavior and problem solving skills. In addition, and going even further to the past, the Chinese military general Sun Tzu included gaming techniques and gaming strategies for training in his masterpiece book called *The art of war*, which dates back from 500 BC. Furthermore, according to Bracken and Shubik (2001), in World War I and World War II games were developed to train soldiers and generals in order to increase their strategic skills towards the enemy.

With the new digital era that started in the 1970s, a new form of games was born: video games. From that moment on, video games have become very popular among kids and adults. Video games have changed the way people spend leisure time, up to the point that, nowadays, the video game industry is even more profitable than the film industry (Boyle et al., 2011). In fact, according to Harteveld (2011), people spent 9 billion hours playing solitary in 2003, almost 1300 times more than the time it took to build the Empire State Building in New York City (7 million hours). This boost of video games has also influenced the kind of games discussed in the previous paragraph, games for military purposes and training. New technologies have been used ever since to develop more effective and more realistic games for such purposes. In addition, the flexibility that video games provide, has made many other areas to start using video games for training purposes as well. Serious games can be located in that context. In fact, Michael and Chen (2005) have defined serious games as follows:

"A serious game is a game in which education (in its various forms) is the primary goal, rather than entertainment". (p.17)

Education is a key concept in serious games. Regular games do not have education as a primary goal, while serious games do. Michael and Chen (2005) make clear in their definition that the education concept that they are referring to must be seen in its various forms. If we only restrict education to its most obvious form in which the student learns from the teacher, then we would be missing a great part of serious games. In fact, Zyda (2005) sees education in the serious games definition as any activity that *"involves pedagogy: activities that educate or instruct, thereby imparting knowledge or skills"* (p.26).

Despite the previous definition given, there is still great differences between scholars on what a serious game really is (Boyle et al., 2011). The latter is basically caused by the fact that gaming involves many disciplines, like computer science, psychology, media studies, philosophy, or even disciplines in which the game is embedded, like chemistry, biology, energy, mechanics, health, management and so forth (Harteveld, 2011). Some scholars have restricted the serious games definition to an educational process in which the players are the only ones who learn from the game, while some other scholars have broadened it including to possible learners any other actor that may be involved in the game environment (Meijer, 2009, Duke and Geurts, 2004). Despite the different definitions of serious games, what it is agreed between scholars is the fact that serious games do not have entertainment as a primary goal. In this thesis research we will not restrict the definition of serious games; instead, we will accept that serious games are a tool for education in which any actor involved in the game environment could learn from it. In the following section we will go deeper into this matter, presenting different types of serious games in which different people can learn from them.

3.2. Types of serious games

Like in the case of the serious games definition in which many different ones can be found in the literature, many different types and classifications of serious games can be distinguished as well. In fact, the variety of different classifications of serious games is even greater than the variety of definitions (Breuer and Bente, 2010). Mayer et al. (2016), for example, divided serious games in four types in terms of their final purpose: serious games as tools, for innovations purposes, as a persuasive methods and for self-organization purposes. Michael and Chen (2005), on the other hand, classified serious games in terms of application areas: military, government, educational, corporate, healthcare, political, religious and art games. In this thesis research we have decided to follow the classification Meijer (2009) provided due to the fact that it is directly applicable to the purpose of this thesis. He classified serious games in just two different branches: serious games for non-research purposes and serious games for research purposes. The following sections explain each of those two different types of serious games

3.2.1. Serious games for non-research purposes

The first type of serious games Meijer (2009) analyzes, is serious games for non-research purposes. This type of serious games involves learning, skill acquisition and training. Some scholars have also called this type of serious games as game-based learning (Boyle et al., 2011). This type of game is characterized by the fact that players are the individuals who learn from playing it. The game has as its main goal to teach players certain knowledge, skills or train them in specific matters. Depending on how the game is designed, the learning process can take different paths or use different learning theories.

When we talk about learning, skill acquisition or training in serious games for non-research purposes, we are focusing the different learning processes on players. In fact, players are the ones who learn when this type of serious games is used. Players are exposed to a game that has as a main objective to not only deliver certain knowledge or message to the player, but also to make them acquire and internalize it. The great majority of serious games that have been designed and implemented so far are serious games for non-research purposes.

Several examples can be given of previous serious games that have used a non-research purpose approach. Harteveld (2011) mentioned different applications that have been addressed: business and management, health, energy, military, politics and society, safety and crisis response, science and education. For example, some games in the area of health have been designed to increase players' knowledge about illness or health problems, support patients' rehabilitation, increase physical activity, encourage healthy lifestyle and help in therapy processes (Zwikael and Gonen, 2007). In addition, games designed for business and management have aimed at helping players to improve their decision-making process skills in complex real world scenarios (Boyle et al., 2011). Military approaches have also been addressed by serious games in, for instance, training soldiers on how to use military equipment and in providing a tool for helping returning soldiers suffering from post-traumatic stress disorder (Marsh, 2011). If we now take a look at serious games focused on energy consumption, different games have been designed to teach players how to responsibly consume energy or to produce in them behavioral changes towards energy conservation or environmental friendliness (Geelen et al., 2012, Bång et al., 2009, Gustafsson et al., 2009, Cowley et al., 2011). For example, *Power Exchange* is a mobile phone game for teenagers designed to test if players may learn how to reduce energy consumption in the long term in real life (Geelen et al., 2012). Players played the *Power Exchange* game for a week, period in which they interacted with other players in the game environment. Ten weeks after the game was finished their real life energy consumption showed a reduction of 14% on average. New habits related to energy conservation were created in the players that they didn't show before starting the game. The *Energy Battle* is another serious game aiming to reduce energy consumption in the long term (Geelen et al., 2010). The game was developed by Shifft, a spin-off company of Delft University of Technology. In the *Energy Battle* 17 different households from the city of Rotterdam in the Netherlands competed with each other for a money prize on what household could reduce their energy consumption the most. On average, the energy reduction during the game period was 24%, with the highest being 45%. One month after the game was finished and all the prizes given to the winners, the energy consumption of the same households showed an average reduction of 8% compared to the energy consumption before starting the *Energy Battle* game. The two previous examples show that serious games for non-research purposes in the energy consumption field can contribute to the long term household energy reduction.

3.2.2. Serious games for research purposes

The second type of serious game Meijer (2009) analyses is serious games for research purposes. In this second type of serious game the focus of the learning process is not the player anymore. Instead, the researcher is the one who learns from the players' behavior while playing the game. As a matter of fact, serious games for research purposes can be used as a way to generate hypothesis or theories, to test preexisting hypotheses and to evaluate multi-agents' behavior when people play in multiplayer settings. The purpose of a serious game for research purposes is to gather qualitative or quantitative data from the behavior of the participants in order to analyze it afterwards and get conclusions about how they behaved while playing the game. In other words, the game can be seen as a "laboratory experiment" in which researchers "play" with variables. By performing different tests and changing different variables, researchers could analyze under which circumstances people perform better or how people organize themselves given a set of conditions or stimuli. When a game is repeated sufficiently, the outcomes of sessions can be used as valuable data to test hypothesis (Harteveld, 2011). This approach gives policy makers a new tool to analyze complex problems from a different perspective and to test policies' acceptance in people (Meijer, 2009, Van Daalen et al., 2014). Apart from policy making, one may also think that serious games for research purposes could also be used to test new products or services from a commercial perspective in the future customers, in order to forecast the product or service market penetration before it is released to the market.

Despite the potential of serious games for research purposes in contributing to academic and scientific studies, they have not been used as much as games for non-research purposes. In fact, just a few examples of studies that have used this type of serious games were found in the existing literature. Psychological and social phenomena have used serious games, for instance, to analyze consumers' preferences when dealing with trade-offs and to test players' reactions after emotion-relevant events (Camerer and Fehr, 2002, van Reekum et al., 2004). Serious games to analyze socio-economic theories have also been designed to test how people trade commodities in which the quality attribute is not known (Meijer et al., 2008). Serious games for research purposes in the health area have also been designed to, for instance, measure respiratory capacity in healthy and asthmatic preschool children (Vilozni et al., 2005). Policy making has also used serious games for research purposes to test different policies, such as testing different regulatory styles in the energy sector and infrastructure development policies in railway systems (Kuit, 2002, Lo et al., 2013, Meijer, 2012). In addition, different supply chain management processes in different countries have also been analyzed and investigated by using this type of serious games (Zúñiga-Arias et al., 2007).

3.3. Strengths and weaknesses

As any other methodology, serious games have strengths and weaknesses. In this section we will analyze and explore the most important strengths and weaknesses that both types of serious games have.

3.3.1. Serious games for non-research purposes

After a literature review of serious games for non-research purposes, we have identified 4 main strengths. First, due to the high flexibility the serious games methodology provides, they can be applied to many disciplines. In fact, as previously discussed, serious games have been applied to business and management, health, military, politics and society, public policy, safety and crisis response, science and education, and so forth (Harteveld, 2011). Second, serious games for non-research purposes are more engaging and didactic than the conventional learning techniques, supporting the idea that this type of serious games are a good tool to effectively support learning processes (Boyle et al., 2016). Third, depending on how the game is designed, serious games for non-research purposes can have a high psychological impact on players, a feature that can be used to support learning processes in players (Ritterfeld et al., 2009). In fact, the psychologists Boyle et al. (2011) have analyzed how psychological aspects of players are influenced by video games in their research called *"The role of psychology in understanding the impact of computer games"*. Fourth, impossible or unreal situations can be simulated by video games that in real life would be impossible to get due to safety issues, cost and time constraints. Teaching people how to behave during an earthquake can be achieved by simulating different earthquakes in a video game. It is clear that in real life is impossible to produce controlled earthquakes, problem that is solved if they are simulated by using video games (Susi et al., 2007).

In addition, the literature review on serious games for non-research purposes also revealed 6 main weaknesses. First, after players are exposed to the game it is difficult to clearly state whether the learning process took place in the player or not. In fact, several researchers have tried to define methodological guidelines in order to solve this difficulty of assessing successfulness of serious games for non-research purposes. Surveys, questionnaires, tracing players' ex-post behavior are some tools that have been used. However, there is still no clear answer on how to be sure that players really learned from the game or not (Connolly et al., 2014). Second, the learning process in players may take long. Using video games do not ensure that the learning process goes faster than traditional learning techniques. In this sense, using serious games for non-research purposes may even be more expensive than traditional techniques, making them, eventually, not economically profitable (Marsh, 2011). Third, it is difficult to engage players to play a video game for learning, considering that, by definition, serious games do not have entertainment as their primary goal. If players are not engaged to the learning process it is difficult to actually succeed in teaching them the required knowledge or skill. As a result, player engagement in serious games for non-research purposes has also been analyzed by scholars in order to somehow include the "fun" into the serious games goals (Haferkamp et al., 2011). Fourth, demographic characteristics of players highly affect the penetration of the game into the possible players. Ritterfeld et al. (2009) shows that age and gender may have a strong influence on the penetration of games into the possible players. For instance, sports games are more popular among young males than they are in women. Fifth, although the game is designed in a very simple way and with very easy instructions, people with prior knowledge or experience with video games may show better learning experiences when playing serious games for non-research purposes than people without any prior experience (Boyle et al., 2011). Sixth, a serious game for non-research purposes should be as simple as possible in order to make the learning process of players easy. In addition, a serious game should also be realistic enough for the players in order to connect real life situations with the game environment if they are to apply those new skills or knowledge in real life. However, the more realistic a game is the more complex it gets. Balancing this trade-off is not easy to do and game designers must be well aware of this when defining its purpose (Harteveld, 2011).

3.3.2. Serious games for research purposes

As done in the case of serious games for non-research purposes, a literature review of serious games for research purposes was carried out. 6 main strengths were identified. First, serious games for research purposes are a good tool to generate or test theories or hypotheses and for testing players' behaviors under certain stimuli or conditions. This type of serious games become a simulated laboratory for experiments in which several phenomena can be tested (Meijer, 2009). Second, serious games for research purposes do not need to explicitly simulate psychological or sociological factors of players since they bring those characteristics with themselves into the game environment. When the rules of the game and the instructions are well defined, players will bring their own psychological characteristics into the game, such as their preferences, their emotions, their intuition, their soft skills, their social abilities and so forth. Therefore, the game does not have to assume or build up those aspects (Meijer, 2009, Katsaliaki and Mustafee, 2012). Third, experiments can be repeated as many times as the researcher wants due to the simplicity to carry them out. Experiments, or playing sessions, just need players to play the game and there is no need to spend extra resources to repeat experiments. Therefore, the cost of making mistakes are not as high as they are in other methodologies in which experiments are more cost intensive (Meijer, 2009). Fourth, and connected to the previous strength, experiments are flexible and they can be easily adapted (Meijer, 2009). Fifth, several psychological and sociological theories can be tested by using serious games for research purposes. Generally, testing those kind of theories or hypotheses are not easy to explore without assuming many different unmeasurable factors like psychological characteristics of people (Meijer et al., 2008). Sixth, the environment can be perfectly controlled by setting the right system boundaries and defining what it is included in the game and what it is not. The strict definition of system boundaries makes the model robust to external factors that can disrupt the outcomes or produce bias. Therefore, since system boundaries can be easily defined in serious games for research purposes, bias produced by external factors are minimized (Meijer, 2009).

In addition, we have identified 4 main weaknesses of serious games for research purposes. First, no matter how realistic and close to reality a game is, it will always be a game and not the reality. In that sense, people may play in a different way as they would do in real life just because they are facing an unreal simulation and not their real own experience. In that sense, validating the results in real life situations is a major challenge (Harteveld, 2011). Second, the time players need to play the game is usually considerably higher than any other research methodology which needs participants, like surveys. Most of the times in a serious game for research purposes many steps must be accomplished: the instructions of the game must be clearly explained, pre and post gaming experience surveys may be carried out, the gaming session itself and so on. All the steps combined might be too long for a player to voluntarily participate in the research, reason why it is another challenge to find a big enough sample size in order to draw reliable conclusions for the whole population under study (Meijer, 2009). Third, like in the case of serious games for non-research purposes, balancing realism of the game and complexity may be difficult. Especially for this type of serious game in which reality is a very important factor to consider in order to draw reliable conclusions for the whole population, realism should be as high as possible. However, in increasing realism, complexity increases too, reducing the reliability of the results if players found the game too complex. A good trade-off realism/complexity must be achieved, otherwise the results may be biased by a high game complexity or low realism (Harteveld, 2011). Fourth, each experiment or gaming session may be different to one another, resulting in a problem because eventually they may not be comparable. Each gaming session or experiment has its uniqueness characteristics due to the players' feedback and interaction. Researchers could reduce this problem by structuring a very rigid set of instructions and set up (Harteveld, 2011).

3.3.3. Summary

In Table 3.1 we have summarized the most important strengths and weaknesses of serious games for non-research and for research purposes. It provides a short summary of what it has been discussed in this chapter.

Table 3.1. Strenghts and weaknesses of serious games.

Type of serious game	Strengths	Weaknesses
Serious games for non-research purposes	- Applicable to a many different disciplines due to its flexibility - More engaging and didactic than traditional educational processes - They have a high psychological impact on players - They can simulate situations that are impossible to get in real life due to safety, cost, time, etc.	- Difficult to assess or measure whether the learning process took place in the player or not - Players' learning acquisition may take long - Since their goal is not entertainment, players should be highly motivated to play the game - Penetration of the game highly depends on demographic factors of players - The learning process in different people may also be different if players have prior abilities to play video games - Difficult to balance the tradeoff between realism and game complexity
Serious games for research purposes	- They can be used to investigate different theories and phenomena using players' results - They eliminate the need to build in psychological, sociological or cultural assumptions since players bring that into the gaming experience - Experiments can be easily repeated - Experiments are flexible - Psychological and sociological phenomena can be analyzed - System boundaries can be very well defined	- Simulated context, not real life experience (difficult results validation) - Requires more time for participants than other forms of research - Difficult to balance the tradeoff between realism and game complexity - Difficult to compare different playing sessions or experiments due to their unique characteristics.

3.4. Serious games for research purposes as a way to reduce the rebound effect's controversies

As previously discussed in chapter 2, the controversies around the rebound effect concept (its size and its importance for the policy making process) may be reduced if the causes of those controversies are reduced too (methodological issues, different numeric definitions and fuzzy and different system boundaries). In this thesis research we propose that serious games (and specifically serious games for research purposes) may be helpful in reducing rebound effect controversies by improving methodological issues of the previous attempts to assess and measure the effect (and specifically those attempts that have used before/after analysis). In fact, if we take a look at Table 2.6 we can find out that some of the methodological shortcomings of before/after analysis that have analyzed the rebound effect in the past are the strengths of serious games for research purposes in Table 3.1. In the following we will analyze why each methodological shortcoming of quasi-experimental analysis in Table 2.6 may be solved by using serious games for research purposes.

3.4.1. Difficulty to follow ceteris paribus

The first methodological shortcoming of before/after analysis of previous attempts to measure the rebound effect is the difficulty to follow the ceteris paribus principle. When performing before/after analyses, it is very important to change just one variable at a time in order to isolate the effects and calculate a rebound effect as a purely consequence of changes in energy efficiency and not as a consequence of changes in other variables. In real life before/after analyses this is hardly achieved because real life systems are dynamic in which many variables change at the same time and it is impossible to keep track of all the changes that the system may be exposed to. However, serious games for research purposes to carry out before/after analyses are a good tool to change just one variable at a time. Due to the fact that system boundaries can be very well defined and the experiments are carried out in a controlled game environment, the researcher can control all the changes in the important variables in order to make them happen one at a time, or in a ceteris paribus way. Flexibility in the experiments in serious games for research purposes can also contribute to reduce this shortcoming due to the fact that the researcher can totally control the changes inside the system, or even not make any other change to happen different than the energy efficiency improvements.

3.4.2. Difficulty to isolate effects

The second methodological shortcoming of quasi-experimental analyses is the fact that all the effects are hard to isolate. This shortcoming comes very close to the previous one in a sense that because of the dynamic characteristics of real life systems, effects are always intertwined with other effects, making it very difficult to isolate them. As a result, when real life quasi-experimental analyses are used to measure the rebound effect, they hardly achieve a pure rebound effect due to improvements in energy efficiency, and the final calculation may be influenced by some other hidden effects. Again, serious games for research purposes seem to help in solving this shortcoming because of the ease to well define system boundaries and the high flexibility to make different changes to happen inside the game environment.

3.4.3. Ex-post calculation

The third methodological shortcoming of quasi-experimental analyses that have measured the rebound effect in the past is the fact that they have all performed ex-post calculations, not giving any tool for policy makers to analyze the phenomena beforehand. Instead, just indications of how successful energy improvements policies were in the past is the only usefulness of those quasi-experimental analyses. Serious games for research purposes, on the other hand, provide a tool to ex-ante analyze policies that aim to improve energy efficiency, giving policymakers extra information of the possible consequences of those policies. In fact, since serious games for research purposes can be used for testing hypotheses, a hypothesis related to whether energy efficiency improvements are successful or not can be tested beforehand.

3.4.4. Absence of control group

The fourth methodological shortcoming of quasi-experimental analysis that have measured the rebound effect in the past is the fact that they have not had a clear control group in order to compare energy consumption over time after energy efficiency improvements. When we refer to a clear control group we mean a group in which no actions were taken, i.e. a group in which the energy efficiency remained the same. Most of the times energy efficiency is improved in both groups losing the essential characteristic of the control group. As a result, the comparison afterwards loses all the essence of before/after analyses: comparing two groups in which one of them was exposed to one specific stimulus that the other group was not. Serious games for research purposes may solve this issue if the researcher just lets to one specific group improve efficiency while the other remains the same. Using this type of serious game provides the required flexibility the researcher needs to clearly define a control group.

3.4.5. Single energy services analysis

The fourth methodological shortcoming of quasi-experimental analyses that have measured the rebound effect in the past is the fact that they have only measured the rebound effect in one single energy service, leaving possible indirect rebound effects outside the scope. The reason why this has happened is, again, the difficulty to isolate the effects that play a role in the system. When a researcher includes the possible indirect rebound effects between gas consumption and electricity consumption in households, for instance, the variables to control become too many to actually control them all in a proper way. As a result, most of the previous studies we have analyzed in this research that have used before/after analysis to measure the rebound effect have only focused on one single energy service. Nevertheless, when we think that serious games for research purposes provide the researcher a very flexible methodology to easily include or exclude effects from the system boundaries, we think that this methodology may be helpful in order to include some indirect rebound effects while keeping a reduced number of variables inside the system. In fact, and depending on how the serious game is designed, multiple indirect rebound effects can be included without compromising the possibility to isolate all those effects. The key is to clearly define what is in and what is out of the system boundaries and keep track of those indicators to separate the effects at the end of the research if necessary.

3.4.6. Neglection of people's psychological aspects

This may be one of the most important contributions of using serious games for research purposes to measure the rebound effect. So far, attempts to measure the rebound effect by using quasi-experimental analysis have not considered consumers' psychological aspects when they consume energy. Due to the fact that the rebound effect is a purely psychological phenomenon, deeply rooted in the consumers' behavior psychology, not considering this aspect is leaving the essence of it outside the system boundaries. However, serious games for research purposes provide the ease of including people's psychological factors when consuming energy without requiring doing extra assumptions or extra modelling techniques in order to consider them inside the system. Players are the ones who bring their own preferences, emotions and mind set into the game environment. In other words, players themselves bring their psychology into the game. As a result, no invalid or inaccurate assumptions have to be made to include the essence of the rebound effect into the analysis.

3.5. Disadvantages of using serious games for research purposes to assess the rebound effect

Despite all the benefits of using serious games for research purposes to measure the rebound effect, the main disadvantages of this type of serious games, jeopardize the measurement itself. First, since every serious game is a representation of the real world in a simulated environment, players may behave differently than they would do in real life. If this happens, the results of the rebound effect measurement may not be applicable to a real context. However, it must be mentioned that any model is a representation of the real world, none of them being a total mirrored representation of it. As a result, not only serious games, but also discrete, continuous or statistical models try to replicate the real world in a simulated way, producing differences between what the model concludes and what happens in real life. Nevertheless, this disadvantage can be partly investigated by analyzing how players behaved during the game, and if they would behave similarly in real life. This analysis can be carried out by performing pre and post-game questionnaires in order to check whether players effectively behaved according to their real mind set or they acted in a different way. However, in doing so, some inner disadvantages of surveys may be introduced, like the nonresponse bias or the lying bias (see section 2.5.1.3). Second, it is true that playing sessions take longer than other research methodologies, a situation that may cause players to lose focus on the game or to start playing in a meaningless way, introducing undesirable bias to the results. As a result, the chosen serious game for research purposes to measure the rebound effect must be engaging for the players to keep their attention and focus throughout the game session. Detecting players who lost their attention or got bored during the game session in order to exclude them from the final results may improve the overall quality of the research. Pre and post-game questionnaires can be used again to check how motivated players were during the session. Nevertheless, we may introduce, once again, nonresponse bias or lying bias. Third, the perfect balance between realism and complexity in a serious game is difficult to achieve. As previously discussed, a serious game must be realistic enough to make players behave in a similar way as they would do it in real life, and at the same time simple enough to not introduce undesirable bias due to an excessive complexity. Just like the previous two disadvantages of serious games for research purposes, checking how realistic and complex a game was for the players can be measured by performing post-game questionnaires in order to ask players their impressions and specifically how complex and how realistic the game was for them. Fourth, dealing with differences between different experiments or game sessions can also be challenging. When different players get different instructions, different stimuli or different treatments, the game sessions may take different paths, introducing bias to the results when those different experiments are to be compared. As a result, the researcher must make sure that those differences don't happen, by for instance, defining a clear set of instructions and goals for the players, in order to reduce those undesirable differences between different game sessions. Nevertheless, this disadvantage is hard to control due to the fact that many different sources of external stimuli can be present, introducing differences between experiments and game sessions. The best a researcher can do is to not give room for players to improvise and define a straightforward set of instructions and goals. The effectiveness of those set of instructions and goals can also be checked by post-game questionnaires.

To sum up, the general disadvantages of serious games for research purposes will also be present if the rebound effect is analyzed by using this technique. The researcher must be well aware of those disadvantages in order to reduce them, or at least measure them, to check afterwards how important they were. If the measurement of those disadvantages show that they didn't play an important role throughout the research, then the researcher may proceed with the data analysis, data validation or any other further step.

3.6. Chapter's conclusions

Because of all the reasons previously discussed, we believe that serious games for research purposes used as a tool to perform a quasi-experimental analysis to assess the rebound effect, will be a contribution to reduce methodological shortcomings of previous attempts to measure it. As a result, controversies around the rebound effect may also be reduced. The previous conclusion comes despite the fact that new shortcomings may be introduced as well. If those new shortcomings are analyzed and measured, then the researcher might analyze if they played an important role into the results or not.

As a result, this thesis research aims to assess the rebound effect by using serious games for research purposes in order to investigate the contribution of this new methodology. In this point we are going to accept what Aydin et al. (2015) and Sorrell (2009) argue about the uniqueness characteristics of the effect: the fact that the rebound effect must be independently analyzed in each specific case and situation, and that any new study should not be focused on calculating one single and unique rebound effect size because that unique size does not exist. Consequently, new researches that aim to reduce the rebound effect's controversies should be focused on improving the overall quality of the assessments, in a way that the controversies are not reinforced but diminished. Finding one single and specific rebound effect size is not what the scientific community needs; instead, it needs studies that stop trying to find consensus on the absolute magnitude of the rebound effect which distracts the attention of the scientific community on important matters, like evaluating and planning strategies that minimize the rebound effects in the many contexts and realities where they occur (Friedrichsmeier and Matthies, 2015).

4. The NRG game

The serious game that will be used to assess the rebound effect in a way to reduce its controversies is called the NRG game. The NRG game was created at the faculty of Technology, Policy and Management of Delft University of Technology in the Netherlands. In the present chapter the NRG game is described in further detail. This chapter first starts by describing the main characteristics of the game. Second, it describes the most important characteristics of the experiments and the modifications needed to be done to the NRG game to make it a suitable tool to analyze the rebound effect. Third, how the experiments were carried out is also described. Fourth, the type of the rebound effect included in the assessment is analyzed. Finally, a short conclusion of this chapter is given.

4.1. General description

The NRG game was created first as a serious game for research purposes in order to test different interventions such as information, feedback, discounts and subsidies towards energy conservation in households (Mohammed et al., 2015). For the previous purpose, the NRG game simulates the basic decisions people make in their households regarding energy consumption. Players start with a certain budget and they receive an income after each round of the game, simulating the part of the income people spend in paying energy bills, buying new appliances and furniture. As a result, the players' goal is to manage their household in the game in terms of energy consumption. Players must pay gas and electricity bills, can buy new appliances, produce energy using solar panels or wind turbines, sell appliances to get some money back and so forth. In order to check how they are doing in their game performance, they can check their energy consumption and comfort level at any time.

Energy consumption and comfort level are the most important indicators in the game. When buying new appliances, players can decide if they prefer a more luxurious item (that will change the comfort level) or a more eco-friendly one (that will affect the total energy consumption). When players decide to increase the luxurious status of the new item the price increases accordingly, and when players decide to increase the eco-friendly level of the new item the price increases accordingly as well. Therefore, players deal with constant trade-offs between the price of the item and the comfort/eco-friendly level of it. Depending on all the possible decisions players can make, their energy consumption and comfort level changes accordingly, and so do the gas and electricity bills.

The main screen of the game is shown in Figure 4.1. It shows the house each player is playing with and the most important features of it. If players click in number 1, they can go up and down to see their houses' floors. Figure 4.2 shows what players see when they select to go to the second floor of the house, for instance. If players click in number 2, they can see the catalog with all the available things they can buy (Figure 4.3 (a)). If they select one item and click on it, they will see the luxury and eco-friendly selection of that item (Figure 4.3 (b)). In this last selection is where players face the mentioned trade-offs between price, energy consumption and comfort. They can drag and drop the luxury and eco-friendly level of the item. The more luxury level the item has, the more expensive it gets and the more comfort level it adds. The more eco-friendly level the item has, the more expensive it gets and the less energy it consumes. Therefore, they must balance their own desires and interests regarding energy consumption/comfort in terms of the price of each item. If players now click on number 3 in Figure 4.1 they can remove items and get some money back (as if the item is sold). Number 4 in Figure 4.1 shows the money players have available to spend. After each round of the game players receive their annual income (see section 4.2.4). If players click in number 5 in Figure 4.1 they can check their performance in the game: their energy consumption, their energy production (if they decided to produce energy), their total comfort level, the condition of all the appliances and furniture they have in their houses, the total money they are paying in electricity and gas after each round and the annual income they receive after each round.

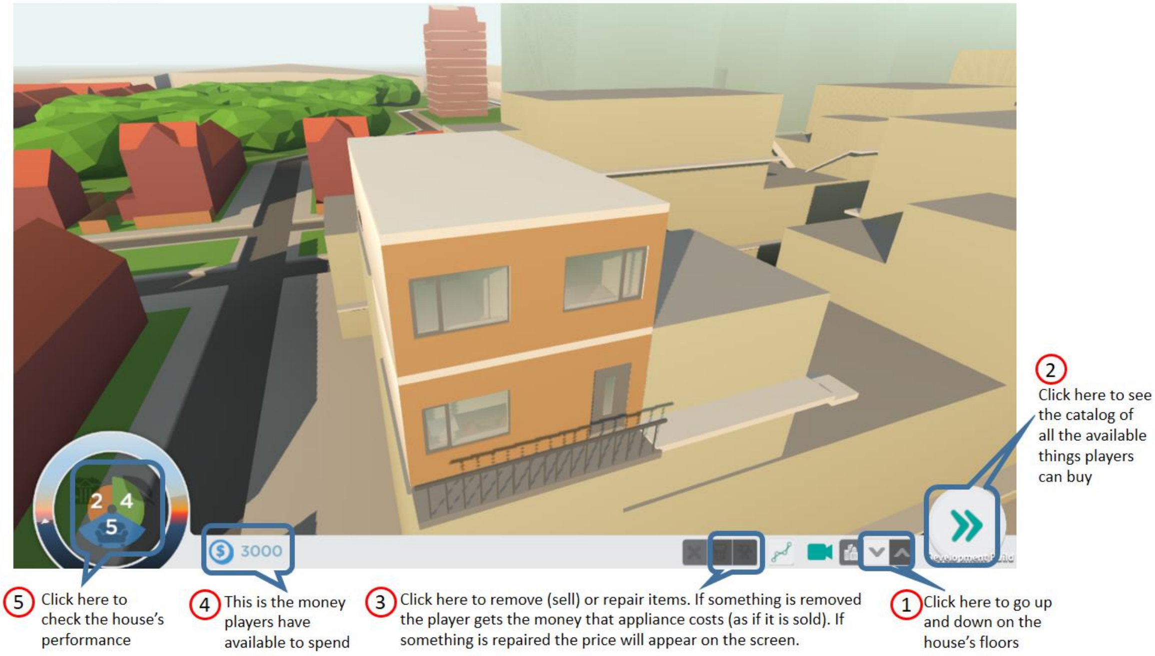


Figure 4.1. Main screen of the NRG game with the main options.

Figure 4.2. Second floor of the house.

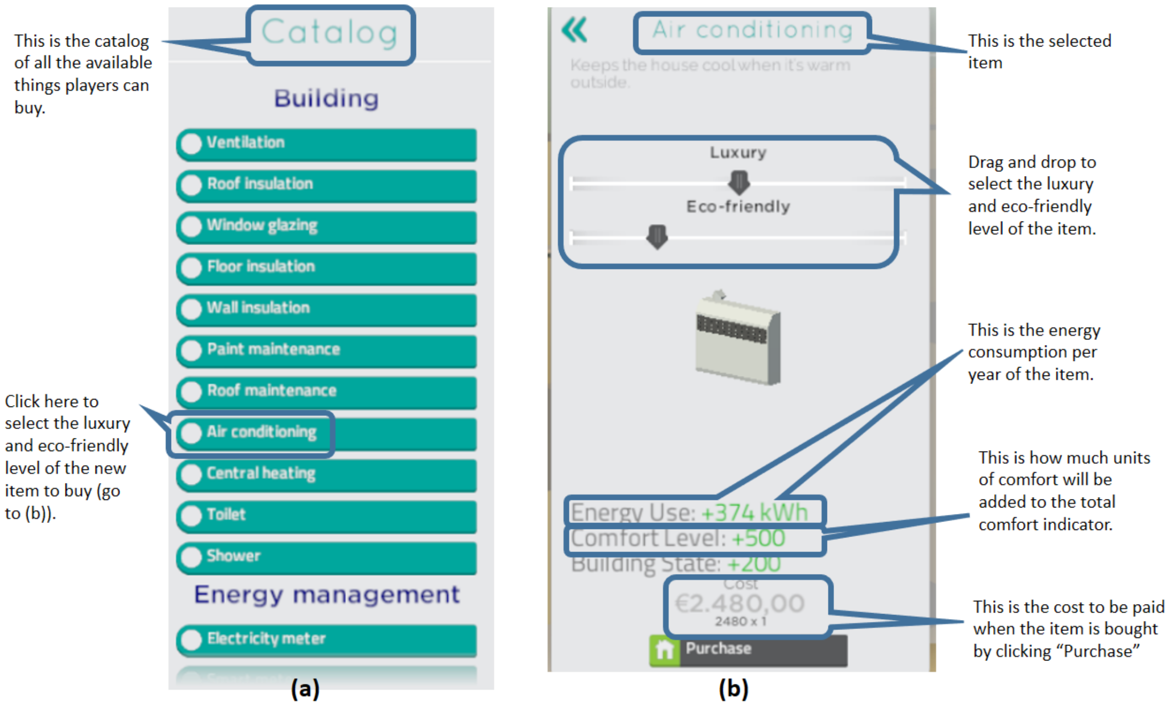


Figure 4.3. (a) Catalog of available things to buy. (b) Luxury and eco-friendly selection of the item to be bought.

In terms of all the things players can do in the NRG game, we can list five different decisions they can take. First, players can remove appliances as if they are selling them and get some money back. This can be done with any gas/electric appliance and furniture in the house. Second, players can improve the efficiency of the appliances they already have by paying some money, and therefore, decreasing the energy consumption of the selected appliance. Third, players can buy new gas/electric appliances and furniture according to their own desires. In the catalog of all the available things to buy in the game (Figure 4.3 (a)) players can check all the 68 available appliances and furniture. Fourth, players can buy new energy management devices, which are devices that manage to reduce the overall energy consumption of a house by a certain percentage. In the NRG game, energy management devices are smart meters, movement sensors or stand-by-killers, if we talk about electricity reduction, or insulation (wall, floor, windows, and so forth) if we talk about gas reduction. Fifth, players also have the opportunity to buy energy production devices which are devices that produce usable energy. In the NRG game, energy production devices are wind turbines, solar panels or heat pumps. To sum up, if we want to analyze what decision influences the players' total energy consumption the most, we should focus on those five decisions previously described. In section 5.5 we break down the total energy consumption in terms of those decisions.

The NRG game gives automatic results of each player. Every decision a player takes is recorded in his/her own output, including the luxury level and eco-friendly level they wanted the new items to be. In addition, the players' output gives the updated total energy consumption and comfort level after each decision. Therefore, it is easy to track those two indicators.

Due to the fact that the NRG game was not designed for the purpose of assessing the rebound effect, several modifications were needed to be done in order to make the NRG game a suitable tool for analyzing this phenomenon. The following sub-chapter shows the main modifications that were needed to make.

4.2. NRG game to assess the rebound effect

Several modifications of the first version of the game were needed in order to make the NRG game a suitable tool to perform a quasi-experimental analysis to investigate the rebound effect. In the following those main modifications are discussed.

4.2.1. Two possible houses options

The NRG game gives players the possibility to choose the house they would like to live in, in terms of three indicators: energy level, comfort level and building state level. According to their selection, they receive a house out of 7 available houses. All the possible houses are different in terms of their energy consumption, type of initial appliances, house size and house comfort.

If we want to analyze the rebound effect using the NRG game, the first thing to consider is finding a way to reproduce and simulate a quasi-experimental analysis or a before/after analysis. For that reason, the possibility to have 7 different house types needs to be reduced to only two possible houses (one of them representing the case “before” and the other one representing the case “after” efficiency improvements). One type of house will represent one efficiency level and the other house type a different efficiency level. If those two groups are compared, an indication of the rebound effect may be obtained.

As a result, just two possible houses will be available for players. The one and only difference of those two houses will be the energy efficiency level. One will be called “low efficiency house” and the other one “high efficiency house”. In order to not introduce selection bias into the sample, each player will randomly receive one type of house and it won’t be possible to change it. The main idea underlying this choice is to simulate the situation in which one single person in real life starts using his/her own real house and after a while the energy efficiency of it changes. In this case, the situation before and after efficiency improvements will be analyzed by two different groups of people, each of them using one specific house with one specific energy efficiency level. The low efficiency houses group will be used as a control group for the analysis, in which their behavior will be taken as the situation in which efficiency does not change, or the case “before” efficiency improvements. However, what this methodology supposes is the fact that the two groups of people are indeed comparable, result that must be analyzed before any further analysis. If we obtain that the two groups are indeed comparable, we could get an estimate of the rebound effect comparing the situation “before” efficiency improvements and the situation “after” efficiency improvements.

It is worth to mention here that this is not a typical before/after analysis due to the fact that a typical one considers that just one group is analyzed. There is one group of people that starts the experiment under a certain energy efficiency condition and at some point in the middle of the experiment that efficiency is improved. In this way the same individuals are exposed to efficiency improvements and their own behaviors before and after efficiency improvements are compared. In this thesis research we propose to perform a modified before/after analysis in which there are no individuals who are exposed to efficiency improvements. Alternatively, we propose to separate the case “before” and the case “after” efficiency improvements by analyzing two different (but comparable) groups of people. The reason behind this decision is rooted in the long duration of the experiments when serious games are used. If one player is exposed to efficiency improvements in the middle of the game, there might not be enough time of the game left to properly analyze the impact of the efficiency improvement. If we can analyze from the beginning of the game the impact of this stimulus, the results may be richer and more conclusions may be obtained. The latter can be achieved by separating the situation “before” and the situation “after” efficiency improvements and assign them into different groups, just as we propose to proceed in this research.

4.2.2. Initial appliances and consumption

As mentioned in the previous sub-section, just two types of houses will be available in the game in order to simulate a modified before/after analysis. In order to not make the same mistake that previous attempts to measure the rebound using quasi-experimental analysis have made, we need to make sure that the ceteris paribus principle is followed, that the rebound effect is isolated by including in it just energy efficiency improvements and that a clear control group is defined (see Table 2.6). Those three aspects are well tackled if we make sure that the one and only difference between the high and low efficiency house is the energy efficiency level of each of them. The NRG game is flexible enough to include this into the game settings. As a result, the two types of houses will be designed in a way that they both have the same appliances at the beginning of the game and the only difference is the efficiency level of those appliances, resulting in a different energy consumption. Since the NRG game also gives an indication about the comfort level of players, this index must be equal in both types of houses. In doing so, the possible differences in the energy consumption of the two types of houses throughout the game will be caused uniquely by a different initial energy efficiency level and not by some other factors that may introduce bias to the results.

The set of initial appliances included in the houses are the ones that were most common in Dutch households in the year 2000 according to Fawcett et al. (2000) and Papachristos (2014). The energy consumption of the high efficiency house was defined to be equal to the average gas and electricity consumption in the Netherlands in 2012, i.e. 3,510 kWh of electricity per year and 1,314 m^3 of gas per year (Gerdes et al., 2014). On the other hand, the energy consumption of the low efficiency house was defined as the difference in energy efficiency between a typical Dutch household in 2012 and in 2000. According to Gerdes (2015), Dutch households in 2012 were on average 30% more efficient than they were in 2000. As a result, the low efficiency house in the game was defined as a typical Dutch household from 2012 keeping the energy efficiency from the year 2000, i.e. a consumption equal to 4,575 kWh of electricity per year and 1,703 m^3 of gas per year. Additionally, the distribution of the total electricity and gas consumption among the selected appliances to be included by default in the houses, was calculated using standard energy consumptions patterns of those basic appliances (check Fawcett et al. (2000), Gerdes et al. (2014), Friendly and Kormylo (2012) and

U.S.DepartmentOfEnergy (2016)). In appendix A a list of all the appliances that houses receive by default with their initial energy consumption is shown.

4.2.3. Prices and consumption update

Due to the rapid changes in the energy sector, energy and appliances prices are dynamic. As a result, a complete review of the prices and consumption of all the appliances available in the game was carried out. For example, the energy prices were updated to the average prices in the Netherlands in 2015, i.e. the electricity price equal to 0,2 €/kWh (Eurostat, 2015a) and the gas price equal to 0,8 €/m^3 (Eurostat, 2015b). In addition, prices of energy production devices and their standard energy production were updated as well according to the latest available technologies and prices from the manufacturers (check TheEcoExperts (2016), CleanTechnica (2016), WindTurbines (2016) and Hansen (2016)). Finally, the rest available appliances were also reviewed. Some of their prices and energy consumption were updated according to Friendly and Kormylo (2012) and the U.S.DepartmentOfEnergy (2016).

4.2.4. Income

A critical factor in order to give the game the required realism is the income players receive after each round. According to Aydin et al. (2015), the rebound effect is strongest among the lower income levels, reason why a good selection of the income is required if we want to test a realistic rebound effect. It is crucial as well that the income players receive reflects just the money people spend in real life on paying energy bills, buying electricity appliances, buying gas appliances and buying furniture, excluding the part of the income people spend on, for instance, food, transportation, leisure activities and so forth.

The Dutch Central Bureau of Statistics (CBS) keeps track of the Dutch expenditure. In 2014, each Dutch household spent on average a total of € 1,753 in energy bills (electricity and gas). In addition, Dutch households spent in 2014 a total of € 9,262 million in new gas and electricity appliances and furniture. If we now divide the previous number by the total amount of households in The Netherlands in 2014 (7.4 million households), we can obtain that each Dutch household spent on average a total € 1,252 in 2014 in new gas and electricity appliances and furniture. If we add up the € 1,753 in energy bills and € 1,252 in new gas and electricity appliances and furniture, we obtain that Dutch households spent on average a total of € 3,005 in those aspects in 2014. All the required information to make this calculation was obtained from CBS (2014b), CBS (2014a) and CBS (2015). Therefore, it was defined that each player receives an income equal to € 3,000 per round, in line with the previous calculation. As a result, it was also defined that one round of the game equals to one year.

4.3. About the experiments

Having described the main modifications of the NRG game in order to make it a suitable tool to assess the rebound effect, this sub-chapter will describe what the experiments are about. As previously described, there will be two groups of players: one group playing the NRG game using the high efficiency house and a second group using the low efficiency house. It is worth to remember again that the one and only difference between those two groups is the energy efficiency, resulting in a different initial energy consumption. In addition, and in order to increase the statistical significance of the results and to get a significant representative behavior of each group, the sample size of each group must be as large as possible. Due to the time it takes for players to complete each game experiment (1 hour on average per player), it is difficult the get a large enough sample size. However, it was decided that at least 25 players must play the NRG game per group, getting a total sample size equal to 50 people.

In order to increase the quality of the results and to test if the two groups are indeed comparable, a pre-game questionnaire is carried out (Appendix B shows the pre-game questionnaire). The idea of it is to test some demographics characteristics of the players and to evaluate their comparability in terms of, for instance, their environmental friendliness. It is natural to think that if we want to compare the energy consumption of two groups they must have the same environmental attitude, otherwise the results may be influenced by other factors different than the rebound effect, like having one group with a different environmental attitude than the other. In section 5.2 this comparison is carried out. After players have answered the pre-game questionnaire, they receive the player's manual with all the instructions and some screen shots of the game for them to know where to find the necessary options in the game. The player's manual can be found in appendix C. After the player's manual, players started playing the NRG game in a single player mode, without any interaction with other players in the virtual gaming environment. In total, participants played 10 rounds of the game (simulating 10 years). The first round of the game was designed to last 8 minutes in order to let players get familiar with the game functioning, and after that each round lasted 4 minutes. However, when players needed more time to finish one specific round, they received some extra time. If we add up all the rounds together, participants played the NRG for approximately 45 min in total. During that time, players were allowed to ask any question they wanted about the game, and to freely interact with the rest of the participants. When round 10 was over, participants answered a post-game questionnaire about their strategies and impressions about the game. This questionnaire was designed to discover the main reasons of why players behaved in a certain way. Appendix D shows the post-game questionnaire. To sum up, one single experiment took approximately 1 hour in total (including questionnaires, reading manual and playing).

Before starting the final experiments with the players, two persons played the game and answered the questionnaires as a trial test. They were interviewed after completing their experiments in order to get their opinions and points for improvements. Those two interviews with the questions asked and their answers can be found in appendix E.

Finally, the game experiments started. The population of this study was defined as "adults living in the Netherlands". After several playing sessions between the 20th of April 2016 and 21st of May 2016, the target of 25 people per group was achieved, basically with students from the Delft University of Technology in the Netherlands. Some playing sessions were carried out with multiple players and some others had to be carried out with just one player at a time due to the difficulty to get participants. Figure 4.4 shows a playing session carried out in the faculty of Technology, Policy and Management in Delft University of Technology on the 10th of May 2016, in which 9 people played at the same time.

Figure 4.4. Playing session in the faculty of Technology, Policy and Management in Delft University of Technology, the Netherlands.

4.4. About the type of the rebound effect

It is worth to mention now that, due to the way the experiments were designed, the rebound effect assessment that will be carried out considers mainly components of the indirect rebound effect (section 2.1). The direct rebound effect is not immediately included in the experiments due to the fact that in the NRG game players do not have the opportunity to increase the number of hours in which they are using certain appliance or increase the energy intensity of those. For instance, players can't decide to turn the TV on more hours per day or use the microwave at a higher temperature. The energy consumption of each appliance is fixed and defined by the initial selection of the eco-friendly level of the items. Nevertheless, what it is well captured by the experiments in the game is the indirect rebound effect. Players have the opportunity to buy or remove appliances any time they want, including in the game energy consumption exchanges between different appliances and between different energy services (gas and electricity). For instance, players may decide to remove the sound system that was included by default in the house they receive in the game, and change it by a game console. In this way they may be replacing energy of one appliance with energy from another appliance. In addition, the indirect rebound effect is also seen when players replace the energy consumption coming from an electric appliance and replace it with energy coming from a gas appliance, or the other way around. This can happen if, for instance, players replace their gas ovens or gas stoves for electric ones. Finally, the economy wide rebound effects are not captured by this research since no macro-economic concepts were included in the game.

4.5. Chapter's conclusions

The serious game that will be used to assess and analyze the rebound effect is called the NRG game. The NRG game simulates the most basic decisions people take in their households when consuming energy (gas and electricity). Therefore, players in the game can not only buy new appliances, sell appliances they don't want or buy energy production devices like solar panels, but also the must pay energy bills in line to their actual energy consumption. In addition, players also receive a certain income after each round (year) of the game in a way to simulate the part of the income people spend to manage the energy consumption of their houses. Due to how the NRG game works and what type of decisions players can take, the rebound effect that will be included in the experiments is mainly the indirect rebound effect. The direct rebound effect is not immediately included due to the fact that players can neither decide to increase the number of hours of use of certain appliance nor increase the energy intensity of them. The previous feature is not included in the game, resulting that the direct rebound effect is not considered in the system boundaries.

The NRG game in its first version is not sufficient to assess the rebound effect, therefore, several modifications had to be done prior the experiments with players (number of houses available, prices revisions, income players receive, set of initial appliances, and so on). The experiments were designed in a way that two different groups of people are analyzed: one first group in which players have a high efficient house and a second group in which players have a low efficient house. The group with low efficiency houses will be the control group and it will be used as the case “before” energy efficiency improvements, and the group with high efficiency houses will be the case “after” energy efficiency improvements. The latter constitutes a modification of a typical before/after analysis in which the case before and the case after are assigned to different groups of people. If comparisons between the two groups are carried out (proving beforehand that they are, indeed, comparable) the rebound effect may be detected.

5. Results

Having finished the playing sessions in which the target sample size of 50 people was achieved, the analysis of the results was carried out. In this chapter the main results are discussed.

5.1. Descriptive statistics

As previously said, a total of 50 people played the NRG game, 25 people using the low efficiency house and 25 using the high efficiency house. If we take a look at the gender distribution of the participants, a total of 18 players were female and 32 were males (Figure 5.1).

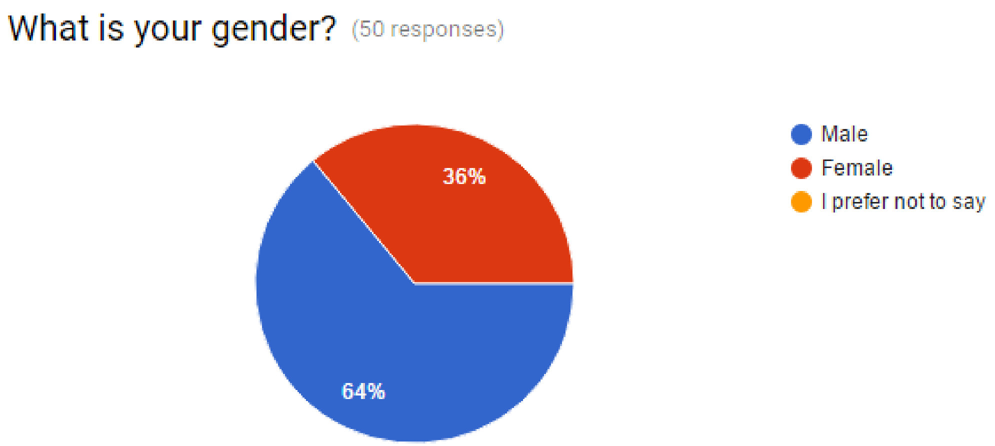


Figure 5.1. Gender distribution.

If we now take a look at the education distribution, we will see that 41 players (82%) have a postgraduate education or currently enrolled in one (Figure 5.2).

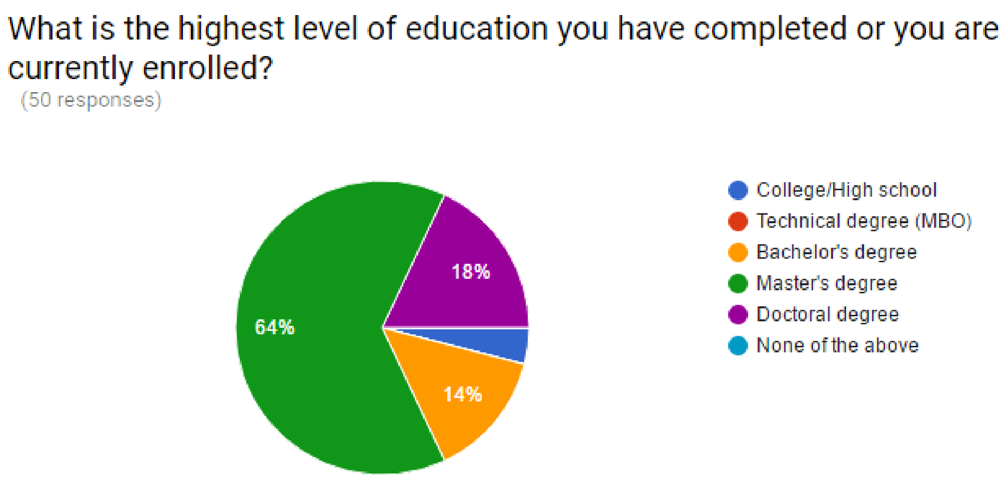


Figure 5.2. Education distribution.

The occupation distribution in Figure 5.3 shows that 32 players (64%) are students and 18 are employed or self-employed (36%).

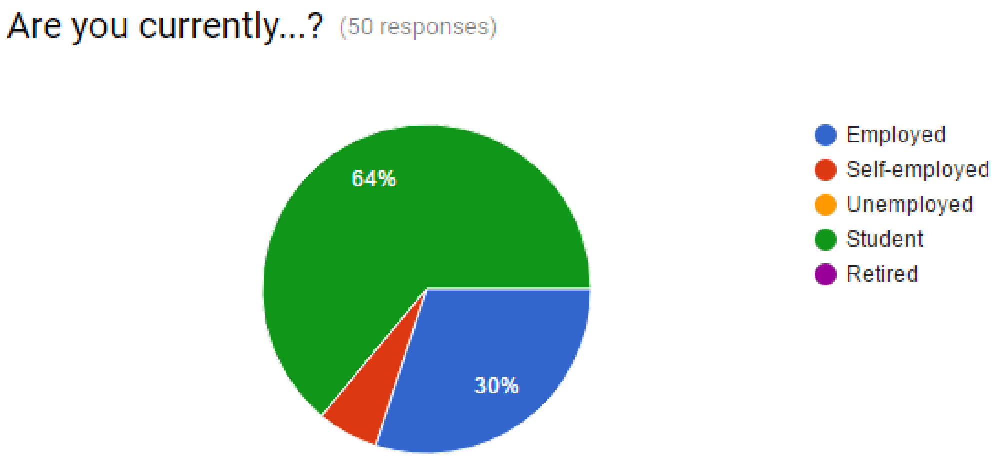


Figure 5.3. Occupation distribution.

The complete descriptive statistics with all the graphs can be found in appendix F with the pre-game questionnaire answers. Additionally, appendix G shows all the post-game questionnaire answers.

5.2. Comparability of the two groups and representativeness

A crucial factor to analyze is the representativeness of the sample and the comparability of the two groups for further analysis. First, we must say that the sample is not a good representation of the population, since the latter was defined as "adults living in the Netherlands". In Figure 5.1, Figure 5.2 and Figure 5.3 we saw that the sample is mainly composed of students, people with postgraduate studies and more males than females. Consequently, the final results of this research can't be applied to the whole population. Alternatively, the population could have been defined as just "students living in the Netherlands", situation that would have been too restrictive and narrowed, producing that some respondents should have been removed from the sample, reducing the sample size and compromising the significance of the results. For the latter reason it was decided that the population was going to be kept as "adults living in the Netherlands".

Second, the comparability of the two groups is crucial if we are to compare them both to analyze the rebound effect, otherwise further comparisons may be biased by differences in those groups. In this research we have decided to analyze this by checking how similar both groups are using four criteria: environmental friendliness, abilities to play video games, how often participants play video games, and whether players pay bills, rent or mortgage in real life. If one of those four factors show statistically significant differences between the two groups, then the comparability of the results may be compromised.

5.2.1. Environmental friendliness

If one group shows a different environmental friendliness attitude than the other it would be difficult to compare their energy consumption, since environmental friendliness and energy consumption may have some kind of significant relation. In this research we have defined the environmental friendliness of players as the average of questions 13 and 14 in the pre-game questionnaire (check appendix B)[2]. Figure 5.4 shows the environmental friendliness index in both groups. The high efficiency group shows an average of 0.76 with a standard deviation of 0.367, while the low efficiency group shows an average of 0.89 with a standard deviation of 0.481.

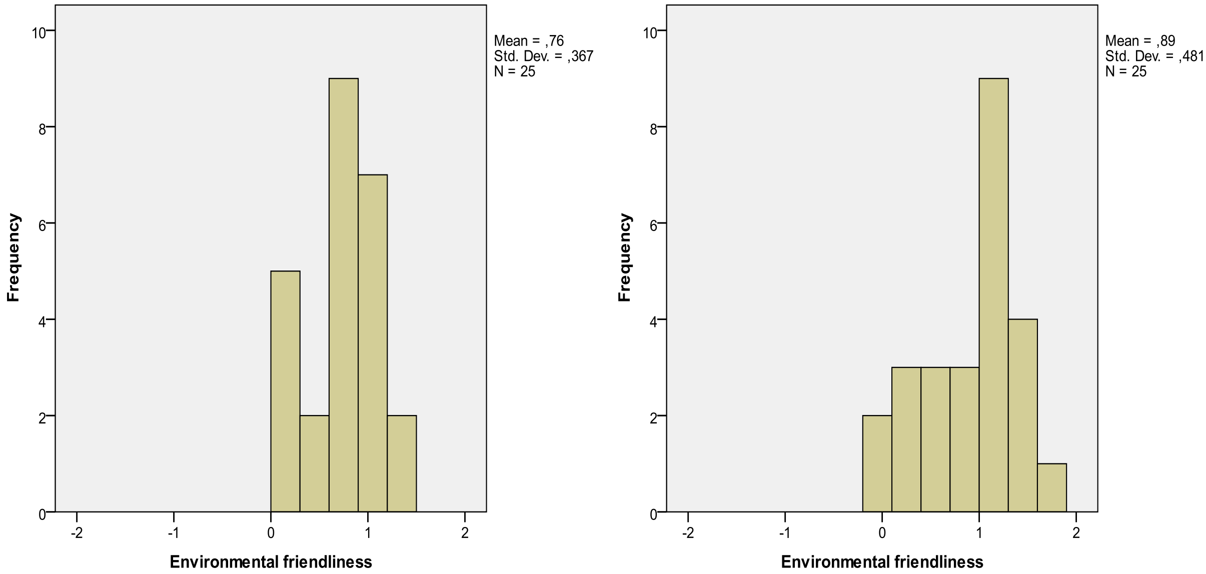


Figure 5.4. Environmental friendliness index high efficiency group (on the left) and low efficiency group (on the right).

If we want to test if both distributions show any significant difference, we can use student t-test for differences in averages if the right conditions are satisfied. First, both groups must be normally distributed, which is proved by testing it with the non-parametric test Kolmogorov-Smirnov test. In this case the null hypothesis H_0 is "the environmental friendliness index follows the normal theoretical distribution", while the alternative hypothesis H_1 is "the environmental friendliness index does not follow the normal theoretical distribution". After performing the test in SPSS, we conclude that the null hypothesis in both groups can't be rejected since the significance value is greater than 5% (16.3% for the high efficiency group and 15.7% for the low efficiency group). As a result, it will be assumed that both group follow a normal distribution. Second, the two samples must be randomly drawn from the population. This condition is difficult to check since players were not chosen to participate in the research in a random way, they expressed themselves their willingness to be part of it. Third, both sample groups are independent. This condition is indeed satisfied because of the way the experiments were carried out, totally isolating both groups to one another in terms of the

[2] The scale used to transform categorical variables into numeric ones to average them is the following:
Totally disagree = - 2; Disagree = - 1; I am not sure = 0, Agree = + 1, Totally agree = + 2

houses they played with in the game. As a result, although the second condition can't be checked, we will assume that we can proceed with the student t-test for differences in averages.

Before performing the student t-test for differences in averages we must define the hypothesis. The null hypothesis H_0 is "there is no difference between the two groups in terms of the environmental friendliness" while the alternative hypothesis H_1 is "there is a difference between the two groups in terms of the environmental friendliness". After performing the test in SPSS we obtain a significance value equal to 29%, leading to the conclusion that H_0 can't be rejected. As a result, no differences in the environmental friendliness index of both groups were detected. Therefore, it will be assumed that both groups can be compared in terms of their environmental friendliness index. The SPSS results can be found in appendix H.1

5.2.2. Abilities to play video games

According to Boyle et al. (2011), people with higher abilities to play video games may show a better performance while playing it. The latter can be translated into the NRG game, for example, in the ease of finding the right options in the game. People with more abilities to play video game may find more and different ways to manage their houses' energy consumption in the game than people with less abilities. As a result, if both groups show some differences in their abilities to play video games then the final results may be biased by this factor and the results may not be comparable to one another. To check this, we have used question 10 of the pre-game questionnaire (appendix B). In this case, the chi-square test must be used because both variables have a categorical level of measurement. Since the conditions to perform this statistical test are not satisfied (more than 20% of the expected counts are smaller than five), some categories were grouped together to reduce the number of expected counts smaller than five[3]. Figure 5.5 shows a bar chart of the abilities to play video games per group with the new categories.

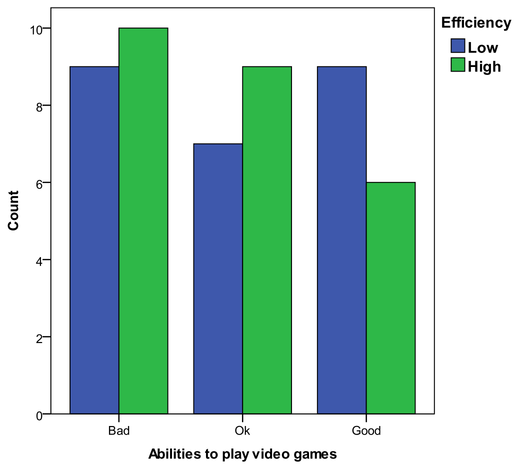


Figure 5.5. Bar chart of the abilities to play video games per group.

[3] The new categories are: Good = "Good" and "Excellent"; Ok = "Ok"; Bad = "Fair" and "Poor".

The null hypothesis for this test H_0 is “there is no difference between the two groups in terms of the participants´ abilities to play video games” while the alternative hypothesis H_1 is “there is a difference between the two groups in terms of the participants´ abilities to play video games”. After performing the chi-square test in SPSS using the new categories, we obtain a significance value equal to 63.7% and there are no expected counts smaller than five (conditions are satisfied). As a result, H_0 can’t be rejected. Therefore, no differences between the two groups in terms of the participants´ abilities to play video games were detected. As a result, it will be assumed that both groups can be compared in terms of their abilities to play video game. The SPSS results can be found in appendix H.2.

5.2.3. How often participants play video games

Similar to what happens in the prior abilities to play video games, participants who play video games more often than others may find more and different alternatives to manage their house in the NRG game. To check that there are no differences in how often participants play video game we have used question 9 of the pre-game questionnaire (appendix B). Figure 5.6 shows a bar chart of the distribution of how often participants play video games per group.

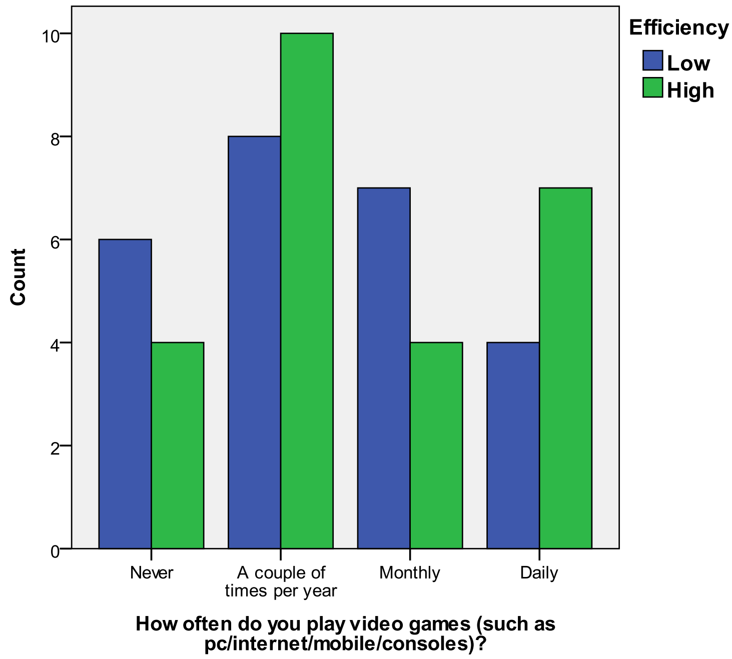


Figure 5.6. Bar chart of how often participants play video games per group.

Once again, the chi-square test will be used since both variables have a categorical level of measurement. The null hypothesis for this test H_0 is “there is no difference between the two groups in terms of how often participants play video games” while the alternative hypothesis H_1 is “there is a difference between the two groups in terms of how often participants play video games”. After performing the chi-square test in SPSS, we obtain a significance value equal to 52.1% and there are no expected counts smaller than five (conditions are satisfied). As a result, H_0 can’t be rejected. Therefore, no differences between the two groups in terms of how often participants play video games were detected. As a result, it will be assumed that both groups can be compared in terms of how often participants play video games. The SPSS results can be found in appendix H.3.

5.2.4. Players paying bills, rent or mortgages in real life

It may happen that players who pay the bills, rent or mortgage themselves in real life show a different behavior in the game. If a person does not pay anything in real life he/she may not care so much about the finances of his/her household in the game, introducing undesirable bias. It is because of the latter reason why it becomes important to check if both groups are formed by the same type of people: people who pay and people who don't pay in their households in real life. In order to check this, question 11 of the pre-game questionnaire will be used (appendix B). Once again, the chi-square test is the right statistical test to use because we are dealing with categorical variables. Just as happened while testing the abilities of players to play video games, the conditions to perform the chi-square test are not satisfied (more than 20% of the expected counts are smaller than five). To overcome this problem, some categories have been grouped together[4]. Figure 5.7 shows a bar chart of the distribution per each group of the variable who pays for the rent, bills or mortgage in your household using the new categories.

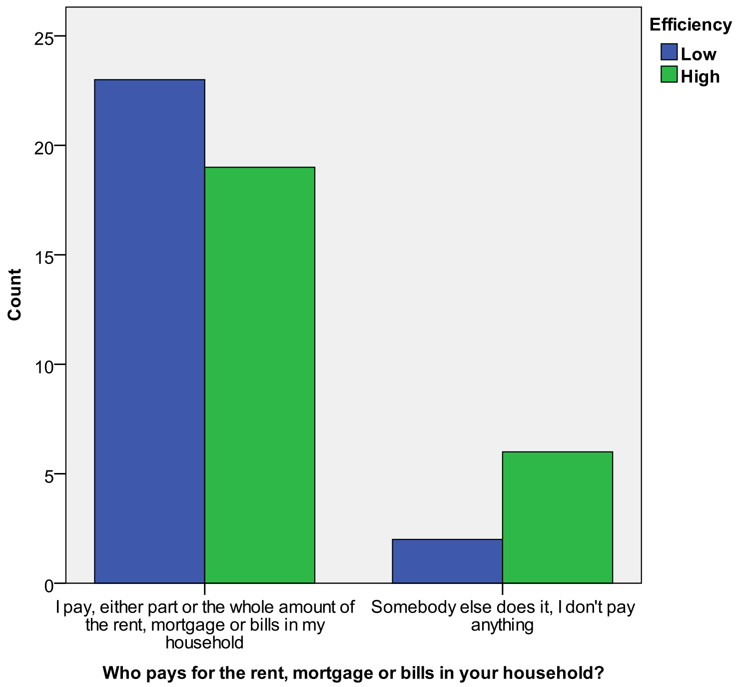


Figure 5.7. Bar chart of who pays for the rent, bills or mortgage in your huseholds per group.

The null hypothesis for this test H_0 is "there is no difference between the two groups in terms of who pays the rent, bills or mortgage in their households" while the alternative

[4] The new categories are: I pay, either part or the whole amount of the rent, mortgage or bills in my household = "I pay the whole amount of the rent, mortgage or bills in my household" and "I pay part of the whole amount of the rent, mortgage or bills in my household"; Somebody else does it, I don't pay anything = "Somebody else does it, I don't pay anything".

hypothesis H_1 is "there is difference between the two groups in terms of who pays the rent, bills or mortgage in their households". After performing the chi-square test in SPSS using the new categories, we see that the conditions, once again, are not met to perform the chi-square test (2 cells have an expected count smaller than 5). As a result, we should analyze the Fisher's Exact Test, which is computed automatically by SPSS when the conditions of the chi-square test are not satisfied. As a result, the significance level of the Fisher's Exact Test is equal to 24.7%, leading us to conclude that H_0 can't be rejected. Therefore, no differences between the two groups in terms of who pays the rent, bills or mortgage in their households were detected. As a result, it will be assumed that both groups can be compared in terms who pays the rent, bills or mortgage in their households. The SPSS results can be found in appendix H.4.

5.3. Energy consumption

Having shown that the two groups of players (using high efficiency houses and low efficiency houses) can be compared to one another, we can visualize and analyze their energy consumption throughout the game. Figure 5.8 shows the total energy consumption of players per round. The upper graphs show the consumption of each player represented by a cross and the lower graphs show the box plot of all of them per round. It must be mentioned that the total energy consumption was calculated transforming the gas in m^3 into kWh in order to get one single unit of energy. The transformation factor used for the conversion was 9.769 kWh per m^3 of gas (Energieconsultant, 2016).

Figure 5.8. Total energy consumption per round per player.

If we analyze the graphs we can see that the high efficiency houses group is represented by a main group of players and a few others separated from the rest. The main group consumes an amount of energy between 10,000 kWh and 20,000 kWh in round 10, and there are 5 players separated from the main group that consume an amount of energy below

6,000 kWh. The latter is represented as well in the box plot of this group in which the range goes between 6,000 kWh to 20,000 kWh and three outliers are below 6,000 kWh (round 10).

If we now move to the low efficiency houses group, we could see that there are two different sub-groups instead of the one big group in the high efficiency houses group. There is one sub-group with 10 players with a consumption between 15,000 kWh and 20,000 kWh in round 10 and a second sub-group with 15 players with a consumption below 11,000 kWh. For the latter reason we can see in the box plot of the low efficiency houses group that the range goes from -1,000 kWh to 20,000 kWh, showing a larger standard deviation than in the high efficiency houses group. Analyzing the pre and post-game questionnaire of players, no indications were found of why those two sub-groups differ, not even analyzing their environmental friendliness indicator. It is worth to point out here that if players have a negative total energy consumption, it does not mean that they do not consume energy. Instead, what this really means is that they are producing more energy than the energy they consume. The energy leftovers are sold to the system and it constitutes an extra source of income.

Despite of having identified two different sub-groups inside the low efficiency houses group, hereinafter the low efficiency group will be treated as a whole without separating it in two. If further analyses are carried out separating this group, the sample size of each of the two sub-groups will be too low (10 and 15 players each) compromising the final significance of the results. Figure 5.9 shows the average energy consumption per group including the 25 players in each one. It can be seen that in round 0 (the moment they just received their given house) the low efficiency houses group consumed a total of 21,213 kWh and in round 10 they reduced their energy consumption by 50%, reaching a total consumption equal to 10,629 kWh. In contrast, the high efficiency houses group started round 0 with a consumption equal to 16,348 kWh and in round 10 they reduced their energy consumption by 32%, reaching a total of 11,130 kWh. As a result, despite their higher energy consumption in round 0, the low efficiency houses group lowered their average energy consumption more than the other group, up to the point that in round 10 they consumed, on average, less than the high efficiency houses group.

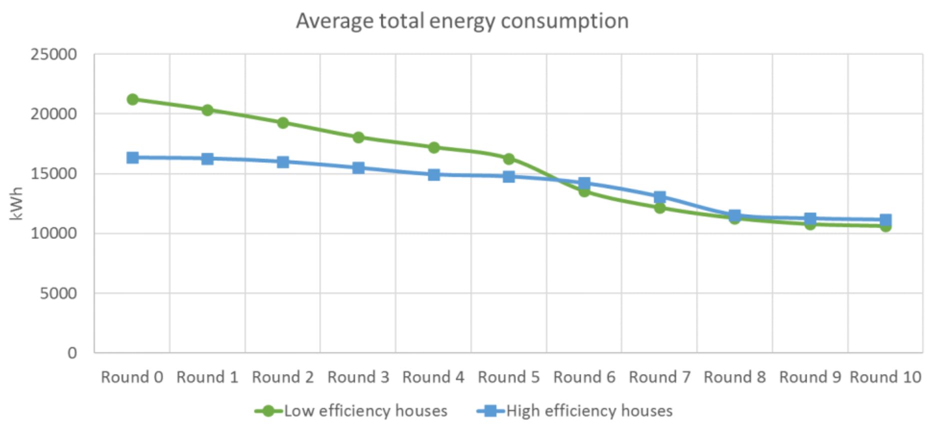


Figure 5.9. Total average energy consumption per group per round.

Despite the previous result in which the difference between the two groups in terms of their average energy consumption diminishes throughout the game, we should statistically check how different those two groups are. This can’t be confirmed by just looking at the averages per round in Figure 5.9 and a measure of spread must be included. The first thing to confirm is to check if the distribution per round per group is normally distributed. In order to check this, the non-parametric Kolmogorov-Smirnov test must be performed. After performing the test in SPSS we can see that just in round 1 both groups are normally distributed, while in all the other rounds there is just one or none group normally distributed (appendix H.5 shows the outcomes given by SPSS). Therefore, student T-test for differences in averages can't be applied to check if there is any significant difference in the averages between the two groups. Therefore, the non-parametric Wilkoxon Rank-Sum test must be applied instead. In this case the null hypothesis H_0 is "there is no difference between the two groups in terms of the total energy consumption" while the alternative hypothesis H_1 is "there is a difference between the two groups in terms of the total energy consumption". Table 5.1 shows the significance values of the statistical test after performing it in SPSS. Appendix H.6 shows the outcome given by SPSS.

Table 5.1. Significance values of the Wilkoxon Rank-Sum test for differences between the groups in terms of their total energy consumption.

Round	Significance	Result
1	0%	H_0 is rejected, both groups are different
2	0%	H_0 is rejected, both groups are different
3	0%	H_0 is rejected, both groups are different
4	0%	H_0 is rejected, both groups are different
5	0.3%	H_0 is rejected, both groups are different
6	55.4%	H_0 can't be rejected, differences between both groups were not found
7	90%	H_0 can't be rejected, differences between both groups were not found
8	82.3%	H_0 can't be rejected, differences between both groups were not found
9	83.9%	H_0 can't be rejected, differences between both groups were not found
10	80.8%	H_0 can't be rejected, differences between both groups were not found

Analyzing Table 5.1 we see that the significance level from round 1 to round 5 is below 5%, resulting that the null hypothesis is rejected. Therefore, from round 1 to round 5 significant differences between the two groups can be identified in terms of their total energy consumption. In contrast, from round 6 to round 10 the significance level of the statistical test is greater than 5%, resulting that the null hypothesis can't be rejected. Therefore, we cannot find significant differences between the two groups in terms of their total energy consumption from round 6 to round 10. The previous result could have been expected considering what was shown in Figure 5.9. Before round 6 the averages per group showed some difference, while from round 6 onwards that difference decreased to almost zero.

5.4. Comfort level

As discussed in Chapter 4, one of the features of the NRG game is the possibility to assign a comfort level to the appliances and furniture that players include in their houses. When players buy a new appliance the can select how luxurious the item is, decision that will immediately affect the overall comfort level of the house they live in. Figure 5.10 shows the evolution of the comfort level indicator per player in each of the two groups (upper graphs) and the boxplots per round for the two groups (lower graphs).

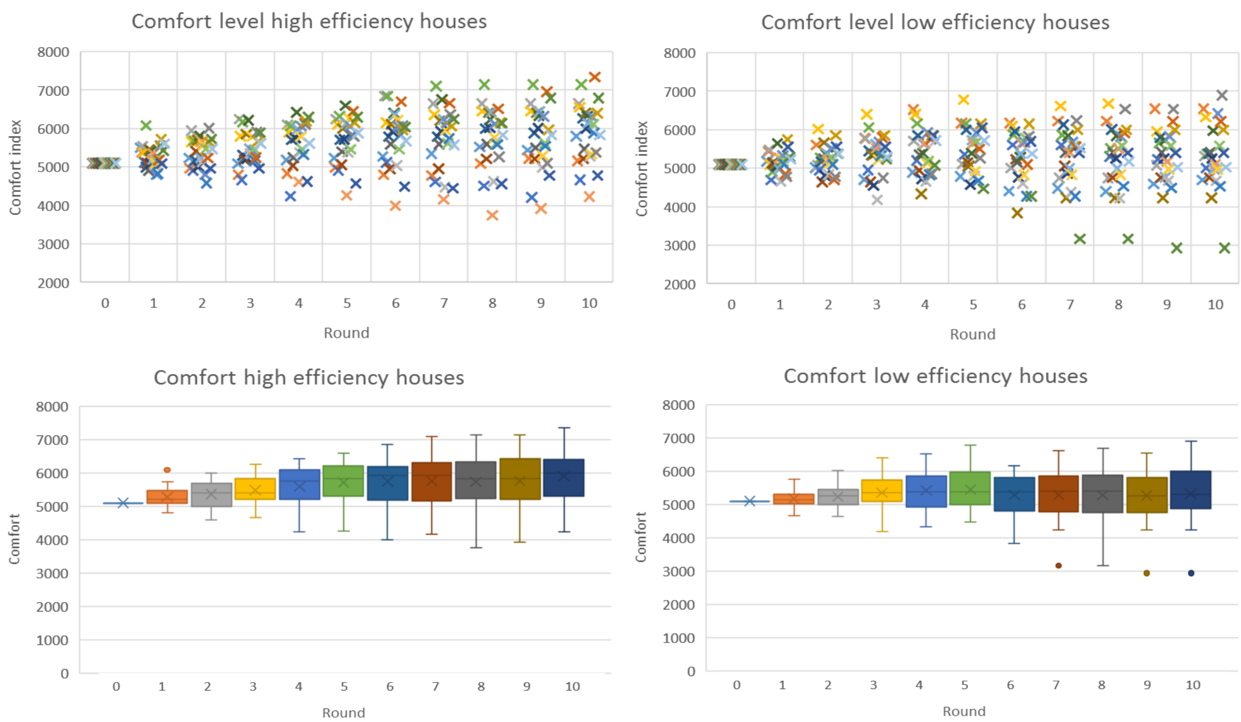


Figure 5.10. Comfort level per round per player.

In this case, we can see that the two groups are composed by large groups of players and we can't identify different sub-groups as we did in the low efficiency houses group related to the total energy consumption. This can be clearly seen in the upper graphs in which one main group is identified and in the boxplots looking at the small range of it.

If the averages of the two groups per round are analyzed, the graph in Figure 5.11 can be made. It shows that both groups started playing the game in round 0 with the same comfort level equal to 5,100 (to isolate the energy efficiency differences between the two groups, all the other factors must be equal, including the comfort level). After finishing the game in round 10, the high efficiency houses group increased, on average, their total comfort level by 16%, reaching a total equal to 5,899. In contrast, the low efficiency houses group increased, on average, their total comfort level by 4%, reaching a total equal to 5,319.

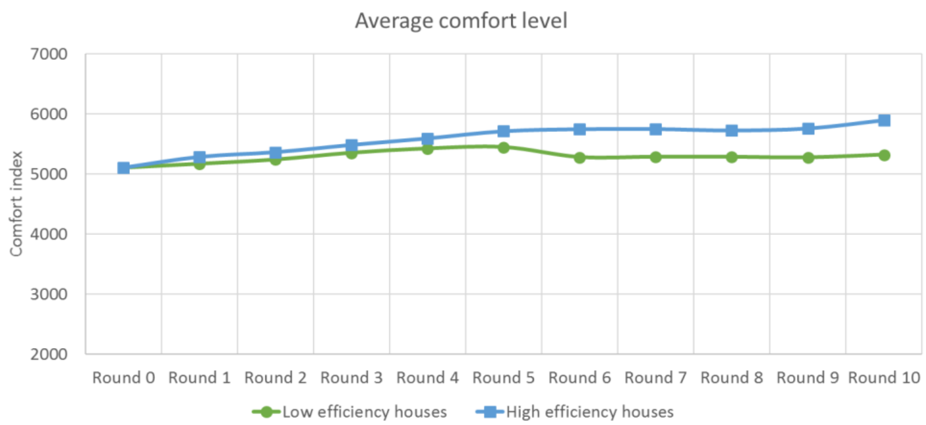


Figure 5.11. Average comfort level per group.

It must be mentioned that the scale used to measure the comfort level doesn't reflect any real pattern. It was chosen that way just to compare different appliances that add more comfort than others. For example, if we compare two different TVs, one with 200 units of comfort and the other one with 100 units of comfort, we can say that the first one gives the user the double comfort than the second one. For instance, the first one may have a greater resolution and a better sound than the other. Therefore, differences in the comfort level just reflect how much the user enjoys using certain appliance or furniture and it is used in this research to compare the luxury level of the appliances included in each house.

As done when analyzing the total energy consumption, to get more significant conclusions we should check how different both groups are in terms of their comfort level. This can't be immediately taken by looking at the averages in Figure 5.11 since they don't consider the spread of the comfort level of each player. Once again, it must be confirmed if the distribution per round per group is normally distributed. The non-parametric Kolmogorov-Smirnov test must be performed. After performing the test in SPSS the null hypotheses can't be rejected, therefore it is assumed that the comfort distribution in each round per group is normally distributed (appendix H.7 shows the outcomes given by SPSS). As a result, if we want to apply the student T-test for differences in averages to check if there are any significant difference in the comfort level between the two groups we must check first the other two conditions. First, the two samples must be randomly drawn from the population. This condition is difficult to check since players were not chosen to participate in the research in a random way, they expressed themselves their willingness to be part of it. Second, both sample groups must be independent. This condition is indeed satisfied because of the way the experiments were carried out, totally isolating both groups to one another in terms of the houses they played with in the game. As a result, although the players were not randomly chosen, we will assume that we can proceed with the student t-test for differences in averages. In this case the null hypothesis H_0 is "there is no difference between the two groups in terms of the comfort level" while the alternative hypothesis H_1 is "there is a difference between the two groups in terms of the comfort level". Table 5.2 shows the significance values of the statistical test after performing it in SPSS. Appendix H.8 shows the outcome given by SPSS.

Table 5.2. Significance values of the student T-test for differences between the groups in terms of their comfort level.

Round	Significance	Result
1	16.3%	H_0 can't be rejected, differences between both groups were not found
2	23.6%	H_0 can't be rejected, differences between both groups were not found
3	32.0%	H_0 can't be rejected, differences between both groups were not found
4	29.7%	H_0 can't be rejected, differences between both groups were not found
5	12.1%	H_0 can't be rejected, differences between both groups were not found
6	1.9%	H_0 is rejected, both groups are different
7	4.2%	H_0 is rejected, both groups are different
8	5.4%	H_0 is rejected, both groups are different
9	3.7%	H_0 is rejected, both groups are different
10	1.5%	H_0 is rejected, both groups are different

Analyzing Table 5.2 we see that the significance value from round 1 to round 5 is greater than 5%, resulting that the null hypothesis can't be rejected. Therefore, we cannot find significant differences between the two groups in terms of their comfort level from round

1 to round 5. In contrast, from round 6 to round 10 the significance value of the statistical test is below 5%, resulting that the null hypothesis is rejected. Therefore, from round 6 to round 10 there are significant differences between the two groups in terms of the comfort level.

5.5. Players' decisions

Having analyzed the two main indicators in the game, the energy consumption and the comfort level, it is interesting to dig deeper into what players did to behave in the way they did. In order to do that, we have broken down the total energy consumption per group into the five different decisions players took that were described in section 4.1. In breaking it down we would be able to see what decisions influenced the total energy consumption the most. As a reminder, those five decisions are: getting rid of initial appliances, improving efficiency of initial appliances, including new electric appliances, including energy management devices and including energy production devices.

The automatic outcomes of the NRG game include just the energy consumption and the comfort level, among others. However, if we want to break the decision down the original outcomes are not enough. To overcome this problem, the programing language Visual Basic in Microsoft Excel had to be used to reconstruct all the decisions players took step by step in order to isolate them to analyze it properly. The following shows the final results after isolating each of the five decisions previously mentioned.

5.5.1. Getting rid of initial appliances

The first decision we have identified that players took to manage their energy consumption is to get rid of the initial appliances they obtained by default at the beginning of the game (appendix A shows the initial appliances players obtained by default in their houses). For example, many players got rid of their TV and replaced it by buying a laptop. The action of selling their TV is included in this decision. Figure 5.12 shows the average reduction of the total energy consumption per group.

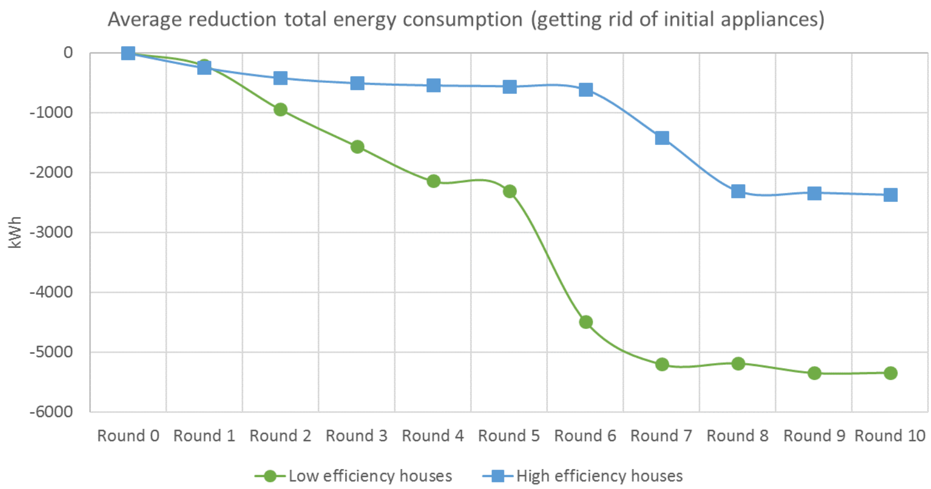


Figure 5.12. Reduction of the total energy consumption by getting rid of the intial appliances.

The previous result shows that after finishing the game in round 10, the low efficiency houses group reduced their total energy consumption, on average, in 5,242 kWh by getting rid of the initial appliances, while the high efficiency houses group reduced their total energy consumption in 2,373 kWh. Despite the difference between the two groups in terms of the average reduction of the total energy consumption, it must be checked if the two groups are indeed different or not, considering the spread between the players. In order to so, it must be first checked if the distribution per round per group is normally distributed. As previously done, the non-parametric Kolmogorov-Smirnov test must be performed. After performing the test in SPSS, just in round 6 in the high efficiency houses group we can assume a normal distribution (appendix H.9 shows the outcomes given by SPSS). Therefore, the non-parametric Wilkoxon Rank-Sum test must be applied instead of the student T-test. In this case the null hypothesis H_0 is "there is no difference between the two groups in terms of the average reduction of the total energy consumption by getting rid of the initial appliances" while the alternative hypothesis H_1 is "there is a difference between the two groups in terms of the average reduction of the total energy consumption by getting rid of the initial appliances". Table 5.3 shows the significance values of the statistical test after performing it in SPSS. Appendix H.10 shows the outcome given by SPSS.

Table 5.3. Significance values of the Wilkoxon Rank-Sum test for differences between the groups in terms of the average reduction of the total energy consumption by getting rid of the initial appliances.

Round	Significance	Result
1	76%	H_0 can't be rejected, differences between both groups were not found
2	56.7%	H_0 can't be rejected, differences between both groups were not found
3	37.2%	H_0 can't be rejected, differences between both groups were not found
4	39.3%	H_0 can't be rejected, differences between both groups were not found
5	15%	H_0 can't be rejected, differences between both groups were not found
6	1.9%	H_0 is rejected, both groups are different
7	0.3%	H_0 is rejected, both groups are different
8	2%	H_0 is rejected, both groups are different
9	1.3%	H_0 is rejected, both groups are different
10	1.3%	H_0 is rejected, both groups are different

Analyzing Table 5.3 we see that the significance value from round 1 to round 5 is greater than 5%, resulting that the null hypothesis can't be rejected. Therefore, we cannot find significant differences between the two groups from round 1 to round 5. In contrast, from round 6 to round 10 the significance value of the statistical test is below 5%, resulting that the null hypothesis is rejected. Therefore, there are significant differences between the two groups from round 6 to round 10. Consequently, both groups show a different behavior from round 6 in terms of the decision to get rid of initial appliances.

5.5.2. Improving efficiency of initial appliances

The second decision we have identified that players took to manage their energy consumption is to improve the efficiency of the initial appliances they obtained by default at the beginning of the game (appendix A shows the initial appliances players got by default in their houses). For example, many players replaced their old central heating systems for less consuming ones. The action of replacing their central heating systems for new ones with a different efficiency is included in this decision. Figure 5.13 shows the average reduction of the total energy consumption per group.

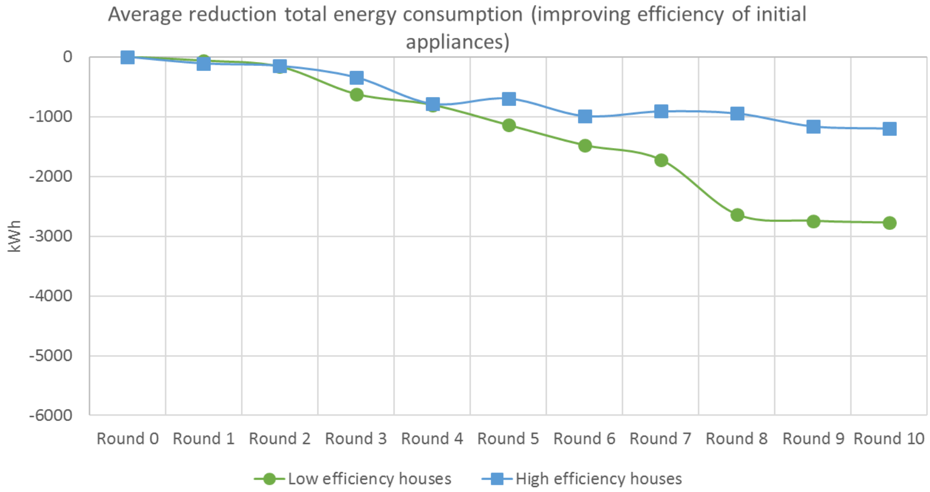


Figure 5.13. Reduction of the total energy consumption by improving efficiency of intial appliances.

The previous result shows that after finishing the game in round 10, the low efficiency houses group reduced their total energy consumption, on average, in 2,773 kWh by improving the efficiency of initial appliances, while the high efficiency houses group reduced their total energy consumption in 1,195 kWh. Despite the difference between the two groups in terms of the average reduction of the total energy consumption, it must be checked if the two groups are indeed different or not considering the spread between the players. In order to so, it must be first checked if the distribution per round per group is normally distributed. As previously done, the non-parametric Kolmogorov-Smirnov test must be performed. After performing the test in SPSS we see that all the distributions do not follow the normal theoretical distribution (appendix H.11 shows the outcomes given by SPSS). Therefore, the non-parametric Wilkoxon Rank-Sum test must be applied instead of the student T-test. In this case the null hypothesis H_0 is "there is no difference between the two groups in terms of the average reduction of the total energy consumption by improving the efficiency of the initial appliances" while the alternative hypothesis H_1 is "there is a difference between the two groups in terms of the average reduction of the total energy consumption by improving the efficiency of the initial appliances". Table 5.4 shows the significance values of the statistical test after performing it in SPSS. Appendix H.12 shows the outcome given by SPSS.

Table 5.4. Significance values of the Wilkoxon Rank-Sum test for differences between the groups in terms of the average reduction of the total energy consumption by improving the efficiency of the initial appliances.

Round	Significance	Result
1	0.2%	H_0 is rejected, both groups are different
2	0.8%	H_0 is rejected, both groups are different
3	0.5%	H_0 is rejected, both groups are different
4	3.2%	H_0 is rejected, both groups are different
5	5.6%	H_0 can't be rejected, differences between both groups were not found
6	80.8%	H_0 can't be rejected, differences between both groups were not found

Round	Significance	Result
7	21.8%	H_0 can't be rejected, differences between both groups were not found
8	2.4%	H_0 is rejected, both groups are different
9	3.5%	H_0 is rejected, both groups are different
10	5.1%	H_0 can't be rejected, differences between both groups were not found

Analyzing Table 5.4 we see a different result to what happened with the previous decision. In rounds 1, 2, 3, 4, 8 and 9 the significance value is below 5%, resulting that the null hypothesis is rejected. As a result, in these rounds the two groups show significant differences. In contrast, in rounds 5, 6, 7 and 10 the significance value is greater than 5%, resulting that the null hypothesis can't be rejected. Therefore, in these rounds the two groups do not show any significant difference in terms of their decision to improve the efficiency of the initial appliances. This result can only be explained by the fact that big standards deviations are shown in the analyzed distributions, resulting that sometimes groups show significant differences, while in some other times that difference is not present anymore. In order to make this clearer, we have also included in Table 5.5 the mean and standard deviation of each group per round.

Table 5.5. Descriptive statistics of the decision to improve efficiency of the initial appliances (values in kWh).

Round	Low efficiency houses		High efficiency houses	
	Mean	**Std. deviation**	**Mean**	**Std. deviation**
1	-64	683	-111	379
2	-165	810	-151	444
3	-622	2,162	-342	1,171
4	-803	2,152	-783	1,890
5	-1,142	2,736	-695	1,647
6	-1,478	3,460	-986	1,949
7	-1,720	3,508	-910	1,959
8	-2,635	3,793	-943	1,970
9	-2,742	3,762	-1,159	1,978
10	-2,773	3,769	-1,195	1,986

Table 5.5 shows that despite of the clear differences in the mean of each group in rounds 5, 6, 7 and 10, the big standard deviations produce that the result of the statistical test indicates that no significant differences between the groups can be found. Therefore, we will assume that the previous result is not conclusive to clearly state that the two groups show significant differences in terms of their behavior in improving efficiency of initial appliances.

5.5.3. Including new electric appliances

The third decision we have identified that players took to manage their energy consumption is to include new electric appliances in their houses. For example, many players bought laptops or dishwashers that were not included by default in their houses. The action of buying new electric appliances is included in this decision. Figure 5.14 shows the average increase of the total energy consumption per group.

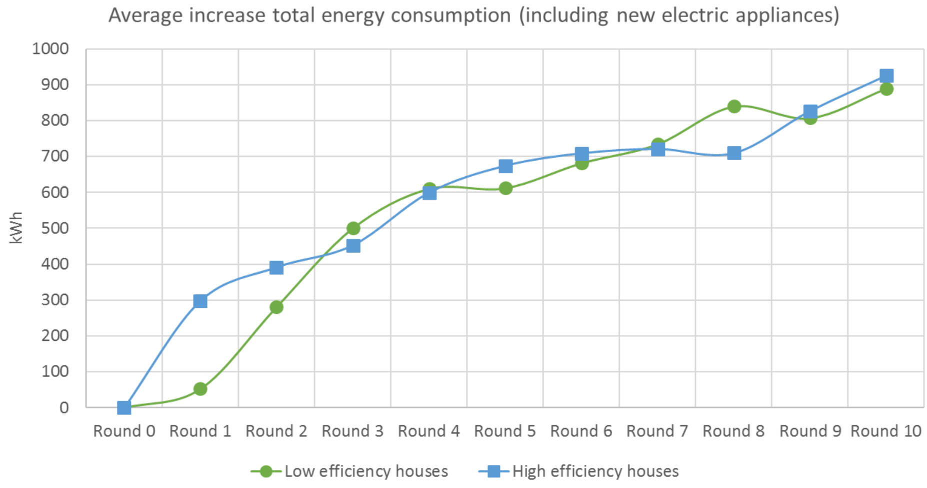


Figure 5.14. Increase of the total energy consumption by including new appliances.

The previous result shows that after finishing the game in round 10, the low efficiency houses group increased their total energy consumption in 889 kWh by including or buying new electric appliances, while the high efficiency houses group increased their total energy consumption in 925 kWh. Despite the small difference between the two groups in terms of the average increment of the total energy consumption, it must be checked if the two groups are different or not considering the spread between the players. Once again, it must be first checked if the distribution per round per group is normally distributed. As previously done, the non-parametric Kolmogorov-Smirnov test must be performed. After performing the test in SPSS we don't find any round of the game in which both groups show a normal distribution (appendix H.13 shows the outcomes given by SPSS). Therefore, the non-parametric Wilkoxon Rank-Sum test must be applied of the student T-test. In this case the null hypothesis H_0 is "there is no difference between the two groups in terms of the average increment of the total energy consumption by including new appliances" while the alternative hypothesis H_1 is "there is a difference between the two groups in terms of the average increment of the total energy consumption by including new appliances". Table 5.6 shows the significance values of the statistical test after performing it in SPSS. Appendix H.14 shows the outcome given by SPSS.

Table 5.6. Significance values of the Wilkoxon Rank-Sum test for differences between the groups in terms of the average increment of the total energy consumption by including new appliances.

Round	Significance	Result
1	4.2%	H_0 is rejected, both groups are different
2	14.9%	H_0 can't be rejected, differences between both groups were not found
3	71.8%	H_0 can't be rejected, differences between both groups were not found
4	90.7%	H_0 can't be rejected, differences between both groups were not found
5	54.1%	H_0 can't be rejected, differences between both groups were not found
6	49.6%	H_0 can't be rejected, differences between both groups were not found
7	64.8%	H_0 can't be rejected, differences between both groups were not found
8	71.2%	H_0 can't be rejected, differences between both groups were not found
9	86.1%	H_0 can't be rejected, differences between both groups were not found
10	69.1%	H_0 can't be rejected, differences between both groups were not found

Analyzing Table 5.6 we see that the significance value in round 1 is below 5%, resulting that the null hypothesis is rejected. Therefore, in round 1 both groups show a significant difference in terms of the increment of the total energy consumption by including new electric appliances. In contrast, from round 2 to round 10 the significance value of the statistical test is greater than 5%, resulting that the null hypothesis can't be rejected. Therefore, no significant differences between the two groups can be found from round 2 to round 10.

5.5.4. Including energy management devices

The fourth decision we have identified that players took to manage their energy consumption is to include energy management devices in their houses. Energy management devices are any device that manages to reduce the overall energy consumption of a house by a certain percentage. In the NRG game, energy management devices are smart meters, movement sensors or stand-by-killers, if we talk about electricity reduction, or insulation (wall, floor, windows, and so forth) if we talk about gas reduction. The action of buying new energy management devices is included in this decision. Figure 5.15 shows the average reduction of the total energy consumption per group.

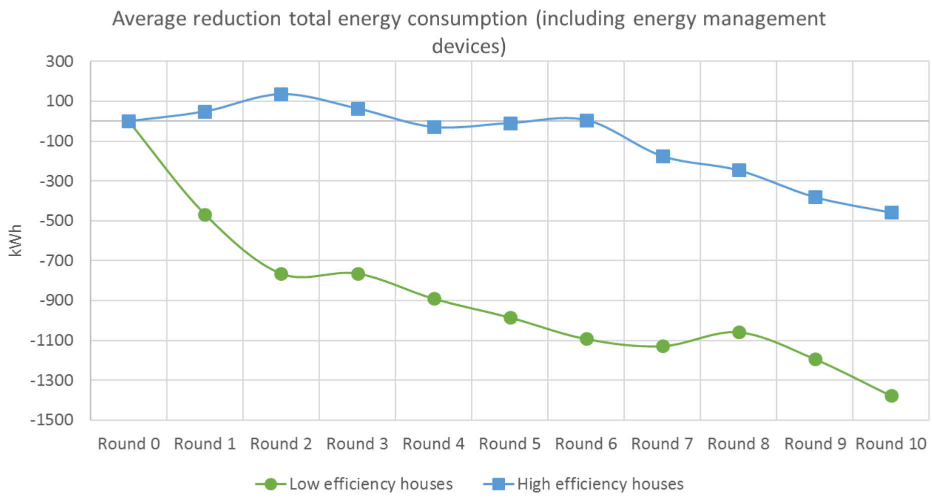


Figure 5.15. Reduction of the total energy consumption by including energy manageent devices.

The previous result shows that after finishing the game in round 10, the low efficiency houses group reduced, on average, their total energy consumption in 1,377 kWh by including energy management devices, while the high efficiency houses group reduced their total energy consumption in 458 kWh. Despite the difference between the two groups in terms of the average reduction of the total energy consumption, it must be checked if the two groups are different or not considering the spread between the players. It must be first checked if the distribution per round per group is normally distributed by applying the non-parametric Kolmogorov-Smirnov statistical test. After performing the test in SPSS we don't find any round of the game in which both groups show a normal distribution (appendix H.15 shows the outcomes given by SPSS). Therefore, the non-parametric Wilkoxon Rank-Sum test must be applied of the student T-test. In this case the null hypothesis H_0 is "there is no difference between the two groups in terms of the reduction of the total energy consumption by including energy management devices" while the alternative hypothesis H_1 is "there is a difference between the two groups in terms of the reduction of the total energy consumption by including energy management devices". Table 5.7 shows the significance values of the statistical test after performing it in SPSS. Appendix H.16 shows the outcome given by SPSS.

Table 5.7. Significance values of the Wilkoxon Rank-Sum test for differences between the groups in terms of the average reduction of the total energy consumption by including energy management devices.

Round	Significance	Result
1	4.2%	H_0 is rejected, both groups are different
2	0%	H_0 is rejected, both groups are different
3	0.2%	H_0 is rejected, both groups are different
4	0.2%	H_0 is rejected, both groups are different
5	0.1%	H_0 is rejected, both groups are different
6	0.1%	H_0 is rejected, both groups are different
7	0.2%	H_0 is rejected, both groups are different
8	0.5%	H_0 is rejected, both groups are different
9	1.3%	H_0 is rejected, both groups are different
10	1.9%	H_0 is rejected, both groups are different

Analyzing Table 5.7 we see that the significance values in all the rounds are below 5%, resulting that all the null hypotheses are rejected. Therefore, both groups show a significant difference in terms of the reduction of the total energy consumption by including energy management devices. As a result, we can assume that the low efficiency houses group reduced their energy consumption by including energy management devices more than the high efficiency houses group.

5.5.5. Including energy production devices

The fifth and final decision we have identified that players took to manage their energy consumption is to include energy production devices in their houses. Energy production devices are any device that produces usable energy. In the NRG game, energy production devices are solar panels, mini wind turbines, micro cogeneration units, and so forth. The action of buying new energy production devices is included in this decision. Figure 5.16 shows the average reduction of the total energy consumption per group.

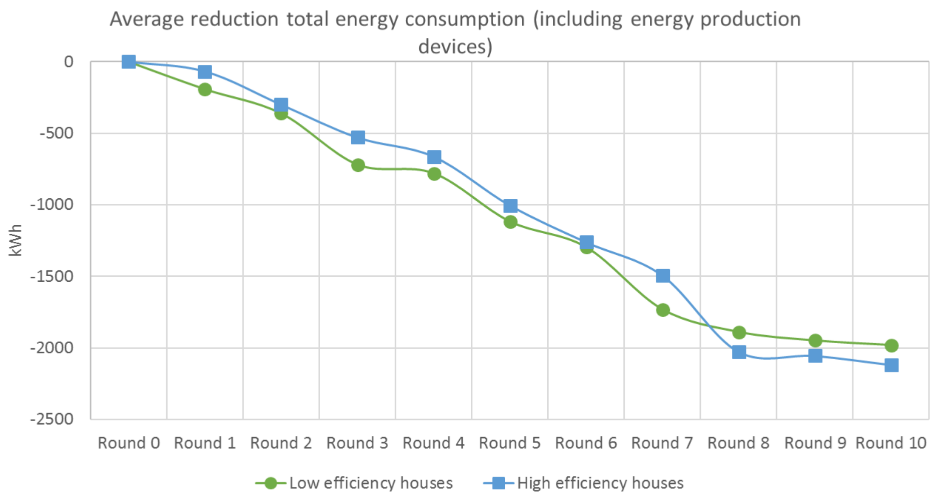


Figure 5.16. Reduction of the total energy consumption by including energy production devices.

The previous result shows that after finishing the game in round 10, the low efficiency houses group reduced, on average, their total energy consumption in 1,980 kWh by including energy production devices, while the high efficiency houses group reduced their total energy consumption in 2,117 kWh. Despite difference between the two groups in terms of the average reduction of the total energy consumption, it must be checked if the two groups are different or not considering the spread between the players. It must be first checked if the distribution per round per group is normally distributed by applying the non-parametric Kolmogorov-Smirnov statistical test. After performing the test in SPSS, just in round 8 and round 10 we can assume a normal distribution in both groups (appendix H.17 shows the outcomes given by SPSS). Therefore, just in round 8 and round 10 we may be able to use the student T-test for differences in averages to check significant differences between the two groups (if the rest of the conditions are satisfied). However, for simplicity, we will proceed with the non-parametric Wilkoxon Rank-Sum test due to the fact that in the majority of the rounds we do not have both groups normally distributed. In this case the null hypothesis H_0 is "there is no difference between the two groups in terms of the reduction of the total energy consumption by including energy production devices" while the alternative hypothesis H_1 is "there is a difference between the two groups in terms of the reduction of the total energy consumption by including energy production devices". Table 5.8 shows the significance values of the statistical test after performing it in SPSS. Appendix H.18 shows the outcome given by SPSS.

Table 5.8. Significance values of the Wilkoxon Rank-Sum test for differences between the groups in terms of the reduction of the total energy consumption by including energy prodction devices.

Round	Significance	Result
1	56.4%	H_0 can't be rejected, differences between both groups were not found
2	78.6%	H_0 can't be rejected, differences between both groups were not found
3	85.6%	H_0 can't be rejected, differences between both groups were not found
4	75.1%	H_0 can't be rejected, differences between both groups were not found
5	92.8%	H_0 can't be rejected, differences between both groups were not found
6	89%	H_0 can't be rejected, differences between both groups were not found
7	41.8%	H_0 can't be rejected, differences between both groups were not found
8	74.1%	H_0 can't be rejected, differences between both groups were not found
9	85.3%	H_0 can't be rejected, differences between both groups were not found
10	79.3%	H_0 can't be rejected, differences between both groups were not found

Analyzing Table 5.8 we see that the significance values in all the rounds are greater than 5%, resulting that the null hypothesis can't be rejected. Therefore, both groups do not show any significant difference in terms of the reduction of the total energy consumption by including energy production devices.

5.5.6. Effects of the decisions on the total energy consumption

Having broken down the decisions players took to manage their energy consumption in the game, we can show them all together to visualize what decisions had more influence over the total energy consumption. Figure 5.17 shows the total energy consumption per group including all the five players' decision we have already analyzed, and it is read as follows. First, line *0) doing nothing* reflects what would have happened if the players had done nothing in their default houses, i.e. the initial energy consumption remains the same throughout the game. Second, line *1) getting rid of initial appliances* reflects the new total energy consumption if just the effect of this action is included into the energy consumption. Third, line *2) improving efficiency of initial appliances* shows the new total energy consumption if this action is added to the previous action reflected by line 1) in a cumulative way. Fourth, line 3*) including new electric appliances* shows the total energy consumption if this action is added to the previous action reflected by line 2). Lines *4) including energy management devices* and *5) including energy production devices* follows the same logic. As a result, line 5) constitutes the final total energy consumption per group including all the five players' decisions.

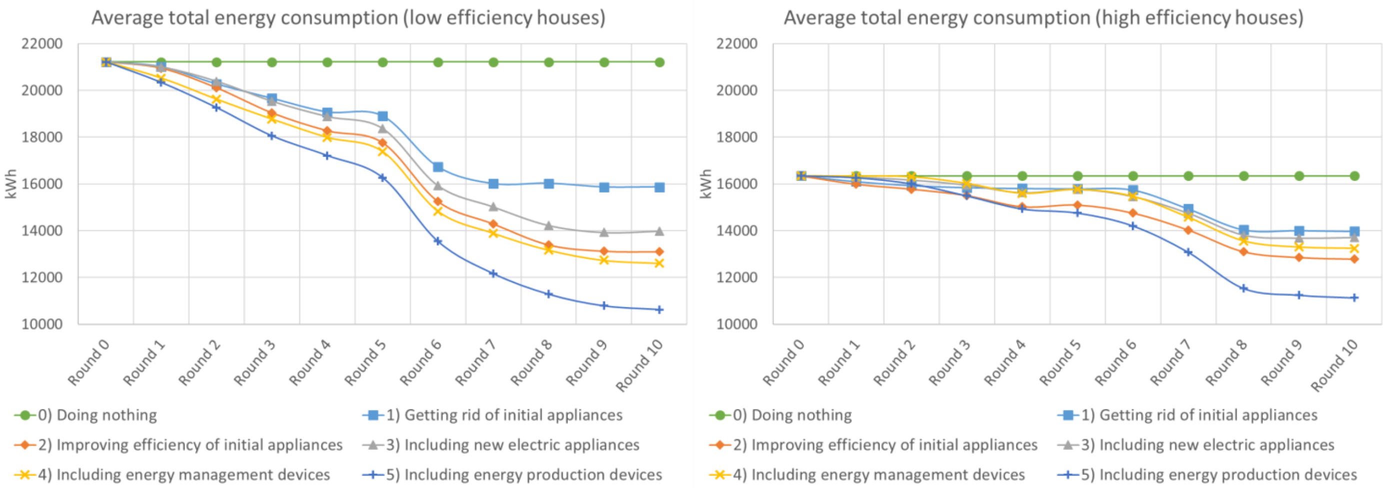


Figure 5.17. Global effects of players' decisions on the total energy consumption.

From Figure 5.17 we can compare the effects of each action over the total average energy consumption per group. The total average reduction of the energy consumption is higher in the low efficiency houses group. In addition, we can also see inside each group that a part of the energy savings from actions *1) getting rid of initial appliances* and *2) improving efficiency of initial appliances* is taken back by action *3) including new electric appliances*. This may be a first sign of the rebound effect inside each group.

5.6. Chapter's conclusions

This chapter starts analyzing if the two groups of people that played the NRG game are comparable, a critical factor if comparisons between them are going to be analyzed. After performing statistical tests, the two groups of people did not show any significant difference in four criteria: environmental friendliness, abilities to play video game, how often participants play video games and whether they pay money in real life for their houses. Those four factors were defined to be crucial for the next comparability of the two groups in order to assess the rebound effect. Since the two groups did not show any significant difference they can be compared in terms of their energy consumption, comfort level and the general decisions they took in the game.

The two most important indicators in the NRG game were analyzed as well: energy consumption and comfort level. First, the energy consumption evolution between the two groups showed significant differences at the beginning of the game. However, from round 6 on, the energy consumption did not show any significant different. Second, the comfort level did not show significant differences between the two groups from round 1 to round 5. However, from round 6 on, the comfort level showed significant differences between both groups. The previous results and their interpretation regarding the rebound effect will be analyzed in the following chapter.

In addition, the energy consumption of each group was broken down into five different decisions players took. Analyzing statistical differences of the two groups in terms of their decisions throughout the game, we found significant differences in *getting rid of initial appliances* and *including energy management devices*. In contrast, no significant differences were detected in *improving energy efficiency of initial appliances*, *including new electric appliance*s and *including energy production devices*. The interpretation of the previous results regarding the rebound effect will be analyzed in the following chapter.

In spite of having analyzed how players behaved in the game, it must be checked later on that the results obtained can be applied into a real world context. In other words, we cannot immediately state that players will behave in real life in the same way they did in the game. The way this will checked is by analyzing the post-game questionnaires that players answered. In chapter 7 some addition analysis will be carried out in order to check the real applicability of the results into a real life context.

Having described and analyzed in this chapter the most important behaviors of players while playing the NRG game, we will continue with the next chapter about the rebound effect. It will be analyzed if signs of the rebound effect are detected and, if so, a calculation of the effect will be carried out.

6. Rebound effect assessment

With the information from the total energy consumption we are now able to analyze the rebound effect comparing the behavior of the two groups throughout the game. This chapter is organized as follows. First, some initial signs of the rebound effect will be described considering the previous general results of the experiments. Second, a calculation of the rebound effect will be carried out trying to discover its size.

6.1. First signs of the rebound effect

With all the analysis previously done we are able to get some initial conclusions about the presence of the rebound effect. Two signs of the rebound effect will be analyzed first: looking at the total energy consumption and looking at the comfort level.

First, if we take a look at the total energy consumption (chapter 5.3) we could see that the initial difference between the two groups diminishes over time, becoming (on average) negative. In other words, the average consumption of the low efficiency houses group is less than the average consumption of the high efficiency houses group in round 10, despite the fact that in round 0 the situation was the other way around. In addition, the statistical tests to check significant differences between the two groups showed that from round 1 to round 5 both groups were significantly different, and from round 6 until round 10 this difference was not significant anymore. As a result, we can confirm that the initial difference of the two groups disappears over time, showing that the low efficiency houses group managed to reduce their energy consumption more than the high efficiency houses group. Since we have shown that the two groups are comparable, the latter can be explained by the rebound effect in which the initial energy savings (difference between the two groups in round 0) diminishes.

Second, we can also analyze the presence of the rebound effect by looking at the comfort level in the two groups (chapter 5.4). We have seen that the comfort level of the two groups don't show any significant difference from round 1 to round 5, situation that changes from round 6 to round 10 in which the high efficiency houses group show a significantly higher comfort level than the low efficiency houses group. The latter is explained by the fact that the high efficiency houses group had more money to spend throughout the game due to their lower initial consumption (and therefore, lower bills to pay after each round). As a result, this group had the opportunity to buy more appliances that increased their overall comfort level than the other group. Although both groups reduced their energy consumption, having more money available produced that the high efficiency houses group reduced their consumption less than the low efficiency houses group due to a higher expenditure in appliances that increased the comfort rather than expenditure in appliances to reduce consumption.

Consequently, we can assume that the rebound effect significantly affected the behavior of the two groups. Having a low efficiency produced players to reduce their energy consumption more than the other group. Similarly, having a high efficiency produced players to increase their comfort level more than the other group.

6.2. Calculation

Having shown the presence of the rebound effect we can now make some calculations with the data collected from the players.

6.2.1. Control group and base case scenario

In order to make calculations of the rebound effect using the two groups of players, we must first define the control group and the base case scenario. First, the control group is defined as the low efficiency houses group, their behavior is the one that will be used to calculate the expected savings and the actual savings, according to equation (1). As a reminder, equation (1) was defined in section 2.2 as follows:

$$Rebound\ effect = \frac{expected\ savings - actual\ savings}{expected\ savings} \times 100 \tag{1}$$

Second, the base case scenario will be defined as the behavior of the high efficiency houses group if the rebound effect was zero. In other words, the base case scenario can be interpreted as the lowest energy consumption the high efficiency houses group could have had (the situation in which no energy savings are lost due to the rebound effect. Therefore, it can also be called the best case scenario). This base case scenario will also be calculated using the defined control group.

Two different ways to calculate the base case scenario will be used: methodology A and methodology B. In the following both methodologies will be explained:

6.2.1.1. *Methodology A*

As previously mentioned, the base case scenario is the behavior that the high efficiency houses group would have to follow having a rebound effect equal to zero. If the real evolution of the energy consumption of the high efficiency houses group differs to that base case scenario, the rebound effect would be detected.

The first way we have proposed to calculate that base case scenario is by looking at the control group (low efficiency houses) and their behavior over rounds. We propose that the base case scenario is equal to the behavior of the low efficiency houses group minus the initial energy consumption difference between the two houses in round 0. In other words, the high efficiency houses group would have had to keep the same difference in energy consumption in round 0 throughout the whole game in order to have a rebound effect equal to zero. As a result, the base case scenario is a curve parallel to the low efficiency houses group energy consumption keeping the initial difference. We can see on the left graph of Figure 6.1 the control group (low efficiency houses group) and the base case scenario or the behavior of the high efficiency houses group if the rebound effect was zero. Another way to understand the base case scenario is by labeling it as the best the high efficiency houses group could have done to reduce their energy consumption (or the best case scenario). In addition, on the right graph of Figure 6.1 we can also see the real evolution of the high efficiency houses group, together with the low efficiency houses group and the base case scenario. It can be seen that the real behavior of the high efficiency houses group differs from the base case scenario. The latter means that the high efficiency houses group could have taken more actions to reduce their energy consumption, showing that the rebound effect was indeed present.

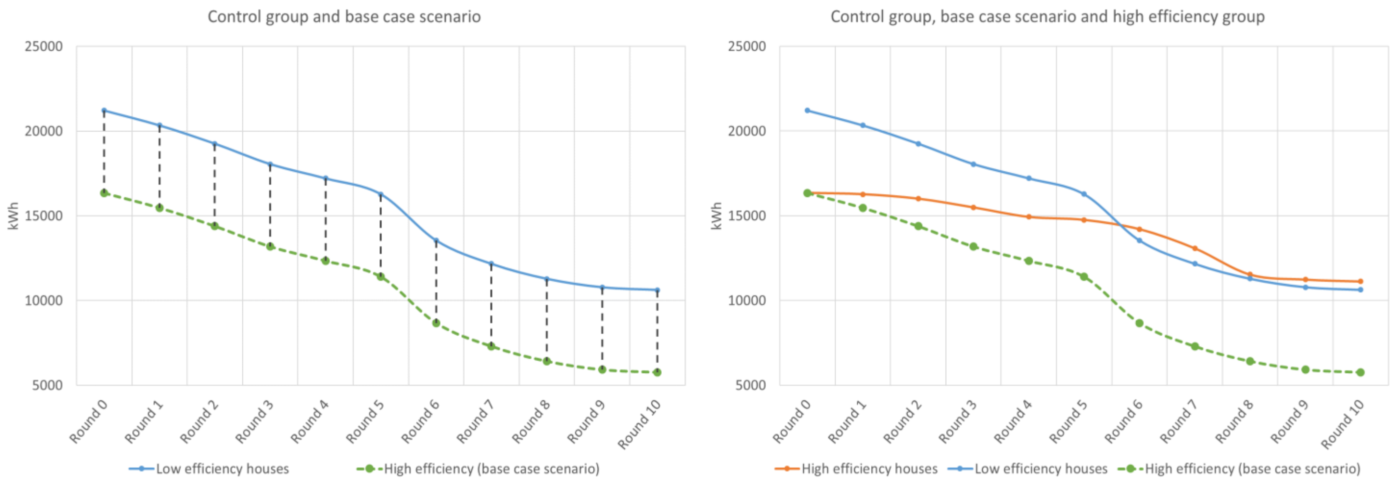


Figure 6.1. High and low efficiency houses group and base case scenario (methodology A).

6.2.1.2. <u>*Methodology B*</u>

Having calculated the base case scenario of the high efficiency houses group using methodology A, we can even go further and improve it. What methodology A has not considered is the fact that the high efficiency houses group has less possibilities to reduce their total energy consumption, due to the fact that they are already high efficient. The latter means that calculating the base case scenario by keeping the same initial difference between the two groups is too restrictive. We must acknowledge the fact that the room for improvements and energy reduction possibilities of the high efficiency houses group are less than the low efficiency houses group. Imagine that person A has a light bulb in which the energy efficiency is 70% while person B has another light bulb with 40% of energy efficiency. Since person A is already more efficient than person B, we can't say that both persons could increase their energy efficiency by 60%. Instead, we must acknowledge that person A has a 30% more to go to be 100% efficient, while person B has 60%. Methodology A assumes that both persons can improve their energy efficiency in the same number by forcing that the base case scenario must keep the same initial difference between the two groups. In order to take this into account, we have also designed methodology B in which the base case scenario calculation considers the difference in the room for improvements between both groups.

The methodology B to calculate the base case scenario goes as follows. First, the difference per round in the room for improvements between the two groups must be calculated (R_i). Equation (2) shows this calculation. It reflects how much possibilities the low efficiency houses group have to reduce their energy consumption over the high efficiency houses group. If R_i is equal, for instance, to 0.5 then the low efficiency houses group has 50% more possibilities to reduce their energy consumption than the high efficiency houses group. In the formula CL_i reflects the actual consumption of the low efficiency houses group in round i, CH_i reflects the actual consumption of the high efficiency houses group in round i and min stands for the minimum possible consumption. Including a minimum consumption acknowledges that both groups have a lowest limit to reduce their energy consumption, i.e. they can't reduce their energy consumption further than that.

$$R_i = \frac{(CL_i - min)}{(CH_i - min)} - 1 \tag{2}$$

Having calculated R_i to reflect how much possibilities the low efficiency houses group has over the high efficiency houses group, we proceed to calculate the real reduction of the energy consumption of the low efficiency houses group per round (RL_i). It is calculated using the actual consumption of the low efficiency houses group in two consecutive rounds. Equation (3) shows this calculation. In the formula, CL_i reflects the actual energy consumption of the low efficiency houses group in round i. If RL_i in one round is, for instance, -0.2 then the low efficiency houses group reduced their energy consumption by 20% between those two consecutive rounds.

$$RL_i = \frac{CL_i}{CL_{i-1}} - 1 \tag{3}$$

Now, we could say that if the low efficiency houses group managed to reduce their energy consumption in two consecutive rounds by a certain percentage (given by RL_i), then the high efficiency houses group could have done the same. However, since the latter group has less room for improvements or less possibilities to reduce their energy consumption, we can't say that. What we can say is that the high efficiency houses group could reduce their energy consumption by the same percentage as the other group, but reduced by the difference between the room for improvements given by R_i. As a result, we can calculate the reduction of the high efficiency houses group in the base case scenario (RH_i^b) using equation (4). It shows that the actual reduction of the energy consumption of the low efficiency houses group per round (RL_i) is reduced by the factor reflecting differences in the room for improvements (R_i). It must be said to clarify this point that RH_i^b does not reflect the real consumption of the high efficiency houses group, instead it reflects the best they could have done in reducing their energy consumption in round i, or what we have already called the base case scenario.

$$RH_{i+1}^b = RL_{i+1} \cdot (1 - R_i) \tag{4}$$

Finally, having calculated the best reduction the high efficiency houses group could have done, we can calculate the energy consumption of the high efficiency houses group in the base case scenario (CH_i^b). To do so, we have to apply the percentage calculated in equation (4) to the energy consumption of the previous round of the same group in the base case scenario. Equation (5) reflects this relation. The initial condition of the formula is the actual consumption of the high efficiency houses group in round 0 (CH_0^b), reflecting the initial state of the high efficiency house.

$$CH_{i+1}^b = CH_i^b \cdot \left(1 + RH_{i+1}^b\right) \tag{5}$$

As a result, by following the previous reasoning we can calculate the base case scenario, or in other words, the behavior that the high efficiency houses group would have followed if the rebound effect was zero or their lowest possible consumption. As a consequence, if we detect that the real consumption of the high efficiency houses group differs from the base case scenario, then the rebound effect would be present.

In order to make the previous methodology clearer, a numeric example will be given. If we focus on round 3, we can see from the data that the real average energy consumption of the low efficiency houses group is $CL_3 = 18{,}053\ kWh$, and in the case of the high efficiency houses group is $CH_3 = 15{,}491\ kWh$. Therefore, we can see that the low efficiency houses group, due to their higher consumption in this round, has more room for reducing their energy consumption than the high efficiency group (given by the R_i factor). Applying equation (2) we get that $R_3 = 16.5\%$ (assuming that the minimum possible consumption is equal to 0 kWh). The latter means that the low efficiency group has 16.5% more possibilities to reduce their energy consumption than the high efficiency group in round 3. In addition, we can also see from the data that the real average energy consumption of the low efficiency houses group in the previous round is $CL_2 = 19{,}260\ kWh$. Therefore, applying equation (4) we conclude that the actual energy reduction of the low efficiency group between round 3 and round 2 is $RL_3 = 6.3\%$, meaning that the low efficiency group managed to reduce their energy consumption by 6.3% in those two consecutive rounds. As a result, we may assume that since the low efficiency group reduced their energy consumption by 6.3%, the high efficiency group could have done the same. However, what we are assuming in this methodology is that the latter is not true and the high efficiency houses group had a 16.5% less possibilities to reduce their consumption than the low efficiency group. Then, we can calculate the maximum reduction the high efficiency group could have achieved in this specific round by applying equation (4). As a result, we find that this number is $RH_3^b = 5.3\%$ (6.3% is reduced by 16.5%). The latter means that despite the reduction of the low efficiency group being equal to 6.3%, the best the high efficiency group could have done between round 2 and round 3 is to reduce their energy consumption by 5.3%. Finally, since this is an iterative method, to calculate the consumption of the high efficiency group in the base case scenario in round 3 (CH_3^b), we must have previously calculated the same factor for round 2 (CH_2^b). CH_2^b resulted to be equal to 15,243 kWh, and applying equation (5) we get that $CH_3^b = 14{,}435\ kWh$. The latter number is the lowest possible energy consumption of the high efficiency group in round 3. Since the real consumption was $CH_3 = 15{,}491\ kWh$, we can see that the high efficiency group could have reduced their energy consumption in 1,056 kWh extra (difference between CH_3 and CH_3^b).

Having explained how methodology B works, we can calculate the base case scenario in all the rounds. In Figure 6.2 we can see on the left the control group (low efficiency houses group) and the base case scenario or the behavior of the high efficiency houses group if the rebound effect was zero. In addition, in the right side of the Figure 6.2 we can also see the real evolution of the high efficiency houses group, altogether with the low efficiency houses group and the base case scenario. It can be seen that the real behavior of the high efficiency houses group differs from the base case scenario. The latter means that the high efficiency houses group could have taken more actions to reduce their energy consumption, showing that the rebound effect was indeed present.

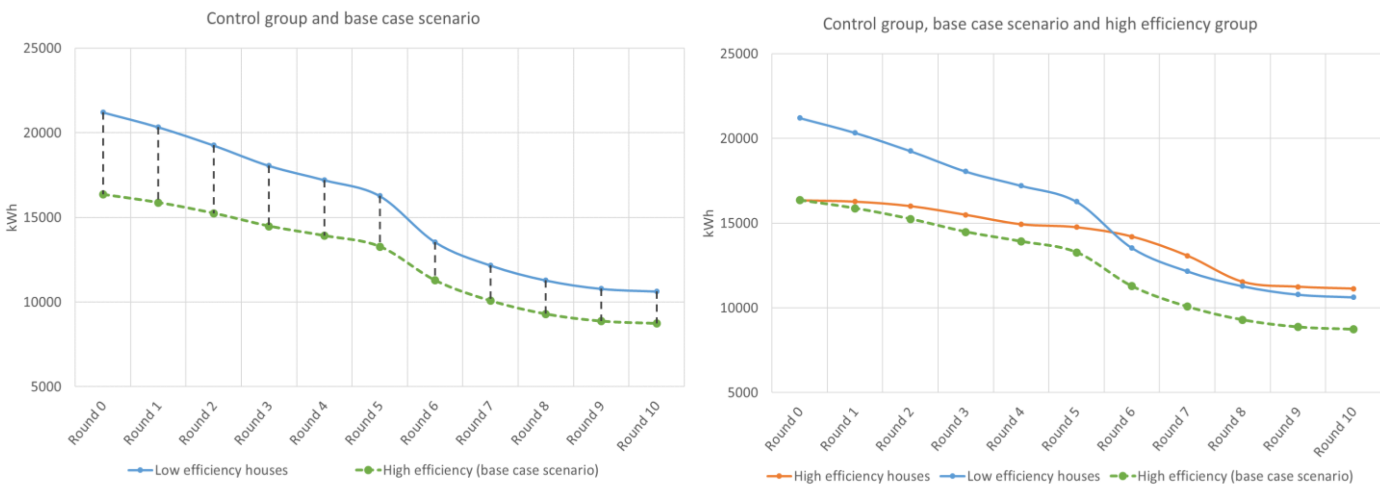


Figure 6.2. High and low efficiency houses group and base case scenario (methodology B).

It must be said in this calculation of the base case scenario under methodology B, that the minimum possible consumption in equation (2) was taken to be equal to 0 kWh. This means that players could have taken their energy consumption all the way down to zero (consumption is equal to production). This result was shown to not be too ambitious for players, since 10% of the players (5 in total, 3 from the high efficiency houses group and 2 from the low efficiency houses group) managed to reduce their energy consumption below zero, producing more energy than they needed. It must be also said that a quick sensitivity analysis was carried out to check the influence of this selection. It was discovered that no major differences in the base case scenario line are obtained if the minimum possible consumption was the absolute minimum consumption of all the 50 players. Therefore, for simplicity, the minimum possible energy consumption was chosen to be equal to zero.

If we now compare Figure 6.1 and Figure 6.2, we can see that the base case scenario in methodology A keeps the same difference with the low efficiency houses group throughout the game, while in methodology B the mentioned difference diminishes over time. The latter result shows the inclusion in methodology B the difference in the room for improvements between the two groups.

6.2.2. Rebound effect

If we now want to calculate the rebound effect we can apply equation (1). In order to do so, we must first identify what stands for the expected savings and the actual savings. First, the expected savings are what we would expect the difference between the two groups to be if the rebound effect was zero. In other words, the expected savings are the gap between the real low efficiency houses group energy consumption and the base case scenario of the high efficiency houses group. In Figure 6.1 and Figure 6.2 the expected savings are the black dotted vertical lines on the left graph, reflecting what the energy savings would have been if the rebound effect was zero (difference between the blue line and the green dotted line in the figures). Second, the actual savings are what actually happened between the two groups, in other words the difference in energy consumption between the low efficiency houses group and the high efficiency houses group (difference between the blue line and the red line in the figures). The results of applying equation (1) in both methodologies are shown in Figure 6.3.

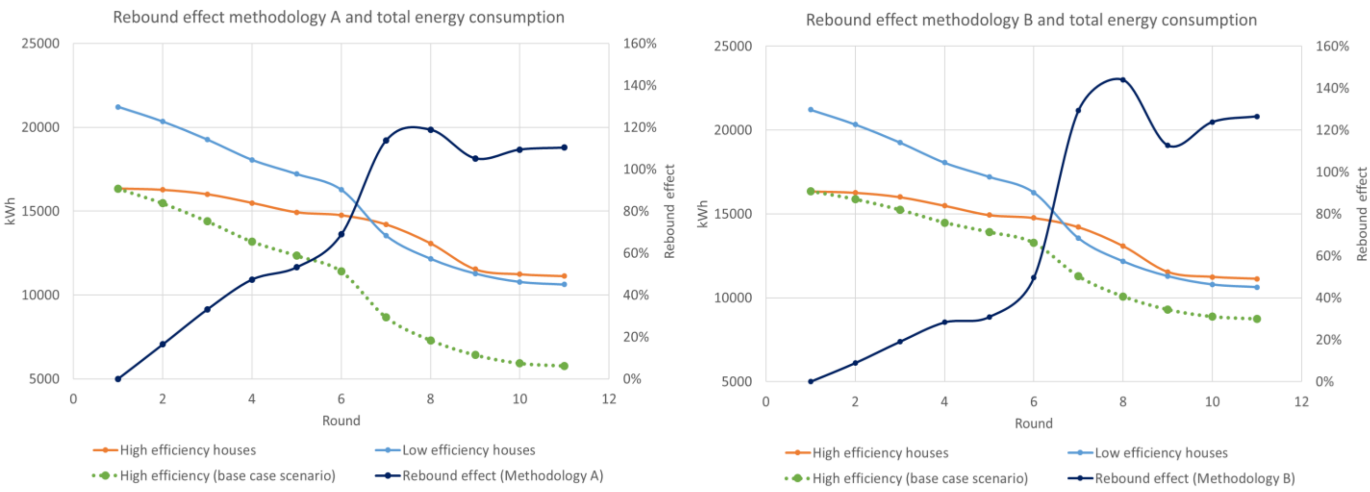


Figure 6.3. Rebound effect calculation using two different methodologies.

In both graphs of Figure 6.3 we can see that the rebound effect is different. Because the expected savings under methodology A are bigger than under methodology B, the final calculation of the rebound effect also differs. What it is equal and unchangeable between both methodologies are the actual savings due to the fact that they only depend to the players' behavior. We can identify that when the low efficiency houses group consumption is higher than the high efficiency houses group consumption, the rebound effect is smaller than 1, meaning that some expected savings are taken back by the rebound effect, but still some savings are accomplished. However, when the low efficiency houses group consumption is smaller than the high efficiency houses group consumption, the rebound effect is bigger than 1, meaning that all the expected energy savings are taken back by the rebound effect and there are no energy savings accomplished whatsoever. The latter is commonly called the back-fire rebound effect, the phenomena in which the rebound effect is greater than 1. To clearly show the differences between the two calculated rebound effects we have constructed Figure 6.4.

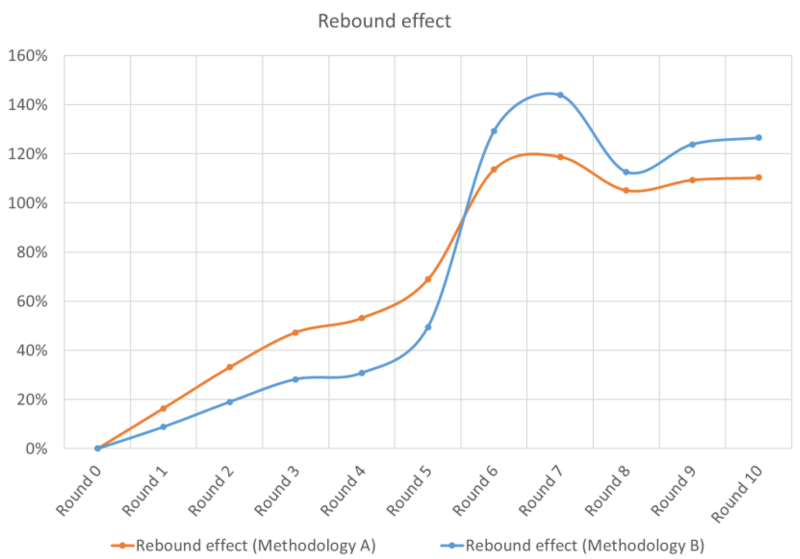


Figure 6.4. Rebound effect using two different methodologies.

As previously mentioned, the differences between the two rebound effects in Figure 6.4 are explained by the difference in the two base case scenario we have calculated. Having different scenarios of the best the high efficiency houses group could have done to reduce their energy consumption changes the expected savings in equation (1). As a result, we end up with a different curve in each case. However, we can also see that those two curves have the same shape and the differences between them are not considerably important. However, proving that those two curves show a statistically significant difference is not possible due to the fact that the rebound effect was calculated using the average energy consumption per group per round. Since players were not exposed to efficiency changes in the middle of the game (remember the modified before/after analysis setup described in section 4.2.1), measures of the rebound effect per player are not available and the only choice left is to analyze the phenomenon just by using average numbers. Therefore, the spread of the curves in Figure 6.4 is not available and no statistical analysis can be carried out to check how different the two rebound effects are.

6.2.3. Analyzing different appliances

Having described the rebound effect on the total energy consumption, we can also explore it a little bit further and try to identify in what appliances the rebound effect is more important. Again, the two defined methodologies will be used for these purposes. The three most consuming appliances will be analyzed next in terms of their average consumption among players: the central heating, the refrigerator and the shower.

6.2.3.1. *Central heating*

The central heating is the appliance that consumes the most in a typical Dutch household, reaching levels to almost 50% of the total energy consumption (gas and electricity) (Friendly and Kormylo, 2012, Gerdes et al., 2014). For the latter reason it is worth to take a closer look to what players did in order to manage their heating systems.

First, we must show that some players replaced their central heating system in the game for heat pumps, devices that use the underground heat to warm up the houses. Although the investment of buying a heat pump in the game was high, players managed to sell their heating systems to partially finance the heat pumps, and therefore, to reduce their gas consumption. Figure 6.5 shows the evolution of the number of players owning central heating systems and heat pumps.

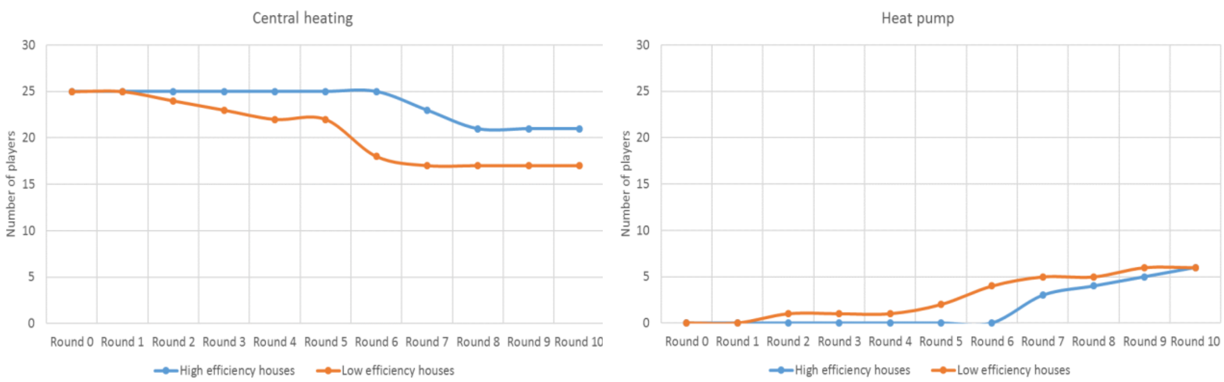


Figure 6.5. Number of players owning central heating and heat pump.

On the left graph in Figure 6.5 we can see that the low efficiency houses group got rid of their central heating systems more than the high efficiency houses group. This result is in line to what we analyzed in chapter 5.5.1 in which we demonstrated that the low efficiency houses group got rid of appliances more than the other group (in terms of the total energy consumption). In addition, the people who kept their heating systems showed a different behavior as well. While people belonging to the low efficiency houses group improved the efficiency of their old heating systems, the high efficiency houses group did not do that. The latter can be seen in Figure 6.6.

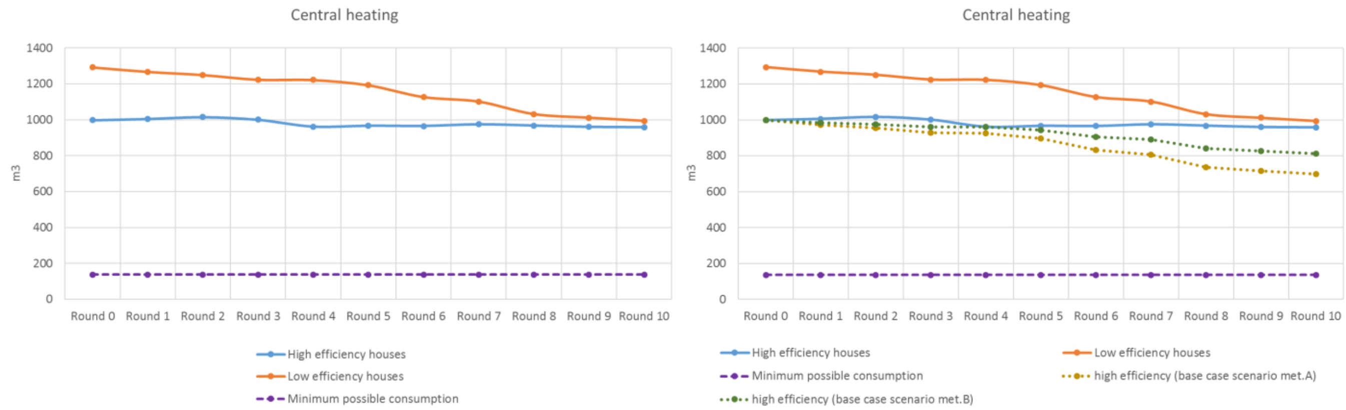


Figure 6.6. Central heating system behavior per group.

The left graph of Figure 6.6 shows that the people who kept their heating systems reduced its gas consumption by 23% (low efficiency houses group) and 4% (high efficiency houses group). In addition, we can also identify in the graph the minimum possible gas consumption of the heating system. If all the measures to reduce gas consumption had been taken by players (the best insulation and the best efficiency of the heating system), the heating system would have consumed a total of 136 m^3 per round. If we now go back to equation (2), we can match the min factor in the formula with the previous number. As a result, and applying once again methodologies A and B, we can calculate the base case scenario (the lowest consumption the high efficiency houses group could have had). The base case scenarios can be seen in the right graph of Figure 6.6. Once again we see that methodology B considers the less room for improvements the high efficiency houses group has over the other group, while methodology A does not consider this. We are now able to calculate the rebound effect of the central heating by using equation (1), in the same way as previously done. Figure 6.7 shows the results of this calculation.

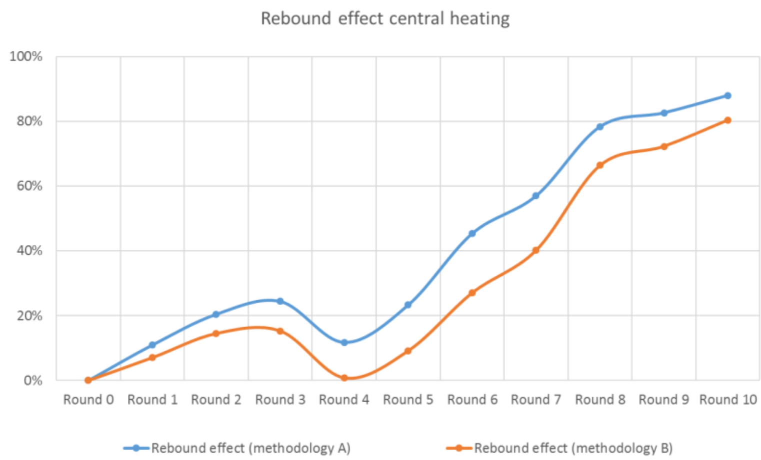


Figure 6.7. Rebound effect central heating.

It can be seen once again that the two rebound effects calculated in Figure 6.7 are similar to one another, not showing any major difference. They are both below 1 due to the fact that the low efficiency houses group consumption is always above the high efficiency houses group consumption (Figure 6.6).

6.2.3.2. Shower

The shower is one of the appliance that consumes the most in a typical Dutch household, reaching levels to almost 15% of the total energy consumption (gas and electricity) (Friendly and Kormylo, 2012, Gerdes et al., 2014). For the latter reason it is worth to take a closer look to what players did in order to manage their shower.

Contrarily to what happened with the central heating in which players replaced it by heat pumps, players did not replace their showers by other appliances due to the lack of alternatives. Just one player was identified to have gotten rid of his/her shower. Of the 49 other players who kept their showers, we can identify that the low efficiency houses group decreased the gas consumption of their showers by 22% while the high efficiency houses group reduced it by 12%. The latter can be seen in the left graph of Figure 6.8.

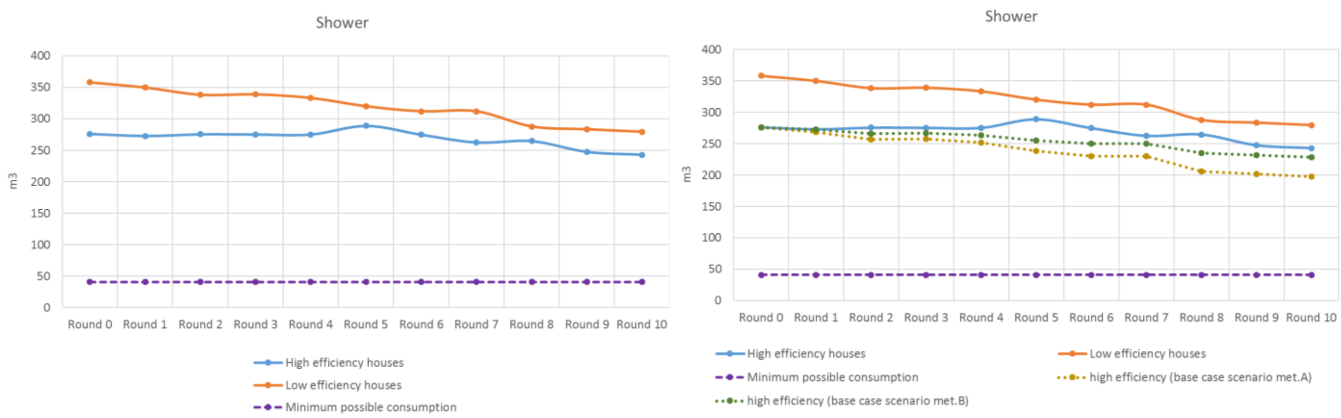


Figure 6.8. Shower behavior per group.

As done with the central heating case, we can calculate the lowest shower gas consumption if all the measures to reduce it were taken (best insulation and best shower efficiency). Therefore, the minimum possible consumption of the shower is equal to 41 m^3 per round. Once again, we can now calculate the best the high efficiency houses group could have done to reduce their shower gas consumption by applying methodologies A and B. In the right graph of Figure 6.8 we can see both base case scenarios. If we now apply equation (1) to calculate the rebound effect of the shower, we can get the behavior shown in Figure 6.9. Once again, both rebound effects under both methodologies are similar in their shape, and methodology A shows a slightly higher value than methodology B. In both cases the rebound effect is less than 1, showing that although some energy savings were taken back by the rebound effect, still some energy savings are accomplished.

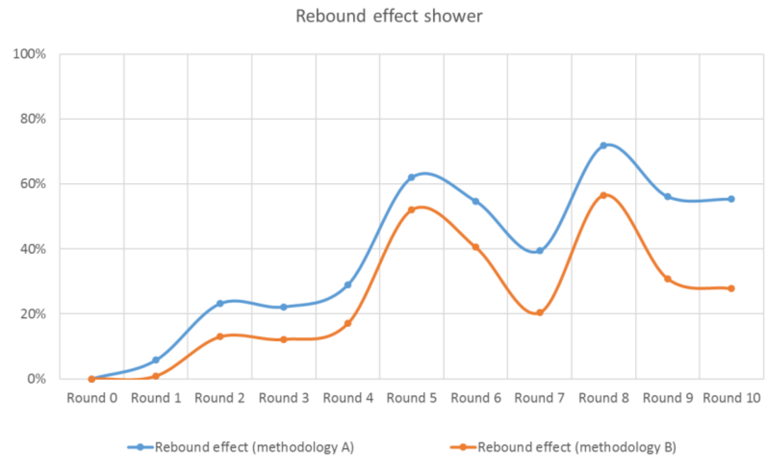


Figure 6.9. Rebound effect shower.

6.2.3.3. <u>Refrigerator</u>

The refrigerator is one of the appliances that consumes the most in a typical Dutch household, reaching levels to almost 5% of the total energy consumption (gas and electricity) (Friendly and Kormylo, 2012, Gerdes et al., 2014). For the latter reason it is worth to take a closer look to what players did in order to manage their refrigerators.

As happened with the shower, players did not replace their refrigerators by other appliances due to the lack of alternatives. Just one player was identified to have gotten rid of his/her refrigerator. Of the 49 other players who kept their refrigerators, we can identify that the low efficiency houses group decreased the electricity consumption of their refrigerators by 34% while the high efficiency houses group reduced it by 28%. The latter can be seen in the left graph of Figure 6.10.

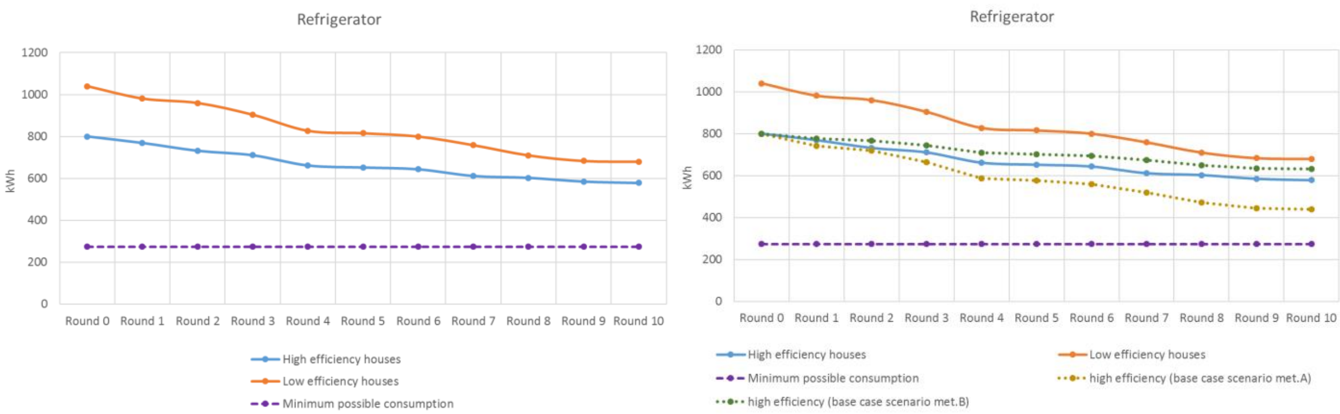


Figure 6.10. Refrigerator behavior per group.

We can again calculate the lowest refrigerator electricity consumption if all the measures to reduce it were taken (best energy management devices and best refrigerator efficiency). Therefore, the minimum possible consumption of the refrigerator is equal to 274 kWh per round. Once again, we can now calculate the best the high efficiency houses group could have done to reduce their refrigerator consumption by applying methodologies A and B. In the right graph of Figure 6.10 we can see both base case scenarios. This time something different happened, the high efficiency houses group base case scenario in methodology B is located above the real consumption of the high efficiency houses group, in contrast to what happened with the central heating (Figure 6.6), shower (Figure 6.8) and the global energy consumption (Figure 6.3). Despite of the previous result, the base case scenario in methodology A is still below the real consumption of the high efficiency houses group. This phenomenon suggests that the final rebound effects using both methodologies may not be similar as they were in the previous cases. Therefore, if we now apply equation (1) to calculate the rebound effect of the refrigerator, we can get the behavior shown in Figure 6.11.

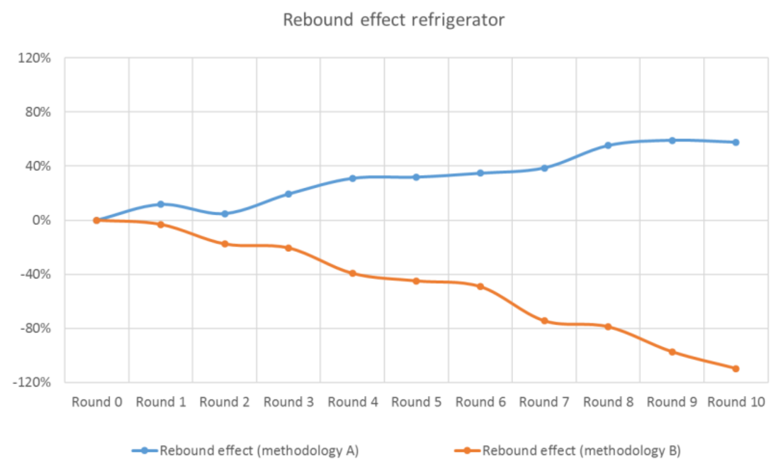


Figure 6.11. Rebound effect refrigerator.

As briefly mentioned before, the rebound effect calculation using methodology A and methodology B shows a totally different behavior, in contrast to what we saw in the previous cases. This case the rebound effect using the base case scenario given by methodology A gives a positive rebound effect reaching almost 60%. Nevertheless, if we see the case using the base case scenario given by methodology B the rebound effect is negative and reaching almost -120% in round 10. The reason why the previous two methodologies differ is explained by the position of the real high efficiency houses group consumption with respect of the two base case scenarios. In the case of methodology B we can see that the players reduced their refrigerators electricity consumption even below to the best they could have done (given by the base case scenario line). As a result, the rebound effect is negative and the actual energy savings were even bigger to the expected ones. In contrast, methodology A supposes that the high efficiency houses group consumption could have gone down even more, reason why the rebound effect is positive, showing that some energy savings were taken back by the rebound effect. Consequently, looking at the refrigerator example we can conclude that the rebound effect calculation really depends on the way it is defined and the methodology used.

6.3. Chapter's conclusions

This chapter starts describing the presence of the rebound effect into the results. First, it is discussed the fact that the total energy consumption of each group was significantly different at the beginning of the game, and after some rounds those differences were not found anymore. The latter was also seen in the average values of the total energy consumption per group, in which the initial difference in round 0 diminished to zero in round 10. Somehow, the low efficiency houses groups managed to reduce their energy consumption more than the high efficiency houses group. Second, it is also discussed the comfort level differences between both groups. At the beginning of the game no significant differences were found, but at some point in the middle of it those differences became significant. Analyzing the average values of the comfort level, we also saw that both groups started the game with the same comfort, but the high efficiency houses group finished the game with a higher value. As a consequence, we can conclude that the rebound effect significantly affected the behavior of the two groups. Having a low efficiency produced players to reduce their energy consumption more than the other group. Similarly, having a high efficiency produced players to increase their comfort level more than the other group.

Having shown that the rebound effect is present in the results, we defined two different methodologies to calculate the base case scenario (or the lowest possible energy consumption of the high efficiency group). One methodology considered that both groups could have reduced their energy consumption the same, and the second methodology considered that the high efficiency group could have reduced their energy consumption less than the low efficiency group, due to the fact that they were already more efficient. Both methodologies showed similar results when the rebound effect was calculated in the total energy consumption. However, when individual appliances were analyzed, both methodologies differed considerably, specifically in the case of the refrigerator. As a result, we showed that the methodology used to calculate the base case scenario affects the final results, confirming once again that the existence of one single rebound effect size should not be the focus of the studies aiming to assess the rebound effect. Instead, studies should focus on reducing the causes of the controversies around the effect, improving the overall assessment of it. What it is crucial in the present assessment is the fact that the rebound effect was detected to be present in the results, rather than purely give a final and single rebound effect size from the experiments.

Having analyzed the rebound effect, next chapter will get some side conclusions on how the players played the game, considering their strategies and motivations for taking certain decisions.

7. Additional analysis

This chapter analyzes some side results apart from the rebound effect itself. The post-game questionnaire will be analyzed now in further detail to extract those side conclusions.

7.1. Realism

As described in Table 3.1, one of the main weaknesses of serious games for research purposes is the fact that the experiments will always be in a simulated context, in which no matter how realistic the game is, the final behavior of players will be in that simulated environment. However, the more realistic the game is, the more probable it is that the players' behavior get closer to what they would do in real life. Therefore, it is crucial to check how players felt the realism of the game and how probable they think they would behave in a similar way in real life. Those two questions were asked in the post-game questionnaire (question 2). Figure 7.1 shows a histogram of each question[5] [6].

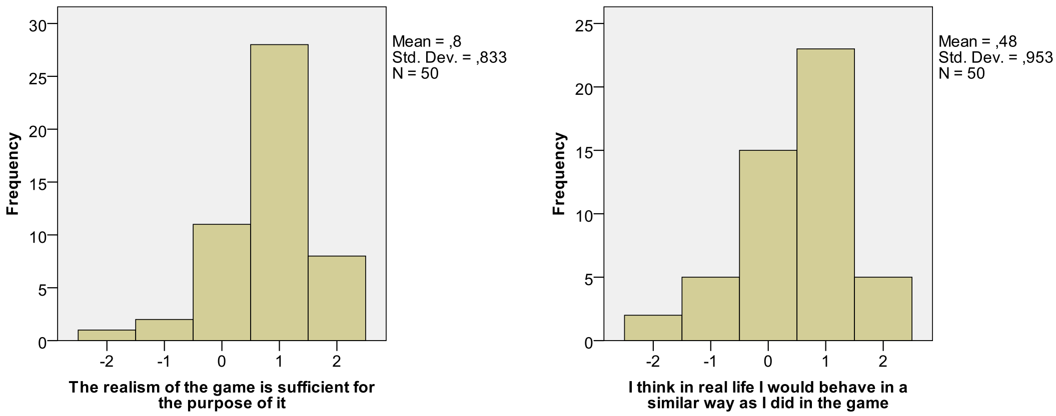


Figure 7.1. Realism and behaving in a similar way in real life.

It can be seen that players rated the realism of the game with an average of 0.8. Despite of the average being almost the "agree" category (they agreed on average that the realism is sufficient), 14 people (28% of the participants) answered that they are not sure if the realism is enough or they (totally) disagree with the statement. This number is high enough to state that the realism of the game is disputable. As a result, the weakness of serious games of being dependent on the realism of the game is expressed in the results and it may influence the validation of the results.

In addition, if we now analyze in Figure 7.1 the question of whether players think they would behave in a similar way in real life, the average is 0.48 (they are, on average, in between of not being sure and being agreed with the statement). Although the average is positive, we

[5] The scale used to transform categorical variables into numeric ones is the following:
Totally disagree = - 2; Disagree = - 1; I am not sure = 0, Agree = + 1, Totally agree = + 2

[6] The purpose of the game was defined in the players' manual (appendix C). Basically, their purpose in the game was 1) play the game as they would do it in real life, and 2) make a house that makes them happy.

can't deny the fact that 22 people (44% of the participants) are not sure or (totally) disagree with the statement. This high percentage affects the validation of the results in a real context. Since players are not completely sure that they would behave in a similar way in real life, we can't apply the previous results into a real context and the results must be kept within the game environment boundaries. Despite of having received a clear instruction of behaving as they would do it in real life, some players did not really do it. Once again, one of the main weaknesses of serious games for research purposes is present.

7.2. Ease and fun to play the game

Another important concept to take into consideration is the ease and fun to play the game. As explained in chapter 3, the trade-off between realism and ease to play must be well balanced. A difficult game may produce that players don't show a realistic behavior due to the difficulty of it. In addition, the more engaging and fun a game is, the more probable it is that players keep their attention throughout the duration of it. As a result, an easy and fun game increases the probability that the results reflect the players' psychological factors and interests.

In the post-game questionnaire, 4 questions about the ease to play the game were asked. We can see a histogram of the responses in Figure 7.2.

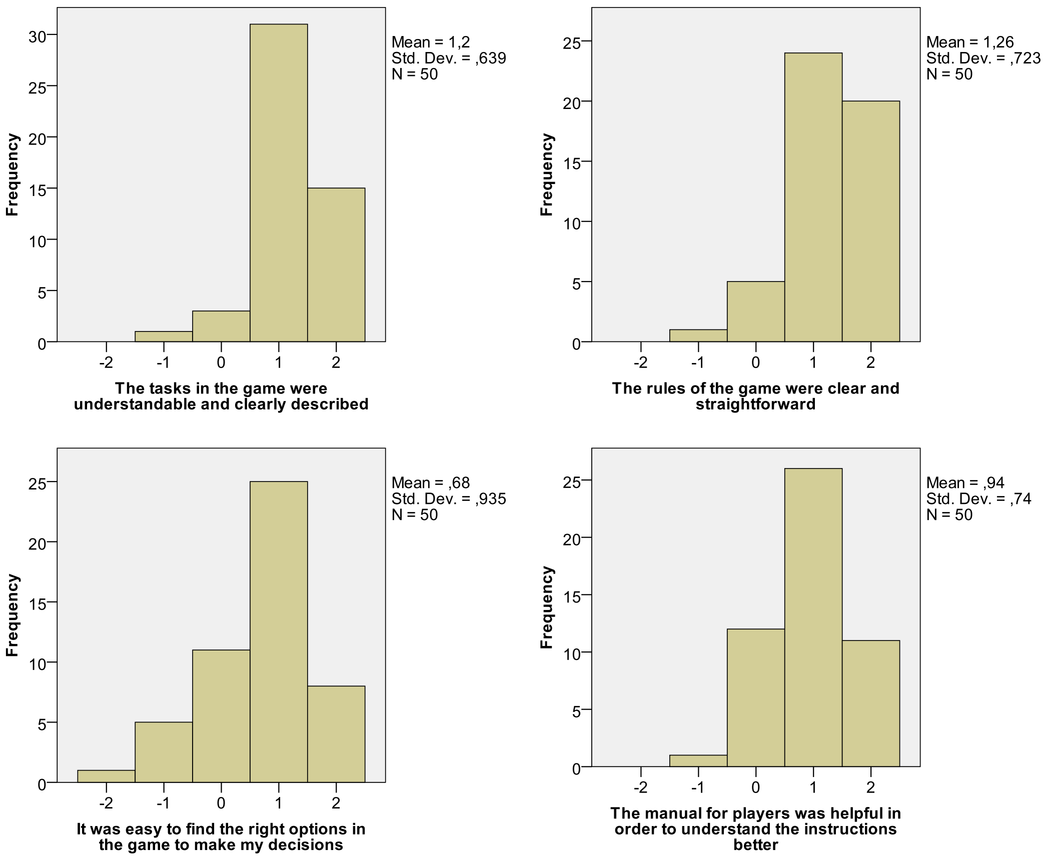


Figure 7.2.Ease to play.

In Figure 7.2 we can see that 46 people (92% of the participants) said that the tasks of the game were understandable and clearly described, 44 people (88% of the participants) said that the rules were clear and straightforward, 17 people (34% of the participants) said that they are not sure or they (totally) disagree with the statement that it was easy to find the right options in the game and 13 people (26% of the participants) said that they are not sure or they (totally) disagree with the statement that the manual for players was helpful. The previous numbers show that for some people the game was not easy, although they rated the rules of the games clear and straightforward. Having a large group of people (34%) who said that it was not easy to find the right options in the game increases the probability that the results do not reflect the real interests of the players, producing problems if we want to translate the results from the game environment into the real world. However, this phenomenon may have happened in the first rounds of the game only, and after some rounds players may have found the right options easier than before, in a process of learning how the game works. As a result, and considering that the whole game took on average 45 minutes, we could say that in the last rounds of the game the ease of the game increased and the 34% of players who said the game was not easy may have decreased.

Another important conclusion to analyze is the fact that players found the rules of the game were clear, somehow contradicting to how the players answered in the question if they would behave in a similar way in real life. One of the main instructions of the game was specifically that one: play the game as you would do it in real life. According to the players, 92% said the tasks were clearly described and 88% said the rules were clear. Despite of having clearly understood the rules, it seems that players didn't followed them so much considering that 22 people are not sure or disagree when they were asked if they would behave in a similar way in real life. The previous result shows that it is not clear if the application the results in a real context is correct.

Now, we can analyze the motivation and fun of the game. Two questions were asked about it in the post-game questionnaire. Figure 7.3 shows a histogram of those answers.

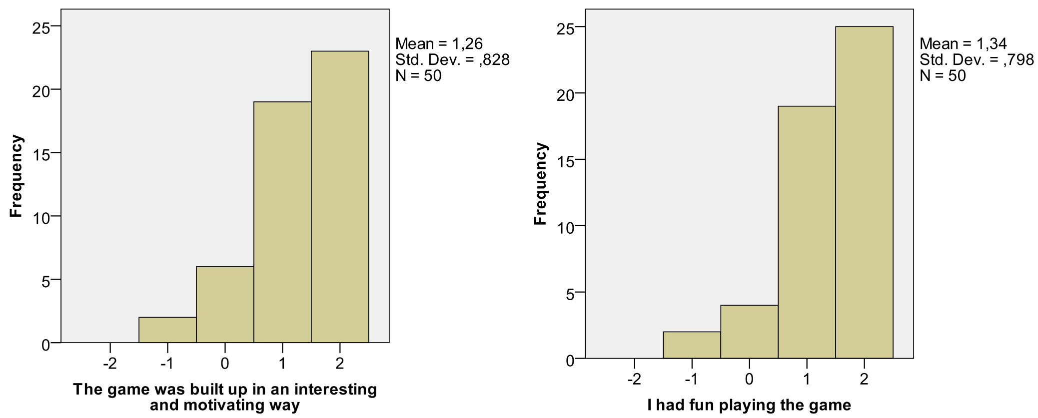


Figure 7.3. Motivation and fun to play.

We can see by looking at Figure 7.3 that 42 people (84% of the participants) said that they (totally) agree that the game was built in an interesting and motivating way, while 44 people (88% of the participants) said that they (totally) agree when they were asked if they had fun playing the game. The previous numbers show that players kept their attention and motivation throughout the game. The latter reflects that their behavior in the game was not influenced by some external factors like boredom which could have decreased the quality of the results. Bored players may have clicked around in the game in a meaningless way, introducing bias and undesirable outcomes.

7.3. Energy management/production devices

It is interesting to take a look at the way players invested in devices to reduce their energy consumption: energy management devices and energy production devices. The post-game questionnaire asked what reasons players had for investing or not investing in these kind of devices. First, we must analyze how many people invested in each one. Table 7.1 shows a cross-table with the number of players investing in them.

Table 7.1. People investing in energy management/production devices.

		Did you invest money in every production devices?		**TOTAL**
		Yes	**No**	
Did you invest money in every management devices?	**Yes**	34	2	36
	No	9	5	14
TOTAL		43	7	50

Looking at Table 7.1 we can see that 36 people (72% of the participants) invested in energy management devices while 43 people (86% of the participants) invested in energy production devices. Despite that difference we must perform statistical tests to check if that difference is indeed significant or it is just caused by sample errors. Since both variables have a categorical level of measurement, we have to apply the chi-square test. The null hypothesis for this test H_0 is “the number of people who invested in energy management devices is the same as the number of people who invested in energy production devices” while the one-tailed alternative hypothesis H_1 is “the number of people who invested in energy management devices is less than the number of people who invested in energy production devices”. Since we find expected counts less than 5 the conditions for the chi-square are not satisfied and we can’t rely on the results. Therefore, the test to perform instead is the Fisher’s exact test. After performing it in SPSS, we obtain a one-tailed significance level equal to 1.4%. As a result, H_0 is rejected and significant differences between the number of people who invested in energy production devices and the number of people who invested in energy management devices were detected. The SPSS results can be found in appendix H.19.

The previous result is interesting considering that energy management devices are more economically profitable than energy production devices, taking into account the energy reduction and the price of it. In the NRG game the same energy savings that solar panels or mini wind turbines bring, could have been achieved by installing energy management devices and by paying less money. For instance, the cheapest energy production device in the game is a mini wind turbine that costs € 1,500 and produces 500 kWh per year. In order to save the same 500 kWh per year, a cheaper option is to buy a smart meter (with an eco-friendly level equal to 0.5) and a stand-by-killer (with an eco-friendly level equal to 1). The latter option costs in total € 575, almost € 1,000 less than the option of buying the mini wind turbine.

As a result, the previous result should have been the opposite if players had behaved in a rational way by analyzing all the possibilities and prices. However, since consumers do not behave rationally they invested more in energy production devices. This result supports the idea that psychological factors of players and their interests were captured by the game considering that they acted irrationally, as consumers do in real life.

If we now analyze some of the reasons of players for investing in both type of devices, we can also get an interesting result regarding what it is more important to them: reducing costs or reducing energy consumption. Figure 7.4 and Figure 7.5 show those two reasons players were asked for investing in energy management devices and energy production devices, respectively.

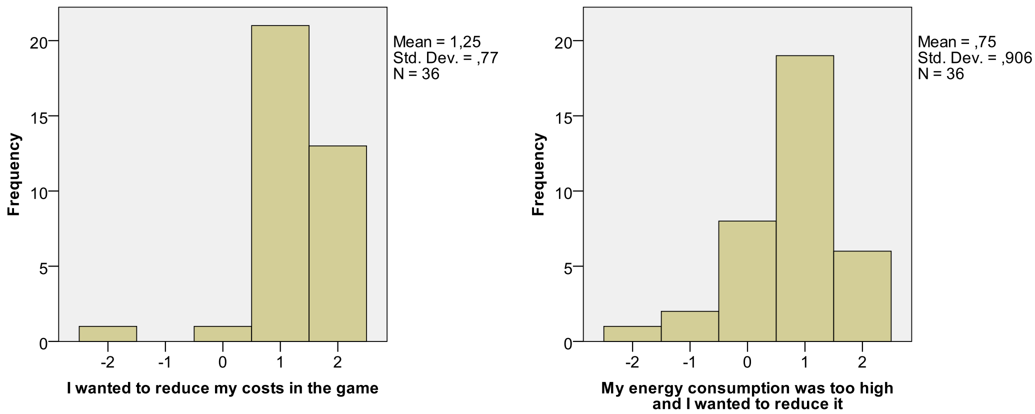


Figure 7.4. Reasons to invest in energy management devices.

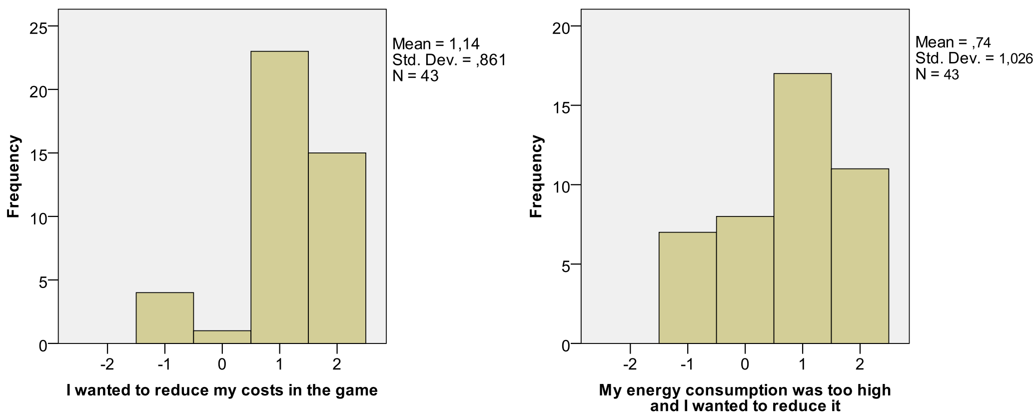


Figure 7.5. Reasons to invest in energy production devices.

We can now compare if there is any significant difference between the two reasons in each type of device. First we need to check if the distribution of those variables follow the normal theoretical distribution. Once again, the non-parametric Kolmogorov-Smirnov test must be performed. In this case the null hypothesis H_0 is "the distribution follows the normal theoretical distribution" while the alternative hypothesis H_1 is "the distribution does not follow the normal theoretical distribution". After performing the statistical test in SPSS we only get significance values below 5%, meaning that none of the previous distributions follow a normal theoretical distribution. The SPSS outcomes can be seen in appendix H.20. As a result, to test significant differences between the two distributions in Figure 7.4 and Figure

7.5 we must perform the non-parametric Wilkoxon Rank-Sum test. In this case the null hypothesis H_0 is "reducing costs is equally important than reducing energy consumption when investing in in energy management/production devices" while the one-tailed alternative hypothesis H_1 is "reducing costs is more important than reducing energy consumption when investing in energy management/production devices". After performing both tests in SPSS, we get that the significance value in the case of the energy management devices is 0.1%, while in the case of energy production devices is 1.8%. The SPSS outcomes can be seen in appendix H.21. In both cases the significance value is below 5%, meaning that the null hypothesis is rejected, meaning that players thought that reducing costs is more important than reducing energy consumption in their reasons for investing in energy management and energy production devices.

The previous result again contradicts what we discovered about more people investing in energy production devices than people investing in energy management devices. We previously stated that energy management devices were economically more profitable than energy production devices, showing that players acted in an irrational way. However, we just showed that players thought that reducing costs in the game was more important than reducing energy consumption as a reason for investing in those kind of devices. Therefore, if players had taken reducing costs with more importance than the other reason, they would have invested more in the more profitable ones. However, the latter did not happen, showing once again the irrational behavior of consumers.

7.4. Strategies and decisions

Another interesting factor to analyze is the strategies players had and the decisions they took. In question 5 of the post-game questionnaire those questions were asked. This time, it is worth to analyze the differences in their answers between the low efficiency houses group and high efficiency houses group.

As previously done, first it must be checked if the answers per group are normally distributed by using the non-parametric Kolmogorov-Smirnov test. After performing the test in SPSS, all the significance values are below 5%, meaning that none of the distributions follow the normal theoretical distribution (appendix H.22 shows the outcomes given by SPSS). As a result, we will proceed with the non-parametric Wilkoxon Rank-Sum test to check if significant differences are detected between groups per each answer. In this case the null hypothesis H_0 is "there is no difference between the two groups in terms of each answer" while the alternative hypothesis H_1 is "there is a difference between the two groups in terms of each answer". This test must be performed as many times as the number of questions (5 in total). Table 7.2 shows a summary of the mean and standard deviation of each answer per group and the significance value of the Wilkoxon Rank-Sum test. Appendix H.23 shows the outcome given by SPSS.

Table 7.2. Strategies and decisions.

Question	Low efficiency houses		High efficiency houses		Significance value
	Mean	Std. deviation	Mean	Std. deviation	
I often checked my gas and electricity consumption	0.84	0.85	0.52	1.295	47.7%
I preferred to buy luxury rather than eco-friendly things	-0.92	0.862	-0.92	0.954	89.1%
I wish I had more money in the game to buy more luxurious things	-0.52	1.194	-0.04	1.207	15.5%
I wish I had more money in the game to buy more eco-friendly things	1.12	0.781	0.8	1.041	27.3%
I wish I had more money in the game to buy more production devices (like solar panels/wind turbines)	1.48	0.823	1.44	0.87	84.8%

After performing the test in SPSS we see that the significance value of each test is greater than 5%, meaning that the null hypotheses can't be rejected. Therefore, no significant differences can be detected between the two groups in terms of how they answered the 5 analyzed questions. It is interesting anyway to point out the differences in the mean of each answer per group. For instance, in the question "I often checked my gas and electricity consumption" the low efficiency houses group had a higher mean than the high efficiency houses group, meaning that on average they checked their electricity and gas consumption more often than the high efficiency houses group. Therefore, we may have thought this as a sign of the rebound effect, in which the people with low efficiency were more worried about how much they were consuming per round. However, the latter conclusion just considers the absolute value of the mean per group. When the spread of the answers is considered no significant differences are detected whatsoever.

7.5. Remembering efficiency

Question 3 of the post-game questionnaire asked whether players remember the initial efficiency of the houses they played with or not (they were told in the players' manual the initial efficiency of their houses, i.e. the group they belong). If they answered yes in question 3, they were also asked if knowing the initial efficiency of their houses influenced them in the way they made their decisions (question 4 of the post-game questionnaire). Therefore, it is worth to analyze what players answered and check if any difference is detected between the two groups. Table 7.3 shows how many players correctly remembered their initial energy efficiency level of the house they played with (either low efficiency or high efficiency).

Table 7.3. Players remembering their initial energy efficiency.

		Do you remember the efficiency of the house you played with?		TOTAL
		Yes	No	
Efficiency group	Low	11	14	25
	High	13	12	25
TOTAL		24	26	50

It can be seen than no major differences between the two groups are detected. Nevertheless, statistical tests to check the latter must be applied to make a conclusive statement. Since both variable have a categorical level of measurement, we have to apply the chi-square test. The null hypothesis for this test H_0 is "the number of people who remembered their initial energy efficiency is the same across the two efficiency groups" while the alternative hypothesis H_1 is "the number of people who remembered their initial energy efficiency is not the same across the two efficiency groups". After performing the test in SPSS, we obtain a significance value equal to 57.1%. As a result, H_0 be can't rejected and no differences between both groups are detected in terms of the number of people who remembered their initial energy efficiency. The SPSS results can be found in appendix H.24.

In addition, and inside the group who answered "yes" in Table 7.3, we can also analyze how many people think they were influenced by knowing it. Figure 7.6 shows the distribution of that answer.

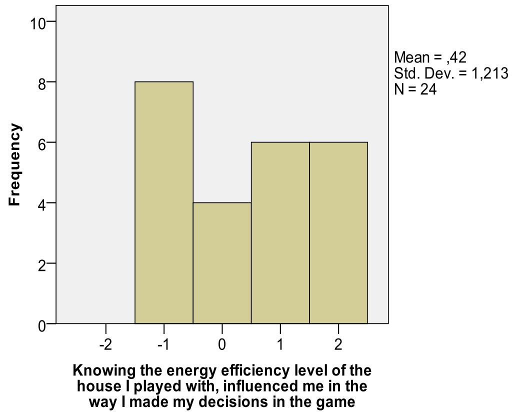


Figure 7.6. People who remembered their initial energy efficiency.

Looking at Figure 7.6 we can say that people who remembered their initial energy efficiency are in between of not being sure and being agreed that they were influenced by knowing this information. If we consider that just 12 people (24% of the participants) said that they were influenced by knowing their initial energy efficiency, we can conclude that the complete behavior of all players is not influenced by this fact. Therefore, what we have analyzed about the rebound effect in chapter 5 and about players' behavior in chapter 6 is an unconscious process that players are not aware. The fact that the low efficiency houses group reduced their energy consumption to levels that in round 10 both groups don't show any significant difference is a result of their own psychology and interests, and it is not because the initial stimuli they were given about the initial efficiency of their houses. Therefore, players' decisions were shaped by economic factors mainly, in which the low efficiency houses group felt the necessity of reducing their energy consumption more than the high efficiency houses group.

Additionally, we can also support the previous result by analyzing if some significant relations are found between the fact of remembering the efficiency of their houses and, for instance, the total energy consumption, energy production or the comfort level. The non-parametric Wilkoxon Rank-Sum test will be used to check this. In this case the null hypothesis H_0 is "there is no difference in the total energy consumption/energy production/comfort level between the group who remembered their initial energy efficiency and the group who didn't remember it" while the alternative hypothesis H_1 is "there is a difference in the total energy consumption/energy production/comfort level between the group who remembered their initial energy efficiency and the group who didn't remember it". This test must be performed 30 times in total (one time per round for each of the three indexes). The SPSS results can be found in appendix H.25. The significance values of the 30 tests are in between 14% and 96.9%, meaning that in all the cases the null hypothesis can't be rejected, meaning that no significant differences can be found in the total energy consumption, energy production or the comfort level between the people who remembered their initial energy efficiency and the people who did not. As a result, we can state that the behavior of the players was not influenced by letting them know their initial energy efficiency, and their behavior is assumed to have been influenced by other factors like their own psychology and interests.

7.6. Environmental friendliness relations

The last relation that will be tested is the possible influence of the environmental friendliness factor of each player over the total energy consumption, energy production and comfort level. There might be a significant relation between them. In order to do this, we will check the scatter plot of each relation per round. In Figure 7.7, Figure 7.8 and Figure 7.9 the scatter plots per round can be found. Analyzing the figures, we can't identify any significant linear relation between the variables and the environmental friendliness of players in any round. The only that can be seen in the graphs is a cloud random behavior, showing no significant linear relations whatsoever. Furthermore, one of the conditions of the Pearson's correlation coefficient test to statistically check this type of relations, is that the analyzed variables visually show some kind of linearity. Since this is not the case, performing the test to calculate correlation coefficients is not applicable. Therefore, having found no significant relations between the variables constitutes an important result in which no matter how players rate themselves in the environmental friendliness index defined in this research, this will not be reflected in the way players actually behaved. It could have been expected though that the more environmental friendly a player is, the more energy production he/she will show, or the less total energy consumption will have. However, the previous relations are not present, confirming that what players did in the game is not influenced on how they see themselves in terms of their environmental friendliness.

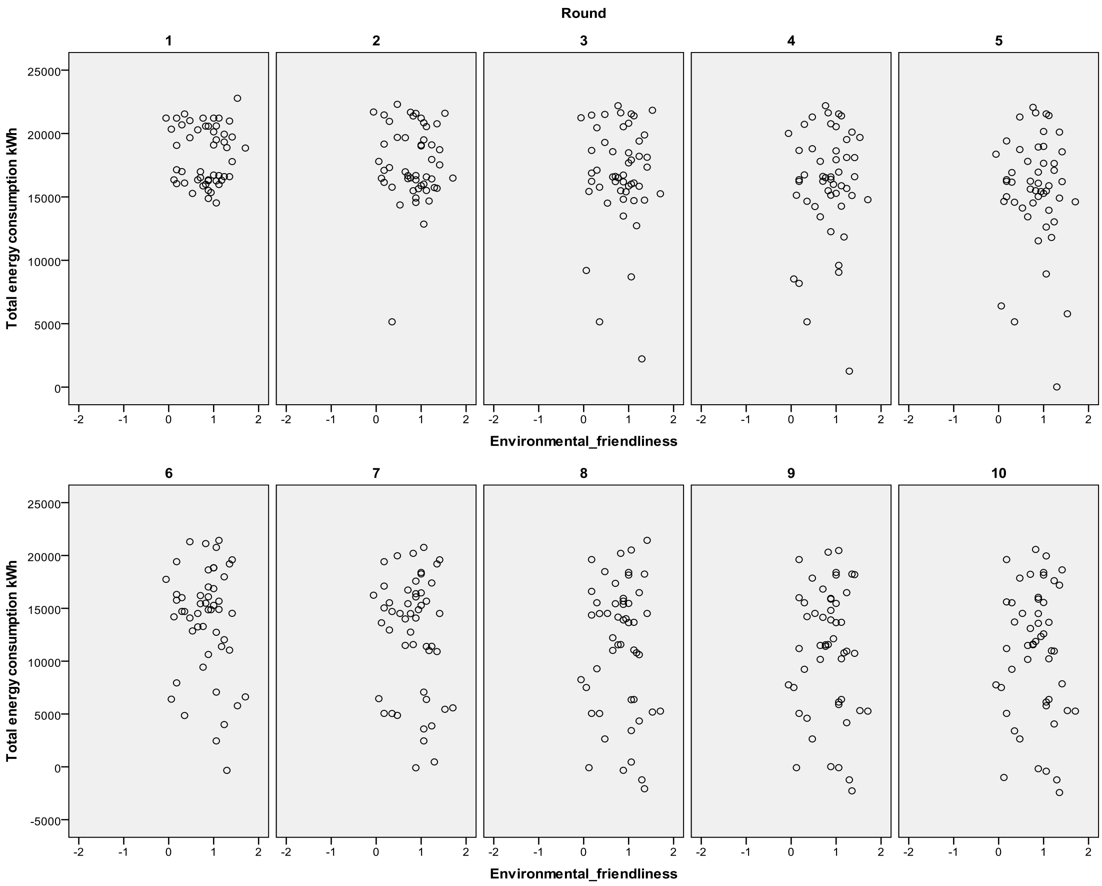


Figure 7.7. Total energy consumption vs environmental friendliness per round.

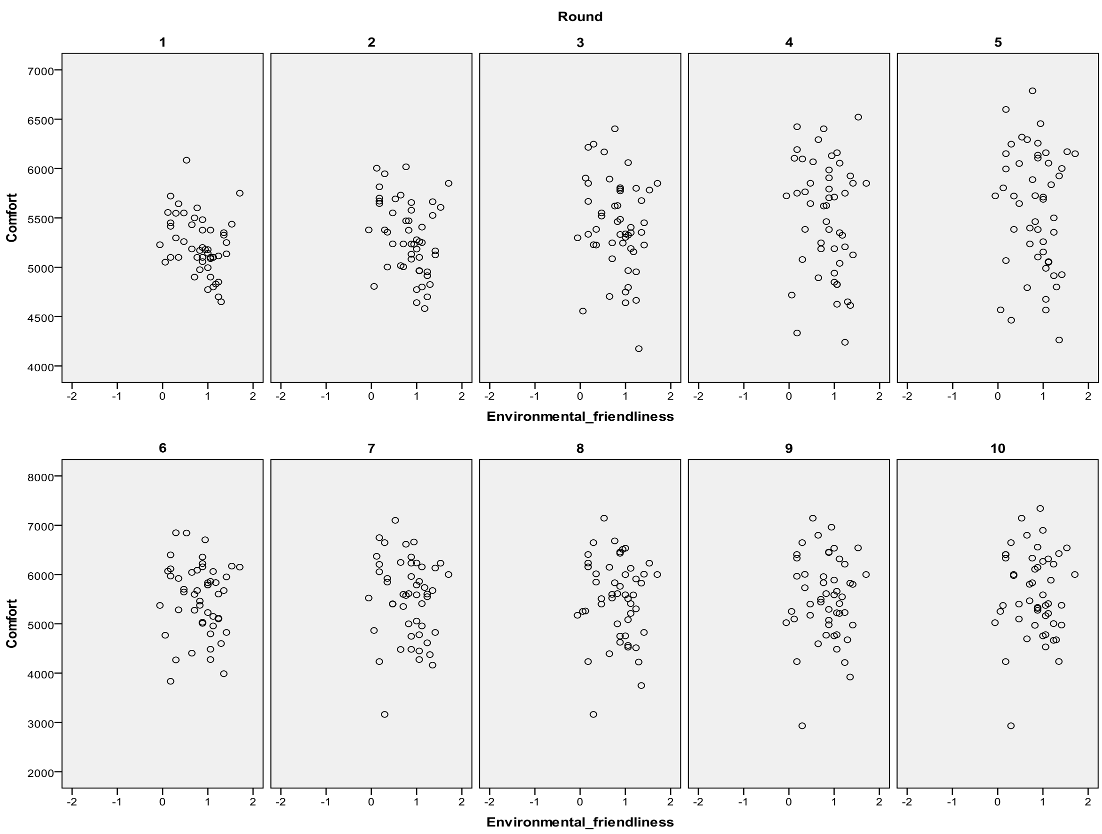


Figure 7.8. Comfort level vs environmental friendliness per round.

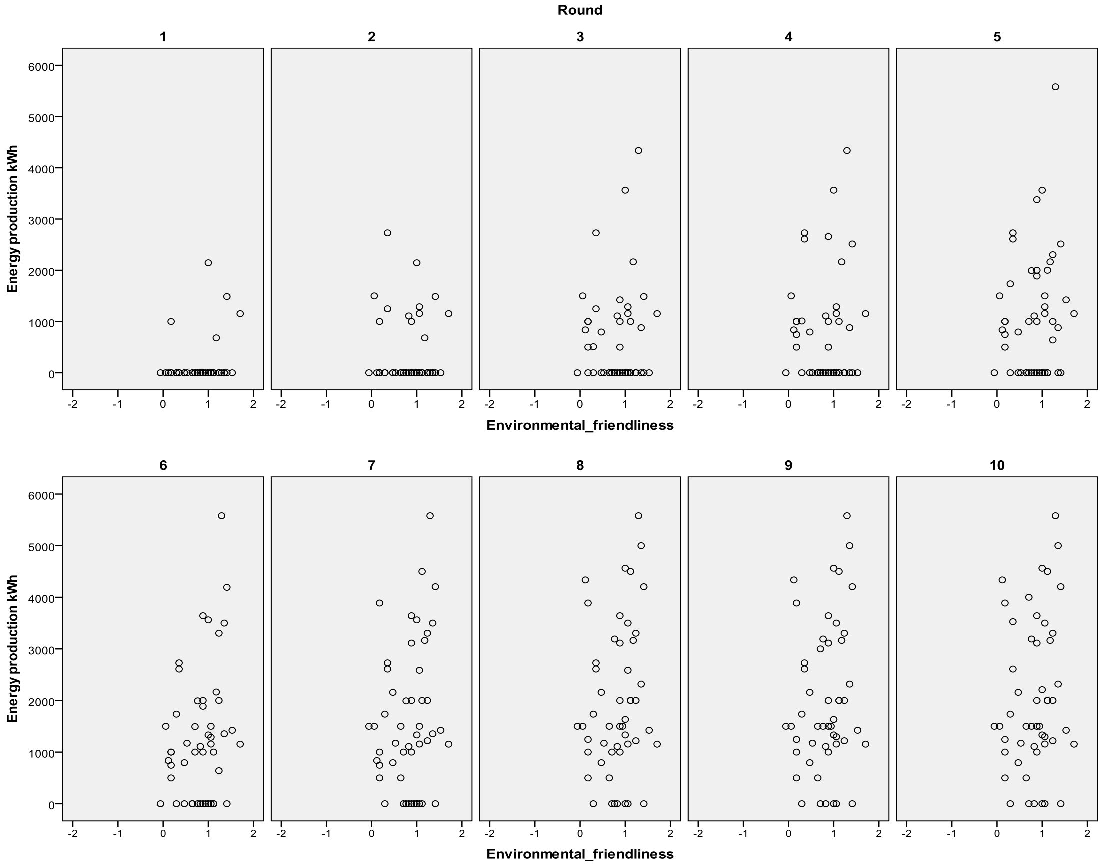


Figure 7.9. Energy production vs environmental friendliness per round.

7.7. Chapter's conclusions

This chapter starts analyzing one of the weaknesses of serious games and if it is present in the results: the realism. The realism of serious game is crucial because it defines if the results from the game may be applicable in a real life context. The realism of the NRG game was felt by players not to be clearly sufficient, and they are not completely sure if they would behave in real life in a similar way as they did in the game. Therefore, the applicability of the results of this research into a real life context is compromised. In addition, another critical factor for the applicability of the results into a real life scenario is the ease of the game. Although players felt that the rules of the game and the tasks were clear and straightforward, a considerable number of players felt that the game was not totally easy. The latter, as happened with the realism of the game, compromises the applicability of the results of this research into a real life context.

In addition, it was also discovered some irrational behavior of players. Players felt that reducing costs was more important than reducing energy consumption, which contradicts the fact that they invested more in the less profitable devices when trying to reduce energy consumption. Having detected some irrational behavior of players constitutes an important finding to support the idea that psychological factors of players were captured by the experiments due to the fact that irrational behavior is always present in real life decisions as well. Finally, it was also discovered that no matter how environmental friendly players think they are, that is not reflected on what they do in the game.

8. Conclusions

This chapter concludes the present thesis research. It shows the main conclusions of it. It is organized as follows. First, the research questions proposed in chapter 1 are answered. Second, limitations of the present research are given. Third, and final, recommendations and ideas for future research are given.

8.1. Research questions

In the introduction of this thesis report the research questions were defined. The main research question was:

How can the rebound effect be adequately assessed in order to reduce the causes of controversies?

In addition, the previous main research question was broken down into three sub-research questions to operationalize it. In answering those sub-questions, we will be answering the main one.

1. What are the main controversies associated to the rebound effect concept?
2. What new and innovative methodology can be proposed to assess the rebound effect in order to reduce the causes of controversies?
3. What are the results and lessons learned when the chosen methodology is used to perform a new rebound effect assessment?

In the following, the previous three sub-research questions are going to be answered.

8.1.1. Question 1: rebound effect's controversies

The rebound effect is the phenomenon in which energy savings, after energy efficiency improvements, are lower than expected. The rebound effect has been broadly analyzed by several scholars, especially in the last 20 years when reducing energy consumption has become critical considering not only the depletion of fossil fuels but also the global warming phenomenon. Despite of being widely analyzed, the rebound effect has been catalogued as a controversial concept by several scholars (Binswanger, 2001, Dimitropoulos, 2007, Greening et al., 2000, Grepperud and Rasmussen, 2004, Hertwich, 2008, Lin and Liu, 2012, Sorrell, 2007). Two main controversies around the rebound effect were identified: the size of the rebound effect and the importance for the policy making process. First, the size of the rebound effect greatly differs between different studies. While some scholars have obtained a certain size of the rebound effect, other scholars have obtained a totally different calculation, even in the same energy service (lighting, transportation or household heating). As a consequence, several authors have been focusing their studies on finding consensus on what the real size of the rebound effect is (Gillingham et al., 2013, Santarius, 2014). However, some other authors believe that the rebound effect is unique for every different situation in which is analyzed, therefore, it is useless trying to find consensus on what the size of the rebound effect is, since this unique size does not exist (Aydin et al., 2015, Sorrell, 2009). Second, the importance for the policy making process is also a controversy. On the one hand, some scholars claim that no matter how big the rebound effect is, some energy savings will be always achieved after efficiency improvements. Therefore, it is not worth, they claim, that policymakers take actions to reduce the rebound effect considering that energy savings are always accomplished (Gillingham et al., 2013). On the other hand, some other scholars take a totally different view, claiming that no matter how low the rebound effect is, policymakers should always take extra actions to reduce it in order to increase energy savings as much as possible (Binswanger, 2001, Dimitropoulos, 2007, Sorrell, 2009).

Additionally, it has also been identified three causes that have produced the previous controversies: methodological issues of previous attempts that have analyzed and measured the rebound effect, different numeric definitions and fuzzy and different system boundaries. This thesis research's objective was to propose a new methodology to assess the rebound effect in order to improve the methodological issues related to previous before/after analysis. Previous before/after rebound effect assessments have shown some methodological issues that have contributed to the rebound effect's controversies to appear: difficult to strictly follow ceteris paribus and make just one change at a time, difficult to isolate effects different from the rebound effect, it constitutes basically an ex-post rebound effect calculation, control groups are not present or they are not clearly defined, just analysis of one single energy service have been carried out and psychological factors of players are not considered into the assessments.

This thesis research proposed a new and innovative methodology to assess the rebound effect by performing a new before/after analysis tackling the previous methodological issues. The methodology that was used is serious games. In the answer of the following research question the reasons of why serious games help to reduce the methodological issues are explained.

8.1.2. Question 2: methodology to assess the rebound effect

This thesis research's objective was to find a new methodology to assess the rebound effect that helps to reduce the causes of controversies, specifically the methodological issues. After identifying previous before/after rebound effect assessments' shortcomings, we can propose that serious games help to tackle them. In fact, the literature review about serious games showed that the advantages of serious games for research purposes (the ones in which hypotheses are tested and the researcher learns from the behavior of players) are the identified shortcomings of previous before/after analysis that have analyzed the rebound effect in the past. First, due to the high flexibility serious games have, the ceteris paribus principle can be followed by making just one change at a time. In fact, system boundaries and control variables can be totally controlled by using serious games, reason why it is possible to just change the energy efficiency of the experiments, keeping all the other variables the same. Second, and related to the previous point, serious games can be used to isolate effects due to its flexibility. In doing so, a pure rebound effect assessment may be obtained, not influenced by third variables that may just introduce bias to the results. Third, since serious games for research purposes can be used for testing hypotheses, the rebound effect can be ex-ante analyzed, giving policymakers a new tool to assess its consequences before implementing efficiency improvements policies. Fourth, clear control groups can be defined by using serious games due to the fact that the experiments are totally controllable by the researcher, and energy efficiency improvements can be applied just to one specific group. Therefore, the group which does not receive the stimulus, or in other words the control group, can be very well defined. Fifth, since the system boundaries can be well defined when serious games are used, different energy services can be analyzed at the same time. So far, the rebound effect has been just analyzed in single energy services. Nevertheless, if serious games are used, different energy services can be included in the same rebound effect assessment. Sixth, serious games do not have to build in the model psychological aspects of people since the players themselves include their own interests and mindset in the experiments. As a result, no inaccurate assumptions about how consumers behave when consuming energy have to be made. As a consequence, if we manage to use serious games for non-research purposes to perform a before/after rebound effect assessment, the identified shortcomings may be tackled and the controversies may be reduced.

8.1.3. Question 3: results and lessons learned

The chosen serious game to perform a new rebound effect assessment was the NRG game. After performing a modified before/after analysis to assess the rebound effect, several lessons were learned and conclusions obtained (in section 8.2.1 we will discuss the limitations about having performed a modified before/after analysis). In the following, those main lessons learned are discussed.

8.1.3.1. *Academic and societal contribution*

After performing the assessment by using the NRG game, the rebound effect was, indeed, detected to be present. Therefore, measuring the rebound effect by using serious games is possible. The modified before/after analysis carried out in this thesis research is indeed an important academic contribution due to the fact that it was possible to tackle the shortcomings of previous before/after analysis that have analyzed the rebound effect, improving the quality of them and reducing one of the causes of the rebound effect controversies: methodological issues. First, it was possible to follow the ceteris paribus principle and make one change at a time. In this case the one and only change and difference between the before and after situation was improvements in energy efficiency. Second, due to the fact that only the energy efficiency was improved and all the other factors remained the same, it was possible to isolate the rebound effect phenomenon and not influence it by some other external factors. Third, a clear control group was defined: the low efficiency houses group. This group was the one used to calculate the base case scenario in which the rebound effect was zero. Fourth, an ex-ante rebound effect assessment was carried out, increasing the possibilities policymakers have to evaluate the importance of the rebound effect before implementing energy efficiency improvements policies. Fifth, it is the first time multiple energy services are analyzed at the same time, including electricity and gas consumption in households altogether. So far, just studies analyzing one single energy service had been carried out. This new study included indirect rebound effects when people transfer their gas consumption, for instance, into electricity consumption. Previous studies have not been able to include this phenomenon. Sixth, psychological factors and the interest of players were included into the assessment. In fact, irrational behavior of players was identified to occur in the game, just as in real life in which consumers act irrationally, driven by their own psychology and interests. This advantage of serious games makes it easy to include this factor into the experiments, contrarily to what happens with some other methodologies that have to make inaccurate assumptions of how consumers behave. The different types of behaviors consumers may show are countless. Therefore, by making people to include themselves their own psychology into the experiments, no inaccurate assumptions are made and the whole spectrum of possible behaviors are included as well. For all the previous reasons, we can state that this rebound effect assessment improved the methodological issues of previous before/after analysis that have analyzed and measured the rebound effect in the past. Therefore, this assessment constitutes and important academic contribution towards the reduction of the rebound effect's controversies.

In addition, we can also identify that this research is an important societal contribution. Policymakers have now a new methodology or tool to ex-ante analyze the rebound effect, considering that this is the first time serious games have been used to measure the effect. New information can be gathered about the consequences of the rebound effect and how people react when they are exposed to energy efficiency improvements. However, it is important that policymakers take into account all the disadvantages of serious games when using this tool, like the difficulty to conclude if the results in the game environment can be applied into a real life scenario. In the rebound effect assessment carried out in this thesis research, it was obtained that the results can't be applied into a real life context due to the fact that players did not think that they would behave in real life in a similar way as they did in the game. In order to guide policymakers about the possible problems to overcome when serious games are used to assess the rebound effect, limitations of this research are explored in section 8.2.

8.1.3.2. *About the assessment of the rebound effect*

The assessment of the rebound effect consisted of a modified before/after analysis, in which one group (low efficiency houses group) was defined as the case "before" and other group (high efficiency houses group) was defined as the case "after". Since those two groups are formed by different people, we had to make sure first if those two groups were indeed comparable in order to carry out the before/after comparison. Four criteria were defined to evaluate if those two groups were statistically similar: environmental friendliness, abilities to play video games, frequency they play video games and whether they pay bills, rent or mortgage in real life. If differences between the two groups in some of the four criteria had been identified, the final comparison would have been compromised. Fortunately, after statistical tests no differences between the two groups were found whatsoever. In order to get the previous result, it was necessary to carry out a pre-game questionnaire that all players answered. This questionnaire was designed to this ultimate purpose, to check how comparable the two groups were.

The sample size of each of the two groups was 25 people, making a total sample size of 50 people. Each of the two groups of players played with a specific house in the NRG game, either low efficient or high efficient. Two main indicators were tracked throughout the game in all the players: energy consumption and comfort level. First, analyzing energy consumption in each of the two groups, it was discovered that the initial difference between them disappeared during the game. At the beginning of the game the energy consumption difference between the groups was significant, while in some point in the middle of the game this energy consumption difference was not found to be significant anymore. The previous result shows a first sign of the rebound effect, in which the low efficiency houses group reduced their energy consumption down to levels in which the initial significant difference vanished. Second, analyzing the comfort level of players, we also discovered that despite of not showing any significant difference in the comfort level between the two groups at the beginning of the game, in some point in the middle of it the two groups split up and significant differences in their comfort level were detected. This is a second sign of the rebound effect, in which the comfort level of the high efficiency houses group increased more than the low efficiency houses group. The latter concludes that having better efficiency helps people to increase their comfort due to the fact that they have more money available to spend because the lower energy bills they have to pay. Therefore, analyzing the energy consumption and comfort level of players, the rebound effect was detected to be present. Due to how the NRG game works and what kind of decisions players can take in the game, indirect rebound effects are included in the system boundaries, and direct rebound effects are not immediately included.

Having detected the rebound effect, a calculation of it was carried out. The first step was to calculate a base case scenario, a behavior that the high efficiency houses group would have behaved like if the rebound effect was zero. Any deviations of the high efficiency houses group from that base case scenario would be labeled as rebound effect. Two different methodologies to define the base case scenario were implemented: 1) considering that the high efficiency houses group have less opportunities to improve their energy efficiency since they are already high efficient, and 2) not considering this difference in the opportunities to improve energy efficiency. First, calculating the total rebound effect (including gas and electricity), the results show that all the energy savings were taken back by the consumers due to a higher energy consumption of the low efficiency houses group over the high efficiency houses group at the end of the game. This shows a back-fire rebound effect, a process in which there are no energy savings left after efficiency improvements (or in other words, a rebound effect greater than 100%). The two methodologies previously mentioned to calculate the base case scenarios showed the same behavior, with only minor differences. Second, calculating individual rebound effects of the most consuming appliances in households, the results differ from the global rebound effect. The central heating and shower showed a rebound effect below 100% and the two methodologies only showed minor differences. Therefore, analyzing only central heating and shower, some energy savings are achieved, although less than the expected ones. However, analyzing the refrigerator, major differences were detected between the two methodologies to calculate the base case scenario. On the one hand, the methodology in which differences in the opportunities to improve energy efficiency are considered, showed a negative rebound effect, meaning that the actual energy savings were bigger than the expected ones. On the other hand, the methodology in which that difference is not considered showed a positive rebound effect, meaning that actual energy savings were less than the expected ones. The latter result shows that the way the base case scenario is calculated is crucial for the final calculation of the rebound effect. This constitutes one of the major limitations of this research that will be discussed on further detail in section 8.2.2. Despite having identified this major limitation, the fact that the rebound effect was indeed detected is not compromised. What is compromised is the reliability of the final number of the rebound effect, not the fact that the rebound effect was identified to be present. Nevertheless, not being able to rely on the number of the rebound effect is not crucial considering the uniqueness characteristics of the rebound effect proposed by Aydin et al. (2015) and Sorrell (2009).

8.1.3.3. *About the players' decisions*

In this research, players' behaviors and decisions were also analyzed. Five main decisions players took were analyzed: getting rid of initial appliances, improving efficiency of initial appliances, including new electric appliances, including energy management devices and including energy production devices. The low efficiency houses group showed more energy reductions by getting rid of initial appliances and by including energy management devices. However, the two groups didn't show any significant difference by improving efficiency of initial appliances, by including new electric appliances and by including energy production devices. The previous results were analyzed by performing statistical tests considering the spread of the different players. As a result, another sign of the rebound effect can be detected, considering that the low efficiency houses group managed to reduce their energy consumption more than the other group by getting rid of appliances and by including energy management devices. Somehow the low efficiency houses group felt the task of reducing their energy consumption more important than the other group.

In addition, this research also discovered that players invested more in energy production devices than in energy management devices, despite of this being not economically profitable. The latter result shows irrational behavior of players in which not all the options or possibilities to reduce energy consumption were analyzed, and their decisions were shaped by their own prior knowledge and interests rather than evaluating the best option. Therefore, it is important that this preference for energy production devices over energy management devices is taken into account, and people receive the knowledge about the usefulness of energy management devices. If people had had the complete information, they may have invested more in the most profitable ones. However, this did not happen due to the lack of information. Therefore, it is important that policymakers take this into consideration when designing policies aiming to reduce energy consumption. People do not have to spend great amounts of money on solar panels to reduce their energy consumption. Instead, they can manage to reduce their consumption in the same amount by including smart meters, movement sensors or better insulation for their windows. Policymakers are advised to inform and educate people about all the possibilities they have to reduce energy consumption, and this may encourage more people towards energy efficiency and energy reduction.

Finally, the influence on players of having informed them beforehand the energy efficiency label of the houses they played with, could not be statistically shown. Therefore, energy policies that inform the energy label of the houses when people rent/buy a property are advised to be reviewed. In the game this influence could not be shown, and it may be that that in real life people are not influenced by these stimuli either. Revisions of the successfulness of these type of policies are advised be carried out by policymakers in order to check their influence on people. Maybe more efficient policies can be designed towards energy reduction using the same investment of letting people know their houses' energy label.

8.2. Limitations

In the following, the most important limitations of this research are discussed.

8.2.1. Modified before/after analysis

It must be said that a typical before/after analysis considers one group of people to be analyzed. The behavior of that group is analyzed before and after the stimulus is applied, in this case energy efficiency improvements. However, in this research this was not the case and two different groups were analyzed: one representing the situation before and another one representing the situation after. This may be a controversial methodological discussion, and if the experiments had been carried out using just one group, the results may have been different than the results this research obtained. As a result, this may be a limitation of this research due to the fact that a non-typical before/after analysis was performed. However, the reason behind this decision is the fact that if just one group had been used, they would have had to receive the energy efficiency improvements in the middle of the game, leaving not enough rounds left of the game to properly analyze the rebound effect. What could have been done is to increase the number of rounds from 10 to 15, for instance, but the game would have lasted too long and it would have been more difficult to gather players and perform the experiments. In addition, players may have lost their interest in the game with longer experiments compromising the quality of the results (the total experiment duration was already long, taking in total 1 hour per player on average). Nevertheless, having two different groups allows us to analyze the rebound effect from the very beginning of the game and without having to increase the duration of the game or the number of rounds. Although this decision constitutes a limitation of this research, the alternative solution would not have been better, reason why it was decided to proceed this way.

8.2.2. Base case scenario methodologies

A second limitation of this research is the way the base case scenario was defined. Two different methodologies were used to calculate this: 1) considering that the high efficiency houses group have less opportunities to improve their energy efficiency since they are already high efficient, and 2) not considering this difference in the opportunities to improve energy efficiency. The two previous methodologies showed similar behaviors, but when the refrigerators were analyzed those two methodologies greatly differed. As a result, we can conclude that the way the base case scenario is calculated is a critical factor for the calculation of the rebound effect. If other methodologies are defined, then new and different results may be obtained. As a result, this research introduced a new numeric definition of the rebound effect, somehow reinforcing one of the causes of controversies around the rebound effect that we have analyzed, having different numeric definitions. As a result, we must point out the necessity that future studies are carried out in line to define a clear definition of the rebound effect in order to reduce the number of definitions that different studies have come up with.

8.2.3. Realism of the game

A third limitation of this research is related to the realism of the game and how players felt this factor. According to what players answered to the question about realism, 14 people (28% of the participants) answered that they are not sure or they (totally) disagree with the statement that the realism of the game is enough for the purpose of it. This is a considerable number of people, therefore the realism of the game is disputable and it may cause some problems in applying the results of the experiments into a real life context. In addition, when players were asked whether they would behave in real life in a similar way as they did in the game, a total of 22 people (44% of the participants) are not sure or (totally) disagree with the statement. This result is high enough to confirm again that the results can hardly be translated into the real world. One alternative to this problem is to have a larger sample size in order to exclude the people who answered this question saying that they would not behave in a similar way in real life. In doing so, the final sample would be formed by just people who would behave in a similar way. In this case this was not possible because the total sample size would have been just 28 people, a small sample size that would have compromised the statistical significance of the conclusions obtained by this research. Future serious games for research purposes should have this in mind from the moment the game is designed, in order to increase the realism of the game as much as possible, and consequently, increase the applicability of the results into a real world context.

8.2.4. Reliability of the rebound effect size

A fourth limitation of this research is the fact that it is hard to rely on the final calculation of the rebound effect, considering that the final number of it highly depends on the methodology used to calculate the base case scenario. We demonstrated that in some cases the rebound effect size highly depends of the chosen methodology. In addition, the rebound effect calculation was carried out with the average values of each of the two groups leaving outside the possibility to statistically analyze the spread of the rebound effect. This had to be performed that way due to the fact that it was not possible to calculate a rebound effect for each player because they played under the same conditions the whole game (modified before/after analysis). Therefore, the rebound effect calculation did not have a measure of spread between participants and it must be calculated by taking just average values into account. Consequently, we can’t really rely on the final calculated size of the rebound effect. However, it must be reminded that obtaining one single and accurate size of the rebound effect is not what this research aimed for. Instead, this research’s objective was to propose a new methodology to assess the rebound effect in order to improve the methodological issues related to previous before/after analysis. Finding unique and accurate rebound effect calculations only distracts the scientific community on important matters (Friedrichsmeier and Matthies, 2015). As a result, not being able to rely on the final calculation of the rebound effect does not compromise the objectives achievement and the academic contribution of this research.

8.2.5. Representativeness of the sample

Due to the fact that this research was carried out basically with students, the sample was not a good representation of the population, which was “adults living in the Netherlands”.

As a result, we can't extrapolate the results obtained from this research to the whole population. If the sample had been a good representation of the population, the experiments would have been too difficult to carry out due to the difficulty to reach people from all the socio-economic groups of the population. Therefore, new studies that use serious games for research purposes to measure the rebound effect should keep this limitation in mind and either try to make the sample a good representation of the population, or simply be as clear as possible about the fact that the results can't be applied to the whole population.

8.2.6. Income influence on players

On section 4.2.4 of this thesis report, the income players would receive after each round of the game was estimated. That net income was the same in both groups of players. However, the fact that one group of players had to pay more for their energy bills than the other group (the low efficiency group), transformed them indirectly into a group of players with less income. As a result, it is not completely clear in this research if the rebound effect that was detected to be present in the results is caused by effectively having a worse energy efficiency or the fact of receiving less income. Therefore, the results may have been influenced by differences in efficiency and differences in the income at the same time, leading to possible bias of the results. This limitation is related as well with the decision of having two different groups of people instead of just one (modified before/after analysis). In a typical before/after analysis this would not have happened since the rebound effect would have been estimated in every single player, and not by comparing different people with different incomes. As a result, it is important that future studies that choose to perform a modified before/after analysis to assess the rebound effect, investigate the influence of the difference in the income received between groups in order to isolate what part of the results are explained by having different efficiencies and what other part of the results are explained by having a different income.

8.3. Recommendations and future research

Different recommendations can be derived from the results of this research. First, serious games can be used by policymakers as a support tool to analyze and ex-ante evaluate the effectiveness of different policies, not only energy policies, but also any other policy that can be simulated by using serious games. In doing so, policymakers may have a new digital laboratory in which they can test policies without having to spend large amounts of money or risking failures in the final implementation of it. This new tool helps not only in testing final policies, but also in detecting sources of failures and possible improvements. Therefore, it is advised that policymakers use this tool and take advantage of its strengths. Second, any future study that uses serious games for research purposes to measure phenomena or test hypotheses should be well aware from the beginning in the possibility that the results from players may not be applicable into a real life context. Therefore, actions to increase this applicability must be taken. Third, the duration of the game should be well analyzed before starting the experiments. Longer games allow researchers to have a richer data set, but at the same time they reduce the number of players willing to participate and they increase the probability that players lose interest in the game becoming noise in the data. Therefore, it must be well balanced the duration of the experiments in order to line it with the objectives of the research. Fourth, if future studies use this type of modified before/after analysis in which one group is the case “before” and another different group is the case “after”, it becomes crucial that the comparability between them is well addressed and confirmed, otherwise it would not be possible to analyze differences in their behavior. In order to do so, key questions must be asked to players in a pre-game questionnaire. Those questions must be well designed towards the purpose to check statistically if the groups can be compared or not. Fifth, in line to what Aydin et al. (2015) and Sorrell (2009) affirmed, any new measurement of the rebound effect must acknowledge the fact that the rebound effect is a unique phenomenon in which different peculiarities may occur in different people. Therefore, policymakers should keep this in mind and do not take for granted that the rebound effect is small or big without analyzing each specific case.

In addition, future research topics can be obtained from this research too. First, it becomes crucial that a new methodology or set of criteria are defined to assess whether the results from serious games for research purposes can be applied into a real life context. In other words, answer the question on how to be sure that players would behave in a similar way as they did in the game, in order to apply the discoveries and conclusions in a real policy making process. Second, it is important too that new studies are carried out to define a set of criteria to take into account when calculating the base case scenarios, or define a clear definition of the rebound effect. Third, according to Aydin et al. (2015), the rebound effect highly depends on the income people receive, being the highest among the low income groups. As a result, a new assessment of the rebound effect using serious games could analyze how income variations among players affect the rebound effect. In addition, and in line with the limitation discussed in section 8.2.6, future studies could also analyze the reasons of the presence of the rebound effect in people: is it because of having a better efficiency or because of having more disposable income?

9. References

ASHTON, T. S. 1966. *The industrial revolution 1760-1830*, In the Hands of a Child.

AYDIN, E., KOK, N. & BROUNEN, D. 2015. Energy Efficiency and Household Behavior: The Rebound Effect in the Residential Sector.

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

BERKHOUT, P. H. G., MUSKENS, J. C. & VELTHUIJSEN, J. W. 2000. Defining the rebound effect. *Energy Policy,* 28**,** 425-432.

BINSWANGER, M. 2001. Technological progress and sustainable development: what about the rebound effect? *Ecological Economics,* 36**,** 119-132.

BOYLE, E., CONNOLLY, T. M. & HAINEY, T. 2011. The role of psychology in understanding the impact of computer games. *Entertainment Computing,* 2**,** 69-74.

BOYLE, E. A., HAINEY, T., CONNOLLY, T. M., GRAY, G., EARP, J., OTT, M., LIM, T., NINAUS, M., RIBEIRO, C. & PEREIRA, J. 2016. An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. *Computers & Education,* 94**,** 178-192.

BRACKEN, P. & SHUBIK, M. 2001. War gaming in the information age: Theory and purpose. *Naval War College Review,* 54**,** 47.

BREUER, J. S. & BENTE, G. 2010. Why so serious? On the relation of serious games and learning. *Eludamos. Journal for Computer Game Culture,* 4**,** 7-24.

BROOKES, L. 2000. Energy efficiency fallacies revisited. *Energy Policy,* 28**,** 355-366.

CAMERER, C. F. & FEHR, E. 2002. Measuring social norms and preferences using experimental games: A guide for social scientists.

CAMPBELL, D. T. & STANLEY, J. C. 2015. *Experimental and quasi-experimental designs for research*, Ravenio Books.

CBS. 2014a. *Income and spending* [Online]. Centraal Bureau voor de Statistiek. Available: http://visualisatie.cbs.nl/en-GB/visualisation/inkomensverdeling [Accessed April 11th 2016].

CBS 2014b. Welvaart in Nederland 2014. Den Haag, The Netherlands: Centraal Bureau voor de Statistiek.

CBS. 2015. *Consumptie naar goederen- en dienstencategorieën; nationale rekeningen* [Online]. Centraal Bureau voor de Statistiek. Available: http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=83069NED&D1=66-71%2c78%2c81-82%2c86-87%2c89%2c92-94&D2=a&HDR=T&STB=G1&VW=T [Accessed April 11th 2016].

CLEANTECHNICA. 2016. *Top Five Micro Wind Turbines* [Online]. Available: http://cleantechnica.com/2008/03/21/the-five-best-micro-wind-turbines/ [Accessed April 7th 2016].

CONNOLLY, T. M., HAINEY, T., BOYLE, E., BAXTER, G. & MORENO-GER, P. 2014. *Psychology, Pedagogy, and Assessment in Serious Games,* Hershey, PA, IGI Global.

COWLEY, B., MOUTINHO, J. L., BATEMAN, C. & OLIVEIRA, A. 2011. Learning principles and interaction design for 'Green My Place': A massively multiplayer serious game. *Entertainment Computing,* 2**,** 103-113.

COX, P. M., BETTS, R. A., JONES, C. D., SPALL, S. A. & TOTTERDELL, I. J. 2000. Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature,* 408**,** 184-187.

DIMITROPOULOS, J. 2007. Energy productivity improvements and the rebound effect: An overview of the state of knowledge. *Energy Policy,* 35**,** 6354-6363.

DOUGHERTY, C. 2008. *Elements of econometrics,* University of London International Programmes in Economics, Management, Finance and the Social Sciences.

DUBIN, J. A., MIEDEMA, A. K. & CHANDRAN, R. V. 1986. Price effects of energy-efficient technologies: a study of residential demand for heating and cooling. *The Rand Journal of Economics***,** 310-325.

DUKE, R. D. & GEURTS, J. 2004. *Policy games for strategic management*, Rozenberg Publishers.

ELMUTI, D. & KATHAWALA, Y. 1997. An overview of benchmarking process: a tool for continuous improvement and competitive advantage. *Benchmarking for Quality Management & Technology,* 4**,** 229-243.

ENERGIECONSULTANT. 2016. *Omrekening van m3 (n) naar kWh* [Online]. Eefde, The Netherlands. Available: http://www.energieconsultant.nl/energiemarkt/energie-berekeningen-uit-de-praktijk/omrekening-van-m3-n-naar-kwh/ [Accessed May 15th 2016].

EUROSTAT. 2015a. *Electricity prices by type of user* [Online]. Available: http://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&language=en&pcode=ten00118&plugin=1 [Accessed April 7th 2016].

EUROSTAT. 2015b. *Gas prices by type of user* [Online]. Available: http://ec.europa.eu/eurostat/tgm/refreshTableAction.do?tab=table&plugin=1&pcode=ten00117&language=en [Accessed April 7th 2016].

FAWCETT, T., LANE, K., BOARDMAN, B., BANKS, N., GRIFFIN, H., LIPP, J., SCHIELLERUP, P., THERIVEL, R., BLOK, K. & VAN BRUMMELEN, M. 2000. *Lower carbon futures for European households*, Environmental Change Institute, University of Oxford Oxford, UK.

FRIEDRICHSMEIER, T. & MATTHIES, E. 2015. Rebound Effects in Energy Efficiency—an Inefficient Debate? *GAIA-Ecological Perspectives for Science and Society,* 24**,** 80-84.

FRIENDLY, D. & KORMYLO, A. 2012. Energy Consumption of Major Household Appliances Shipped in Canada: Trends for 1990–2010. Ottawa, Canada: Natural Resources Canada's Office of Energy Effiency.

FRONDEL, M., RITTER, N. & VANCE, C. 2012. Heterogeneity in the rebound effect: Further evidence for Germany. *Energy Economics,* 34**,** 461-467.

GEELEN, D., BREZET, H., KEYSON, D. & BOESS, S. Gaming for energy conservation in households. Knowledge Collaboration & Learning for Sustainable Innovation: 14th European Roundtable on Sustainable Consumption and Production (ERSCP) conference and the 6th Environmental Management for Sustainable Universities (EMSU) conference, Delft, The Netherlands, October 25-29, 2010, 2010. Delft University of Technology; The Hague University of Applied Sciences; TNO.

GEELEN, D., KEYSON, D., BOESS, S. & BREZET, H. 2012. Exploring the use of a game to stimulate energy saving in households. *Journal of Design Research 14,* 10**,** 102-120.

GERDES, J. 2015. Energy Efficiency Trends and Policies in the Netherlands. ECN, Odyssee Mure.

GERDES, J., MARBUS, S., BOELHOUWER, M., YBEMA, R., SMIT, A. S. & JURJUS, A. 2014. Energietrends 2014. ECN, Energie-Nederland, Netbeheer Nederland.

GILLINGHAM, K., KOTCHEN, M. J., RAPSON, D. S. & WAGNER, G. 2013. Energy policy: The rebound effect is overplayed. *Nature,* 493**,** 475-476.

GREENE, D. 1997. Theory and empirical estimates of the rebound effect for the US transportation sector. *ORNL*.

GREENING, L. A., GREENE, D. L. & DIFIGLIO, C. 2000. Energy efficiency and consumption the rebound effect a survey. *Energy Policy,* 28**,** 389-401.

GREPPERUD, S. & RASMUSSEN, I. 2004. A general equilibrium assessment of rebound effects. *Energy economics,* 26**,** 261-282.

GUSTAFSSON, A., BÅNG, M. & SVAHN, M. Power explorer: a casual game style for encouraging long term behavior change among teenagers. Proceedings of the International Conference on Advances in Computer Enterntainment Technology, 2009. ACM, 182-189.

HAAS, R., AUER, H. & BIERMAYR, P. 1998. The impact of consumer behavior on residential energy demand for space heating. *Energy and buildings,* 27**,** 195-205.

HAAS, R. & BIERMAYR, P. 2000. The rebound effect for space heating Empirical evidence from Austria. *Energy Policy,* 28**,** 403-410.

HAFERKAMP, N., KRAEMER, N. C., LINEHAN, C. & SCHEMBRI, M. 2011. Training disaster communication by means of serious games in virtual environments. *Entertainment Computing,* 2**,** 81-88.

HANSEN, S. 2016. *Heat Pump Price Comparison Guides* [Online]. Available: http://www.heatpumppriceguides.com/#seer-hspf [Accessed April 7th 2016].

HARTEVELD, C. 2011. *Triadic game design: Balancing reality, meaning and play*, Springer Science & Business Media.

HEIJNEN, P. W. 2008. *EPA1311: Research Methods and Data Analysis,* Technology, Policy and Management Faculty, Delft university of Technology, The Netherlands.

HERRING, H. & ROY, R. 2007. Technological innovation, energy efficient design and the rebound effect. *Technovation,* 27**,** 194-203.

HERTWICH, E. G. 2008. Consumption and the Rebound Effect: An Industrial Ecology Perspective. *Journal of Industrial Ecology,* 9**,** 85-98.

HIEMINGA, G. 2013. Saving Energy in the Netherlands. *ING Bank N.V*.

JIN, S.-H. 2007. The effectiveness of energy efficiency improvement in a developing country: Rebound effect of residential electricity use in South Korea. *Energy Policy,* 35**,** 5622-5629.

KATSALIAKI, K. & MUSTAFEE, N. A survey of serious games on sustainable development. Simulation Conference (WSC), Proceedings of the 2012 Winter, 2012. IEEE, 1-13.

KHAZZOOM, J. D. 1986. An econometric model integrating conservation in the estimation of the residential demand for electricity. JAI Press, Greenwich, CT.

KUIT, M. 2002. Strategic behavior and regulatory styles in the Netherlands energy industry.

LIN, B. & LIU, X. 2012. Dilemma between economic development and energy conservation: Energy rebound effect in China. *Energy,* 45, 867-873.

LO, J., VAN DEN HOOGEN, J. & MEIJER, S. Using gaming simulation experiments to test railway innovations: implications for validity. Simulation Conference (WSC), 2013 Winter, 2013. IEEE, 1766-1777.

MADLENER, R. & ALCOTT, B. 2009. Energy rebound and economic growth: A review of the main issues and research needs. *Energy,* 34, 370-376.

MARSH, T. 2011. Serious games continuum: Between games for purpose and experiential environments for purpose. *Entertainment Computing,* 2, 61-68.

MAYER, I., WARMELINK, H. & ZHOU, Q. 2016. The Utility of Games for Society, Business, and Politics. *The Wiley Handbook of Learning Technology*, 406-435.

MEIJER, S. 2009. The organisation of transactions. *Studying supply networks using gaming simulation. Wageningen University*.

MEIJER, S. 2012. Gaming simulations for railways: Lessons learned from modeling six games for the Dutch infrastructure management.

MEIJER, S., HOFSTEDE, G. J., OMTA, S. & BEERS, G. 2008. The organization of transactions: research with the Trust and Tracing game. *Journal on Chain and Network Science,* 8, 1-20.

MEYER, B. D. 1995. Natural and quasi-experiments in economics. *Journal of business & economic statistics,* 13, 151-161.

MICHAEL, D. R. & CHEN, S. L. 2005. *Serious games: Games that educate, train, and inform*, Muska & Lipman/Premier-Trade.

MOHAMMED, I. S., DAALEN, C. E. V. & THISSENA, W. A. H. 2015. NRG: Design of a game to research into and stimulate household investments in energy saving and renewable energy. **Working article**.

OECD/IEA 2014. World Energy Outlook 2014: Executive Summary.

OUYANG, J., LONG, E. & HOKAO, K. 2010. Rebound effect in Chinese household energy efficiency and solution for mitigating it. *Energy,* 35, 5269-5276.

PAPACHRISTOS, G. Residential electricity consumption in the Netherlands: A model-based policy analysis. IST 2014, 5th International Conference on Sustainability Transitions:

Impact and Instututions, Utrecht, The Netherlands, August 27-29; Authors version, 2014.

RITTERFELD, U., CODY, M. & VORDERER, P. 2009. *Serious games: Mechanisms and effects*, Routledge.

ROY, J. 2000. The rebound effect: some empirical evidence from India. *Energy Policy,* 28, 433-438.

SANTARIUS, T. 2014. Der Rebound-Effekt: ein blinder Fleck der sozial-ökologischen GesellschaftstransformationRebound Effects: Blind Spots in the Socio-Ecological Transition of Industrial Societies. *GAIA-Ecological Perspectives for Science and Society,* 23, 109-117.

SAUNDERS, H. D. 2000. A view from the macro side: rebound, backfire, and Khazzoom–Brookes. *Energy policy,* 28, 439-449.

SHAFIEE, S. & TOPAL, E. 2009. When will fossil fuel reserves be diminished? *Energy policy,* 37, 181-189.

SHAFIEE, S. & TOPAL, E. 2010. A long-term view of worldwide fossil fuel prices. *Applied Energy,* 87, 988-1000.

SMALL, K. A. & VAN DENDER, K. 2005. The Effect of Improved Fuel Economy on Vehicle Miles Traveled Estimating the Rebound Effect Using U.S. State Data, 1966-2001. *Policy and Economics*.

SORRELL, S. 2007. The Rebound Effect: an assessment of the evidence for economy-wide energy savings from improved energy efficiency. UK Energy Research Centre London.

SORRELL, S. 2009. Jevons' Paradox revisited: The evidence for backfire from improved energy efficiency. *Energy Policy,* 37, 1456-1469.

SORRELL, S. & DIMITROPOULOS, J. 2008. The rebound effect: Microeconomic definitions, limitations and extensions. *Ecological Economics,* 65, 636-649.

SORRELL, S., DIMITROPOULOS, J. & SOMMERVILLE, M. 2009. Empirical estimates of the direct rebound effect: A review. *Energy Policy,* 37, 1356-1371.

SUSI, T., JOHANNESSON, M. & BACKLUND, P. 2007. Serious games: An overview.

THEECOEXPERTS. 2016. *How Much Do Solar Panels Cost in the UK?* [Online]. Available: http://www.theecoexperts.co.uk/how-much-do-solar-panels-cost-uk#SolarCost [Accessed April 7th 2016].

TURNER, K. 2009. Negative rebound and disinvestment effects in response to an improvement in energy efficiency in the UK economy. *Energy Economics,* 31**,** 648-666.

U.S.DEPARTMENTOFENERGY. 2016. *Estimating Appliance and Home Electronic Energy Use* [Online]. Available: http://energy.gov/energysaver/estimating-appliance-and-home-electronic-energy-use [Accessed April 7th 2016].

UNITEDNATIONS. 2015. *Conference of the Parties: Twenty-first session* [Online]. Available: https://unfccc.int/resource/docs/2015/cop21/eng/l09r01.pdf [Accessed March 27th 2016].

VAN DAALEN, C., SCHAFFERNICHT, M. & MAYER, I. System Dynamics and Serious Games. 32nd International Conference of the System Dynamics Society, Delft, The Netherlands, 20-24 July 2014; Authors version, 2014. System Dynamics Society.

VAN REEKUM, C., JOHNSTONE, T., BANSE, R., ETTER, A., WEHRLE, T. & SCHERER, K. 2004. Psychophysiological responses to appraisal dimensions in a computer game. *Cognition and Emotion,* 18**,** 663-688.

VERSCHUREN, P. & DOOREWAARD, H. 2010. *Designing a Research Project,* The Hague, Eleven International Publishing.

VILOZNI, D., BARAK, A., EFRATI, O., AUGARTEN, A., SPRINGER, C., YAHAV, Y. & BENTUR, L. 2005. The role of computer games in measuring spirometry in healthy and “asthmatic” preschool children. *Chest Journal,* 128**,** 1146-1155.

WANG, H., ZHOU, P. & ZHOU, D. Q. 2012. An empirical study of direct rebound effect for passenger transport in urban China. *Energy Economics,* 34**,** 452-460.

WEI, T. 2010. A general equilibrium view of global rebound effects. *Energy Economics,* 32**,** 661-672.

WINDTURBINES. 2016. *Wind Turbines Ordering / Prices - Ireland* [Online]. Available: http://www.windturbines.ie/domestic/ordering.asp [Accessed April 7th 2016].

WORLDBANK. 2015. *Data Total Population* [Online]. Available: http://data.worldbank.org/indicator/SP.POP.TOTL/countries?display=graph [Accessed March 22nd 2016].

ZÚÑIGA-ARIAS, G., MEIJER, S., RUBEN, R. & JAN HOFSTEDE, G. 2007. Bargaining power and revenue distribution in the Costa Rican mango supply chain: a gaming simulation approach with local producers. *Journal on Chain and Network Science,* 7**,** 143-160.

ZWIKAEL, O. & GONEN, A. 2007. Project execution game (PEG): training towards managing unexpected events. *Journal of European Industrial Training,* 31**,** 495-512.

ZYDA, M. 2005. From visual simulation to virtual reality to games. *Computer,* 38**,** 25-32.

Appendix A: Appliances energy consumption

Gas appliances	**Consumption low efficiency house (m^3)**	**Consumption high efficiency house (m^3)**
Gas stove	22	17
Gas oven	30	23
Shower	358	276
Central heating	1,293	998
TOTAL	**1,703**	**1,314**

Gas multiplier devices[7]	**Multiplier low efficiency house**	**Multiplier high efficiency house**
Gas boiler	0.9996	0.9506
Wall insulation	0.9984	0.9312
Roof insulation	0.9984	0.9312
Glazing insulation	0.9984	0.9312
TOTAL	**0.9948**	**0.7675**

Electric appliances	**Consumption low efficiency house (kWh)**	**Consumption high efficiency house (kWh)**
Refrigerator	1,040	800
TV	615	480
HiFi system	270	220
Washing machine	630	455
Microwave	480	360
Lighting	900	660
Kitchen appliances	640	535
TOTAL	**4,572**	**3,510**

Miscellaneous[8]
Dining table
Dining chair
Sofa
Bookshelf
Twin bed
Toilet

[7] Gas multiplier devices are devices that reduce the overall gas consumption in the house by multiplying the consumption by a factor below 1. The first table in this appendix indicates the final gas consumption of each gas appliances after multiplying its consumption by the gas multiplier factor.

[8] Both types of houses received the same miscellaneous with the same comfort level. Those miscellaneous consume neither electricity nor gas.

Appendix B: Pre-game questionnaire

NRG game pre-questionnaire

Thanks for participating in this research, which is part of a TU Delft master thesis. We will treat all the information received confidentially, and we will not disclose names and answers of our respondents.
Thanks again and let's get started!

*Required

1. **What is your name? (First name and last name)** *

   ....

2. **How old are you?** *

   ....

3. **In what city of the Netherlands do you live?** *

   ....

4. **What is your gender?** *
   Mark only one oval.
   - ◯ Male
   - ◯ Female
   - ◯ I prefer not to say

5. **Where are you from? (name the country)** *

   ....

6. **What is the highest level of education you have completed or you are currently enrolled?** *
   Mark only one oval.
   - ◯ College/High school
   - ◯ Technical degree (MBO)
   - ◯ Bachelor's degree
   - ◯ Master's degree
   - ◯ Doctoral degree
   - ◯ None of the above

7. **Who is the owner of your current dwelling?** *
Mark only one oval.

- Myself
- I am renting from a housing corporation
- I am in a private rental

8. **Are you currently...?** *
Mark only one oval.

- Employed
- Self-employed
- Unemployed
- Student
- Retired

9. **How often do you play video games (such as pc/internet/mobile/consoles)?** *
Mark only one oval.

- Never
- A couple of times per year
- Monthly
- Daily

10. **How would you rate your abilities to play video games?** *
Mark only one oval.

- Poor
- Fair
- Ok
- Good
- Excellent

11. **Who pays for the rent, mortgage or bills in your household?** *
Mark only one oval.

- I pay the whole amount of the rent, mortgage or bills in my household
- I pay part of the whole amount of the rent, mortgage or bills in my household
- Somebody else does it, I don't pay anything

12. **What is the energy label of your home?** *

Mark only one oval.

- A
- B
- C
- D
- E
- F
- G
- I don't know

13. **To what extent do you agree or disagree with the following statements?** *

Mark only one oval per row.

	Totally Disagree	Disagree	I am not sure	Agree	Totally Agree
We need to worry about global warming					
We should take part of the global warming problem and do things by ourselves to reduce it (like recycling, using green energy, reduce car use and so forth)					
I consider myself as an environmental friendly person					
I think the effect of household energy consumption has an important impact on the global warming problem					

14. **How likely it is that you do the following actions?** *

Mark only one oval per row.

	Very unlikely	Unlikely	Neutral likelihood	Likely	Very Likely
Wash the dishes with cold water	⬭	⬭	⬭	⬭	⬭
Turn the lights off if nobody is using them	⬭	⬭	⬭	⬭	⬭
Unplug appliances that are not being used	⬭	⬭	⬭	⬭	⬭
Set the thermostat at a maximum 18°C	⬭	⬭	⬭	⬭	⬭
Use the drying rack instead of the dryer machine to dry clothes	⬭	⬭	⬭	⬭	⬭
Take shorter showers	⬭	⬭	⬭	⬭	⬭
Wash the dishes by hand instead of the dishwasher	⬭	⬭	⬭	⬭	⬭
Install house insulation (roof/walls/floor)	⬭	⬭	⬭	⬭	⬭
Buy energy production devices in the future (solar panels, wind turbines or the like)	⬭	⬭	⬭	⬭	⬭
Buy energy efficient appliances	⬭	⬭	⬭	⬭	⬭
Use the public transportation or take the bike instead of the car	⬭	⬭	⬭	⬭	⬭
Recycle product packaging (plastic/paper/glass/cans)	⬭	⬭	⬭	⬭	⬭
Spend extra money to buy products that do not harm the environment	⬭	⬭	⬭	⬭	⬭

Appendix C: Player's manual

Manual for NRG game players (high energy efficiency houses[9]):

1) The NRG game simulates the energy (electricity and gas) consumption in households. You are going to get a house in which you will play the game. You are asked to play the game as if it was **your own real house (imagine that you just bought that house)**. You will have to make your own choices/decisions regarding energy consumption. **Your purpose is to make it a house in which you are happy to live in** (you define what it makes you happy).
2) The following table shows the inventory of things/appliances that your house already has. This inventory is meant to provide you with the most common things a household has.

Gas things/appliances	Electricity things/appliances	Miscellaneous
Gas stove	Refrigerator	Dining table
Gas oven	TV	Dining chair
Shower	HiFi system	Sofa
Central heating	Washing machine	Bookshelf
Gas boiler	Microwave	Twin bed
Roof insulation	Lighting	Toilet
Glazing insulation	Kitchen appliances	
Wall insulation		

3) The house you will get consumes a total of **3510 kWh of electricity per year** and **1314 m^3 of gas per year**. The latter means that you will pay **702€ per year in electricity** and **1051€ per year in gas**[10]. **The energy efficiency of this house is high (suppose that the house you used to live was less efficient than this one)**. When you start buying/removing/repairing your own things in the game, that consumption will change, and so will your electricity and gas bills.
4) In the game environment you are free to buy, repair, remove or sell anything you want. It is up to you what you do with your house in terms of energy consumption. In order to do that, you will play the game in **10 different rounds (each round is equal to one year)**. You will receive an **income of 3000 € per round (year)**. This money is meant to be used in paying energy bills and buying/repairing appliances in your house (i.e. it excludes the part of the income that you would spend in food, leisure, transportation and so forth).
5) You will start the game in round 1 with 3000€ and after each round you will get another 3000€, but you will have to pay for electricity and gas. **Paying bills and receiving your**

[9] The manual for the players who played with the low efficiency house is the same as this one with minor changes. The only difference is the point number 3 in this manual. Instead of **3510 kWh** must be **4575 kWh**; instead of **1314 m^3** must be **1703 m^3**; instead of **702€ per year in electricity** must be **915€ per year in electricity**; instead of **1051€ per year in gas** must be **1362€ per year in gas**; instead of **the energy efficiency of this house is high (suppose that the house you used to live was less efficient than this one)** must be **the energy efficiency of this house is low.**

[10] In 2015 the average price of electricity and gas was 0,2 €/kWh and 0,8 €/m^3, respectively.

yearly income is done automatically so you don't have to worry about it. You will receive each new round an amount of money that is equal to 3000€ (yearly income) minus the bills. However, you can always check how much you will pay in bills (see screen shots at the end of the manual).

6) All the things you have in your house will **deteriorate** after each round. You can **repair** your things, **sell** them to get some money back, **buy** new things or just **keep using** the deteriorated items. When something deteriorates below 50% you will receive a message saying that the item is broken. However, **that doesn't mean you can no longer use them**, it only means that your lifestyle and house state have decreased. You define if you want to still use a TV that the remote has broken or keep sleeping in a bed that the mattress is not so soft anymore. It is up to you what you do with the deteriorated things. Remember, do whatever that makes you happy.
7) There are just 2 instructions: **1) play the game as you would do it in real life and 2) make a house that makes you happy**. Take your own strategies and play. If you want to buy luxury things, then do it. If you prefer to buy more eco-friendly things, then do it. There are no good or wrong strategies as long as you play as you would do it in real life. **Don't be afraid of clicking around** to discover the game. If you buy something that it doesn't make you happy, then you can just sell it and get the money back. So explore as much as you want.
8) You will play 10 rounds, **approximately 45 min in total**. The first round will be 8 minutes for you to discover and get familiar with the game and after that the duration of each round will be around 4 minutes. If you need more time in some rounds, just ask for it and you will get extra time. If you want to go faster, just ask for it too.
9) After each round has finished, the **next round will start automatically**. You will receive your income and pay the bills automatically as well.
10) In addition to this manual, you have also received some screen shots that help you to find the most important things in the game. Try to figure it out by yourself but if you have any question you can just check the screen shots or ask someone.

Let's get started! Thank you for your cooperation!

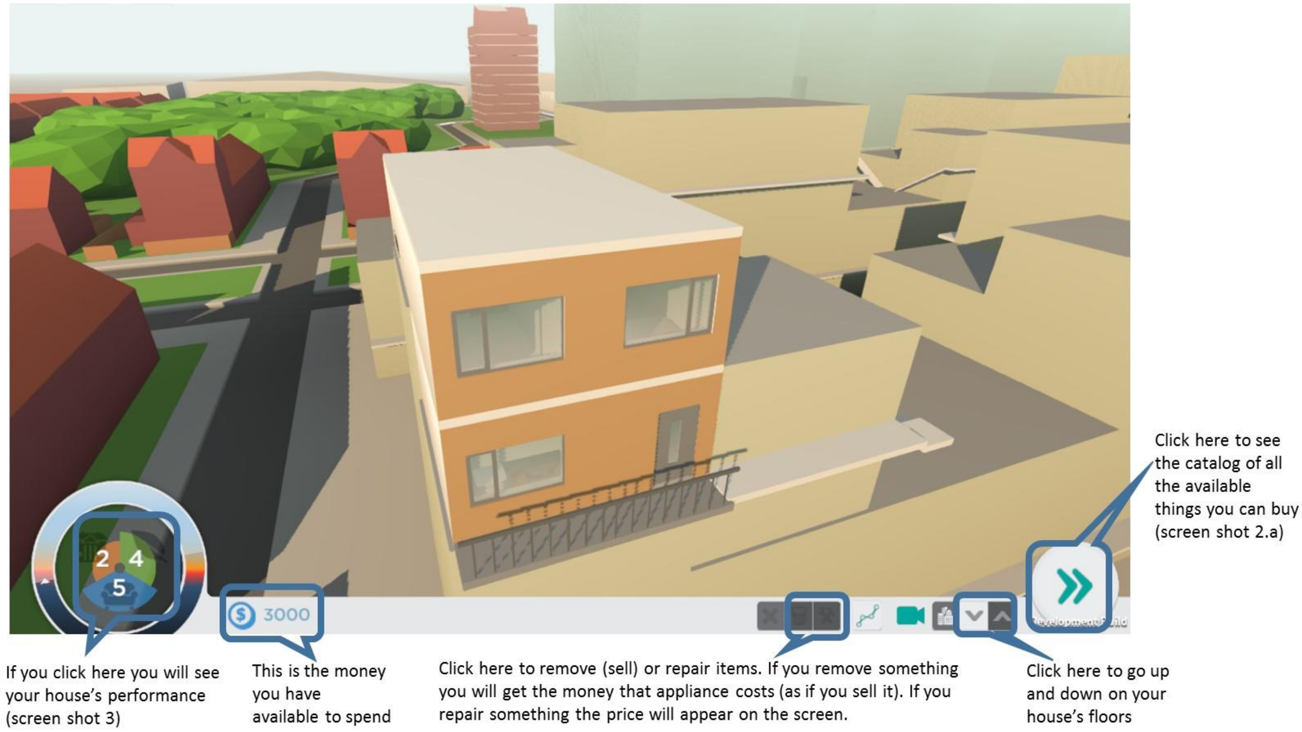


Figure 0.1. NRG game screen shot 1.

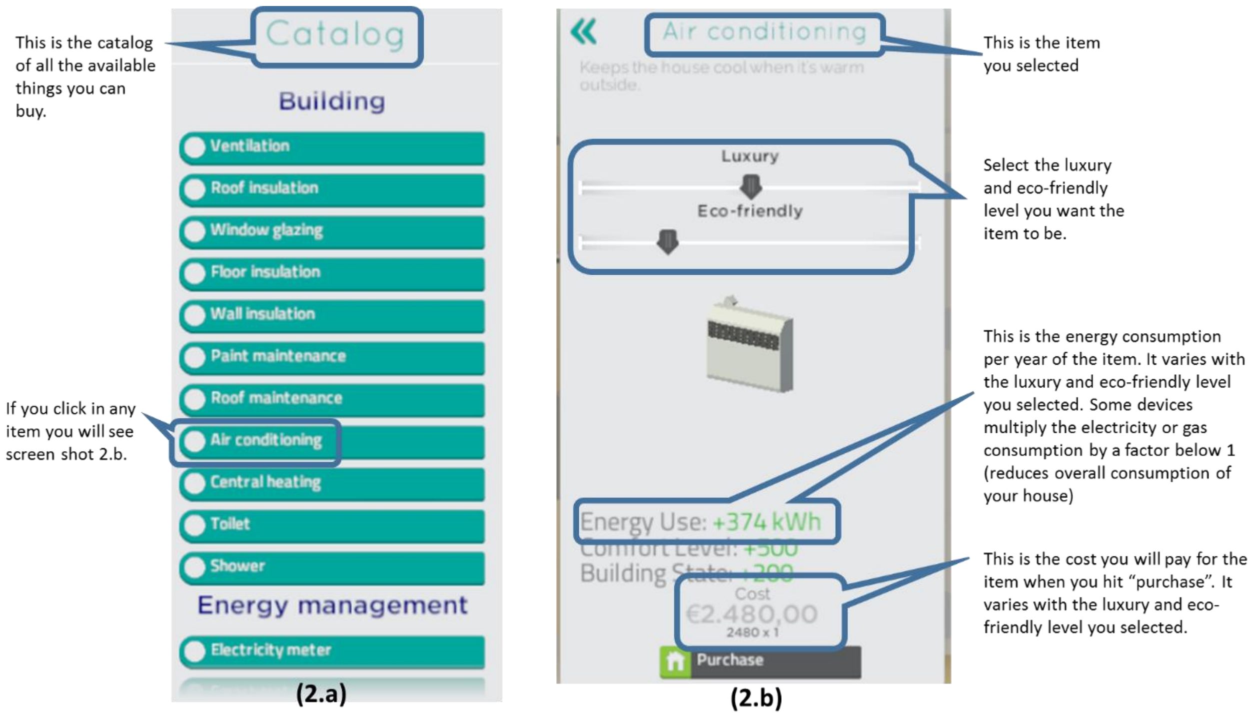


Figure 0.2. NRG game screen shot 2.a and 2.b.

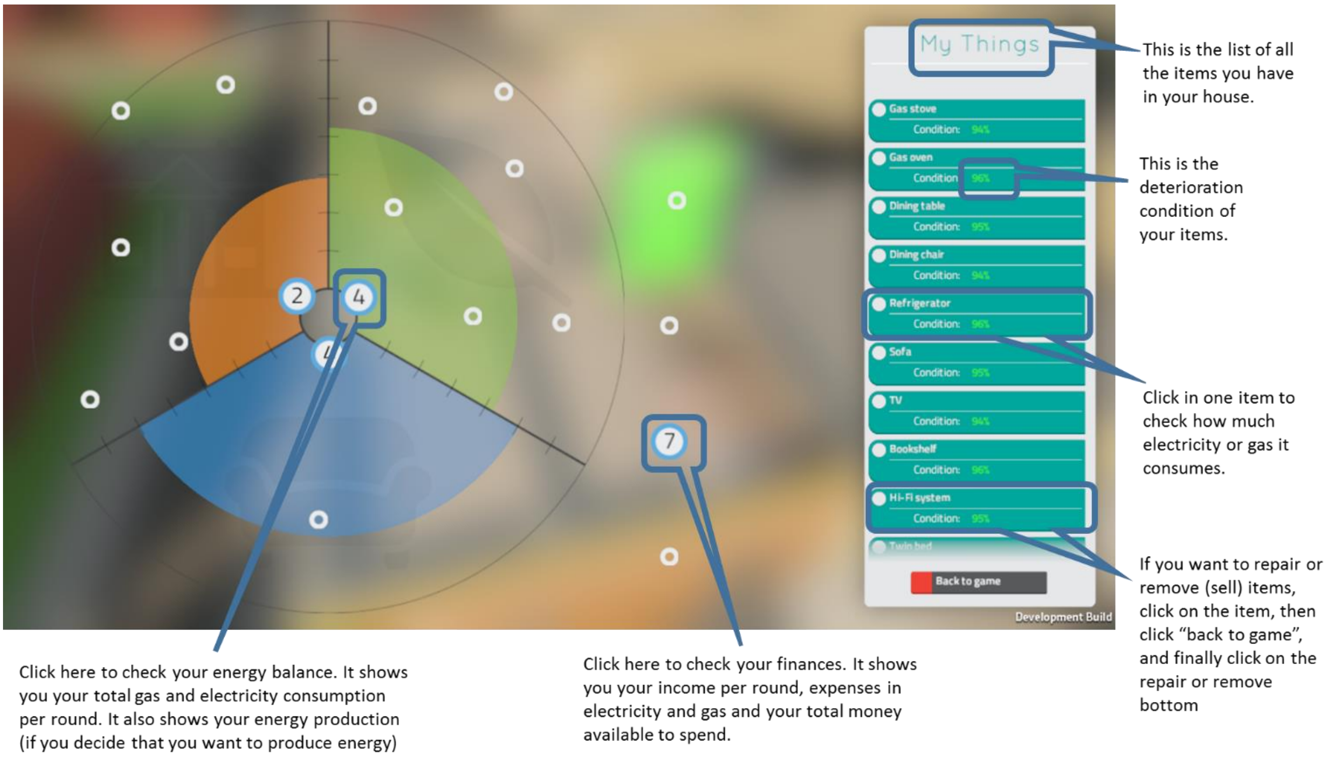


Figure 0.3. NRG game screen shot 3.

Appendix D: Post-game questionnaire

NRG game post-questionnaire

By answering this questionnaire you will finish your participation in this research.
We highly appreciate the time you gave us.
Thank you and lots of success for you!

*Required

1. **What is your name? (First name and last name)** *

....

2. **Game experience: To what extent do you agree or disagree with the following statements?** *

Mark only one oval per row.

	Totally Disagree	Disagree	I am not sure	Agree	Totally Agree
The tasks in the game were understandable and clearly described	⬭	⬭	⬭	⬭	⬭
The rules of the game were clear and straightforward	⬭	⬭	⬭	⬭	⬭
It was easy to find the right options in the game to make my decisions	⬭	⬭	⬭	⬭	⬭
The level of detail in the game is sufficient for the purpose of it	⬭	⬭	⬭	⬭	⬭
The realism of the game is sufficient for the purpose of it	⬭	⬭	⬭	⬭	⬭
The game was built up in an interesting and motivating way	⬭	⬭	⬭	⬭	⬭
The manual for players was helpful in order to understand the instructions better	⬭	⬭	⬭	⬭	⬭
I had fun playing the game	⬭	⬭	⬭	⬭	⬭
I got bored in some parts of the game	⬭	⬭	⬭	⬭	⬭
The duration of the game was too long	⬭	⬭	⬭	⬭	⬭
The number of rounds of the game were too many	⬭	⬭	⬭	⬭	⬭
I think in real life I would behave in a similar way as I did in the game	⬭	⬭	⬭	⬭	⬭

3. **What was the energy efficiency level of the house you play with?** *
Mark only one oval.

- ◯ High efficiency
- ◯ Low efficiency
- ◯ I don't remember *Skip to question 5.*

4. **Do you agree or disagree with the statement:** *
Mark only one oval per row.

	Totally Disagree	Disagree	I am not sure	Agree	Totally Agree
Knowing the energy efficiency level of the house I played with, influenced me in the way I made my decisions in the game	◯	◯	◯	◯	◯

5. **Game strategies and decisions: To what extent do you agree or disagree with the following statements?** *
Mark only one oval per row.

	Totally Disagree	Disagree	I am not sure	Agree	Totally Agree
I often checked my gas and electricity consumption	◯	◯	◯	◯	◯
I preferred to buy luxury rather than eco-friendly things	◯	◯	◯	◯	◯
I wish I had more money in the game to buy more luxurious things	◯	◯	◯	◯	◯
I wish I had more money in the game to buy more eco-friendly things	◯	◯	◯	◯	◯
I wish I had more money in the game to buy more production devices (like solar panels/wind turbines)	◯	◯	◯	◯	◯

6. **Did you invest money in energy management devices (stand-by killer/movement sensors/smart meters) in the game? (energy management devices do NOT include energy production devices like solar panels or wind turbines. There will be a separate question about it)** *
Mark only one oval.

- ◯ Yes
- ◯ No *Skip to question 9.*

7. **The reason why I invested money in energy management devices (stand-by killer/movement sensors/smart meters) in the game is because:** *

Mark only one oval per row.

	Totally Disagree	Disagree	I am not sure	Agree	Totally Agree
I wanted to reduce my costs in the game	⬭	⬭	⬭	⬭	⬭
My energy consumption was too high and I wanted to reduce it	⬭	⬭	⬭	⬭	⬭
Energy management devices were cheaper than other things to reduce consumption	⬭	⬭	⬭	⬭	⬭
I was curious about how much money I could save by buying energy management devices	⬭	⬭	⬭	⬭	⬭
I care about the environment and I wanted to reduce my environmental impact	⬭	⬭	⬭	⬭	⬭
No particular reason, I just felt like trying it out	⬭	⬭	⬭	⬭	⬭

8. **Were there any other reasons for investing in energy management devices (stand-by killer/movement sensors/smart meters) in the game? If so, please describe it shortly**

....

....

....

....

....

Skip to question 11.

9. **The reason why I didn't invest money in energy management devices (stand-by killer/movement sensors/smart meters) in the game is because:** *

Mark only one oval per row.

	Totally Disagree	Disagree	I am not sure	Agree	Totally Agree
I preferred to spend my money in more appliances for my house	◯	◯	◯	◯	◯
The energy savings were too small compared with the investment (it wasn't economically profitable)	◯	◯	◯	◯	◯
I didn't have any money left	◯	◯	◯	◯	◯
The cost of energy management devices were too high for me to afford them	◯	◯	◯	◯	◯
I don't think energy management devices really reduce energy consumption	◯	◯	◯	◯	◯
I don't think my small contribution is worth to reduce global warming	◯	◯	◯	◯	◯
I didn't know energy management devices were available to buy in the game	◯	◯	◯	◯	◯

10. **Were there any other reasons for not investing in energy management devices (stand-by killer/movement sensors/smart meters) in the game? If so, please describe it shortly**

....

....

....

....

....

11. **Did you invest money in energy production devices (solar panels/wind turbines/heat pumps/micro co-generation) in the game? (energy production devices do NOT include energy management devices. You already answered a question about it)** *

Mark only one oval.

◯ Yes

◯ No   *Skip to question 14.*

12. **The reason why I invested money in energy production devices (solar panels/wind turbines/heat pumps/micro co-generation) in the game is because:** *

Mark only one oval per row.

	Totally Disagree	Disagree	I am not sure	Agree	Totally Agree
I wanted to reduce my costs in the game	◯	◯	◯	◯	◯
My energy consumption was too high and I wanted to reduce it	◯	◯	◯	◯	◯
Energy production devices were cheaper than other things to reduce consumption	◯	◯	◯	◯	◯
I was curious about how much money I could save by buying energy production devices	◯	◯	◯	◯	◯
I care about the environment and I wanted to reduce my environmental impact	◯	◯	◯	◯	◯
No particular reason, I just felt like trying it out	◯	◯	◯	◯	◯

13. **Were there any other reasons for investing in energy production devices (solar panels/wind turbines/heat pumps/micro co-generation) in the game? If so, please describe it shortly**

....

....

....

....

....

Stop filling out this form.

14. **The reason why I didn't invest money in energy production devices (solar panels/wind turbines/heat pumps/micro co-generation) in the game is because:** *

Mark only one oval per row.

	Totally Disagree	Disagree	I am not sure	Agree	Totally Agree
I preferred to spend my money in more appliances for my house	⬭	⬭	⬭	⬭	⬭
The energy savings were too small compared with the investment (it wasn't economically profitable)	⬭	⬭	⬭	⬭	⬭
I didn't have any money left	⬭	⬭	⬭	⬭	⬭
The cost of energy production devices were too high for me to afford them	⬭	⬭	⬭	⬭	⬭
I don't think energy production devices really reduce energy consumption	⬭	⬭	⬭	⬭	⬭
I don't think my small contribution is worth to reduce global warming	⬭	⬭	⬭	⬭	⬭
I didn't know energy production devices were available to buy in the game	⬭	⬭	⬭	⬭	⬭

15. **Were there any other reasons for not investing in energy production devices (solar panels/wind turbines/heat pumps/micro co-generation) in the game? If so, please describe it shortly**

....

....

....

....

....

Appendix E: Interviews NRG game trial session

Interviewee 1: L.F.

1) Answer the following questions in terms of how your felt the about the duration of each round. The time assigned for round X was:

   Round 1:
   Too little – Little – **Enough** – Much – Too much
   Round 2:
   Too little – Little – **Enough** – Much – Too much
   Round 3:
   Too little – Little – **Enough** – Much – Too much
   Round 4:
   Too little – Little – Enough – **Much** – Too much
   Round 5:
   Too little – Little – **Enough** – Much – Too much
   Round 6:
   Too little – Little – **Enough** – Much – Too much
   Round 7:
   Too little – Little – **Enough** – Much – Too much
   Round 8:
   Too little – Little – **Enough** – Much – Too much
   Round 9:
   Too little – Little – Enough – Much – Too much
   Round 10:
   Too little – Little – Enough – Much – Too much

2) How did you feel about the number of rounds in the game?

   Too little – Little – **Enough** – Much – Too much

3) How did you feel about total time needed to play the game?

   Too little – Little – Enough – Much – Too much

4) Did you get bored at some point during the game? In what round? Why?
   No.

5) Do you think the time given affected your performance in the game?
   Not the time, the money affects your performance.

6) Did you understand the instructions of the game? If not, why?
   Yes.

7) Did you understand how the game works? If not, why?
   Yes.

8) Do you think the manual helped you to understand how the game works and what the instructions were? If not, why?
   Yes.

9) What do you think it can be improved/changed in the way the game is carried out?
   There is too much money available.

10) Did you check constantly the consumption and the state of your appliances?
    Yes, in each round.

11) Did you lose interest in the game throughout the game? Why?
    No.

12) In which round do you think you were familiar with the game functioning?
    Round 3.

13) Did you have fun playing the game?
    Too little – Little – **OK** – Much – Too much

14) If you had to rate the overall game experience in a scale from 0 (lowest) to 10 (highest), what would it be?
    9

Interviewee 2: S.T.

1) Answer the following questions in terms of how your felt the about the duration of each round. The time assigned for round X was:

   Round 1:
   Too little – Little – **Enough** – Much – Too much
   Round 2:
   Too little – Little – Enough – **Much** – Too much
   Round 3:
   Too little – Little – Enough – **Much** – Too much
   Round 4:
   Too little – Little – Enough – Much – **Too much**
   Round 5:
   Too little – Little – Enough – Much – **Too much**
   Round 6:
   Too little – Little – **Enough** – Much – Too much
   Round 7:
   Too little – Little – **Enough** – Much – Too much
   Round 8:
   Too little – Little – **Enough** – Much – Too much
   Round 9:
   Too little – Little – **Enough** – Much – Too much
   Round 10:
   Too little – Little – **Enough** – Much – Too much

2) How did you feel about the number of rounds in the game?

   Too little – Little – **Enough** – Much – Too much

3) How did you feel about total time needed to play the game?

   Too little – Little – **Enough** – Much – Too much

4) Did you get bored at some point during the game? In what round? Why?
   I got bored when I had to wait some rounds to get money to buy solar panels (round 4)

5) Do you think the time given affected your performance in the game?
   No.

6) Did you understand the instructions of the game? If not, why?
   Yes.

7) Did you understand how the game works? If not, why?
   Yes.

8) Do you think the manual helped you to understand how the game works and what the instructions were? If not, why?
   Yes, but the screen shots were not very helpful because you discover the same things clicking around by yourself.

9) What do you think it can be improved/changed in the way the game is carried out?
   I didn't have a real purpose or goal to play. I would recommend to explicitly say in the manual what is the purpose of the players, like "build the house you would like to live" or something like that.

10) Did you check constantly the consumption and the state of your appliances?
    The consumption maybe once or twice. The state just when my things broke down.

11) Did you lose interest in the game throughout the game? Why?
    A little bit when I had to wait to buy the solar panels (Round 4).

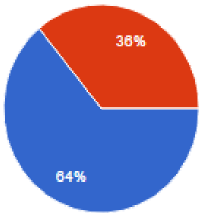
12) In which round do you think you were familiar with the game functioning?
    Round 2.

13) Did you have fun playing the game?
    Too little – Little – **OK** – Much – Too much

14) If you had to rate the overall game experience in a scale from 0 (lowest) to 10 (highest), what would it be?
    7

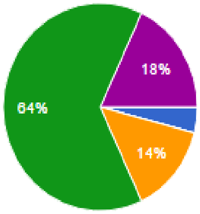
Appendix F: Pre-game questionnaire answers

What is your gender?



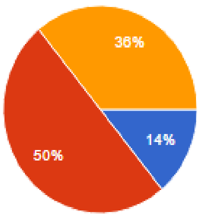
Male	32	64%
Female	18	36%
I prefer not to say	0	0%

What is the highest level of education you have completed or you are currently enrolled?



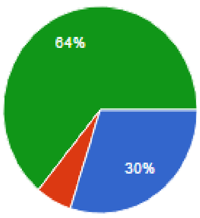
College/High school	2	4%
Technical degree (MBO)	0	0%
Bachelor's degree	7	14%
Master's degree	32	64%
Doctoral degree	9	18%
None of the above	0	0%

Who is the owner of your current dwelling?



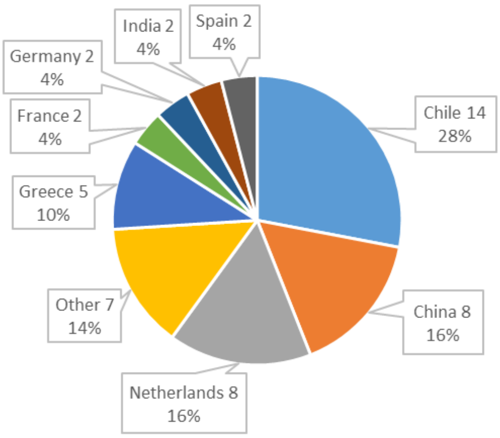
Myself	7	14%
I am renting from a housing corporation	25	50%
I am in a private rental	18	36%

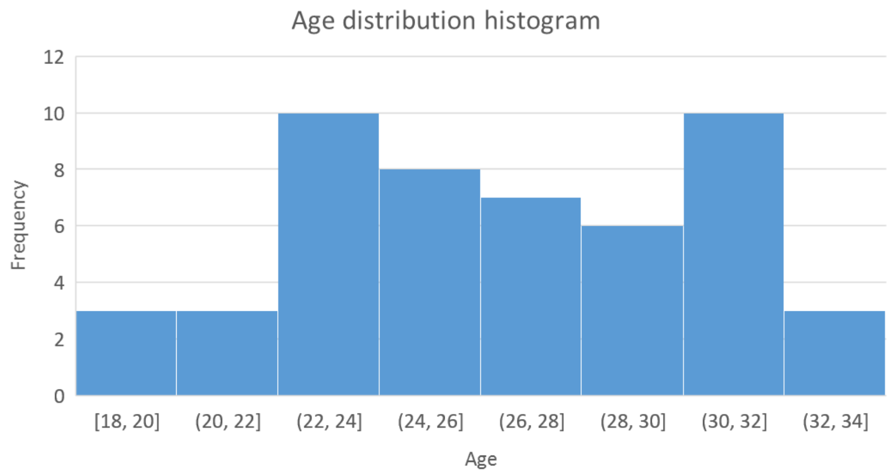
Are you currently...?

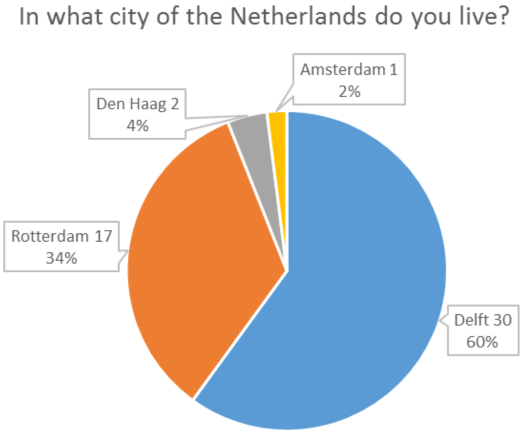


Employed	15	30%
Self-employed	3	6%
Unemployed	0	0%
Student	32	64%
Retired	0	0%

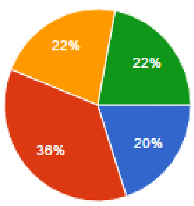
Nationality distribution





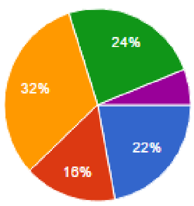


How often do you play video games (such as pc/internet/mobile/consoles)?



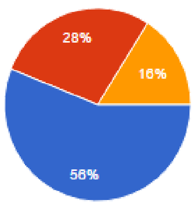
Never	10	20%
A couple of times per year	18	36%
Monthly	11	22%
Daily	11	22%

How would you rate your abilities to play video games?



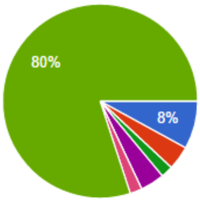
Poor	11	22%
Fair	8	16%
Ok	16	32%
Good	12	24%
Excellent	3	6%

Who pays for the rent, mortgage or bills in your household?



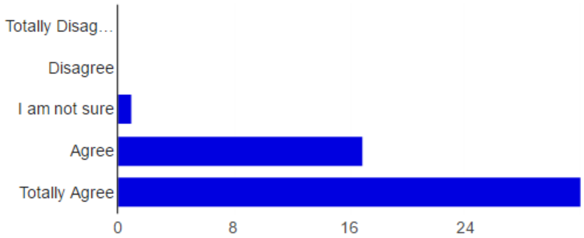
I pay the whole amount of the rent, mortgage or bills in my household	28	56%
I pay part of the whole amount of the rent, mortgage or bills in my household	14	28%
Somebody else does it, I don't pay anything	8	16%

What is the energy label of your home?



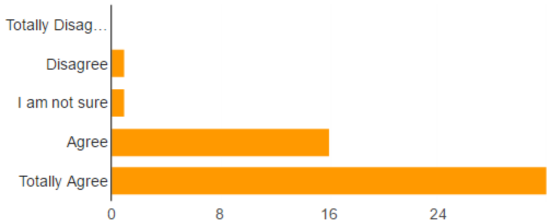
A	4	8%
B	2	4%
C	0	0%
D	1	2%
E	2	4%
F	0	0%
G	1	2%
I don't know	40	80%

We need to worry about global warming [To what extent do you agree or disagree with the following statements?]



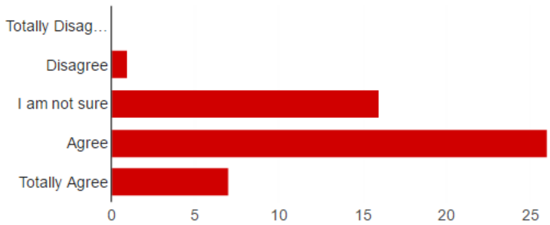
Totally Disagree	0	0%
Disagree	0	0%
I am not sure	1	2%
Agree	17	34%
Totally Agree	32	64%

We should take part of the global warming problem and do things by ourselves to reduce it (like recycling, using green energy, reduce car use and so forth) [To what extent do you agree or disagree with the following statements?]



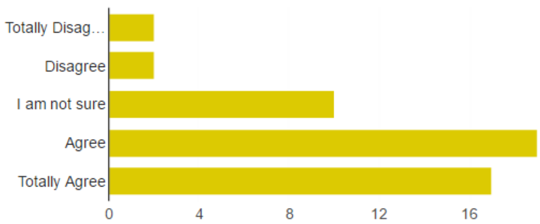
Totally Disagree	0	0%
Disagree	1	2%
I am not sure	1	2%
Agree	16	32%
Totally Agree	32	64%

I consider myself as an environmental friendly person [To what extent do you agree or disagree with the following statements?]



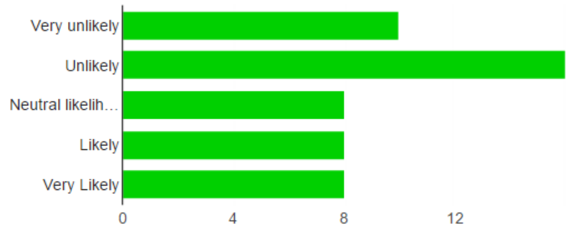
Totally Disagree	0	0%
Disagree	1	2%
I am not sure	16	32%
Agree	26	52%
Totally Agree	7	14%

I think the effect of household energy consumption has an important impact on the global warming problem [To what extent do you agree or disagree with the following statements?]



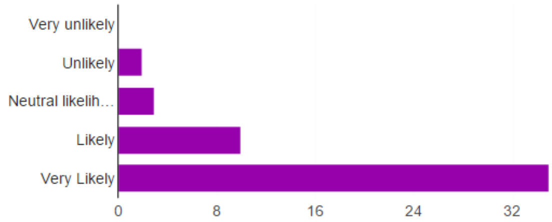
Totally Disagree	2	4%
Disagree	2	4%
I am not sure	10	20%
Agree	19	38%
Totally Agree	17	34%

Wash the dishes with cold water [How likely it is that you do the following actions?]



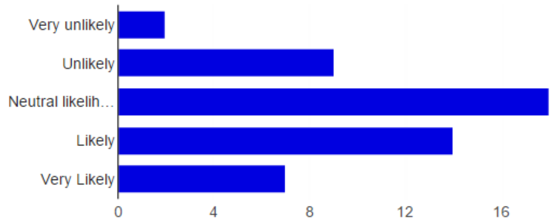
Very unlikely	10	20%
Unlikely	16	32%
Neutral likelihood	8	16%
Likely	8	16%
Very Likely	8	16%

Turn the lights off if nobody is using them [How likely it is that you do the following actions?]



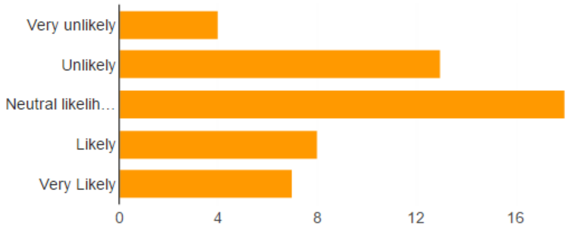
Very unlikely	0	0%
Unlikely	2	4%
Neutral likelihood	3	6%
Likely	10	20%
Very Likely	35	70%

Unplug appliances that are not being used [How likely it is that you do the following actions?]



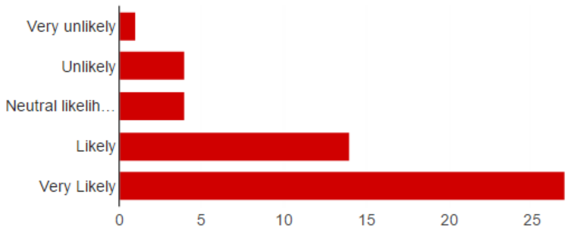
Very unlikely	2	4%
Unlikely	9	18%
Neutral likelihood	18	36%
Likely	14	28%
Very Likely	7	14%

Set the thermostat at a maximum 18°C [How likely it is that you do the following actions?]



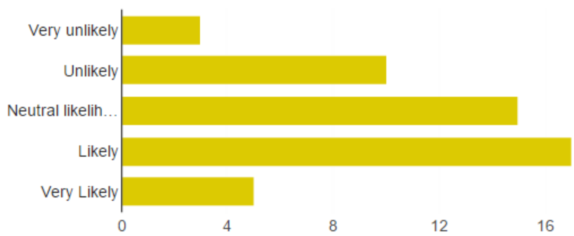
Very unlikely	4	8%
Unlikely	13	26%
Neutral likelihood	18	36%
Likely	8	16%
Very Likely	7	14%

Use the drying rack instead of the dryer machine to dry clothes [How likely it is that you do the following actions?]



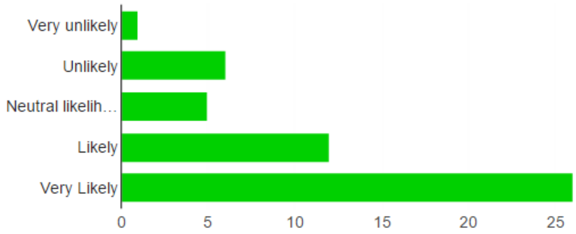
Very unlikely	1	2%
Unlikely	4	8%
Neutral likelihood	4	8%
Likely	14	28%
Very Likely	27	54%

Take shorter showers [How likely it is that you do the following actions?]



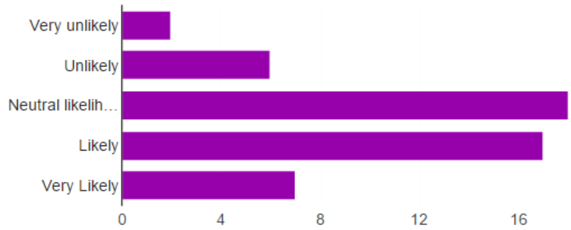
Very unlikely	3	6%
Unlikely	10	20%
Neutral likelihood	15	30%
Likely	17	34%
Very Likely	5	10%

Wash the dishes by hand instead of the dishwasher [How likely it is that you do the following actions?]



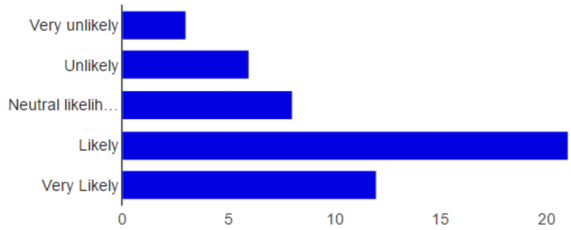
Very unlikely	**1**	2%
Unlikely	**6**	12%
Neutral likelihood	**5**	10%
Likely	**12**	24%
Very Likely	**26**	52%

Install house insulation (roof/walls/floor) [How likely it is that you do the following actions?]



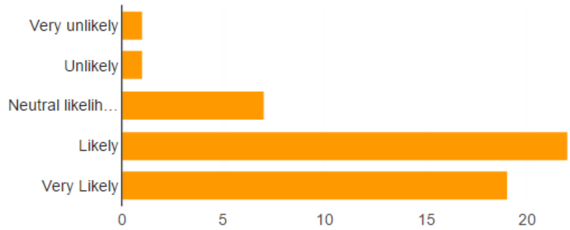
Very unlikely	**2**	4%
Unlikely	**6**	12%
Neutral likelihood	**18**	36%
Likely	**17**	34%
Very Likely	**7**	14%

Buy energy production devices in the future (solar panels, wind turbines or the like) [How likely it is that you do the following actions?]



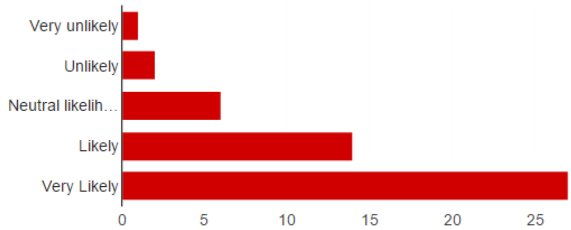
Very unlikely	**3**	6%
Unlikely	**6**	12%
Neutral likelihood	**8**	16%
Likely	**21**	42%
Very Likely	**12**	24%

Buy energy efficient appliances [How likely it is that you do the following actions?]



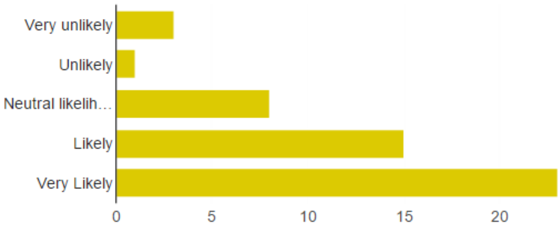
Very unlikely	**1**	2%
Unlikely	**1**	2%
Neutral likelihood	**7**	14%
Likely	**22**	44%
Very Likely	**19**	38%

Use the public transportation or take the bike instead of the car [How likely it is that you do the following actions?]



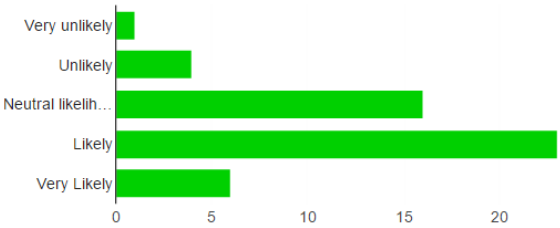
Very unlikely	**1**	2%
Unlikely	**2**	4%
Neutral likelihood	**6**	12%
Likely	**14**	28%
Very Likely	**27**	54%

Recycle product packaging (plastic/paper/glass/cans) [How likely it is that you do the following actions?]



Very unlikely	**3**	6%
Unlikely	**1**	2%
Neutral likelihood	**8**	16%
Likely	**15**	30%
Very Likely	**23**	46%

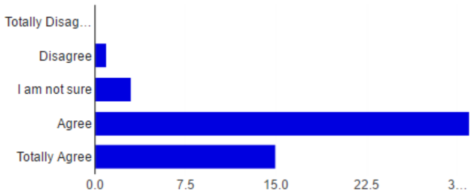
Spend extra money to buy products that do not harm the environment [How likely it is that you do the following actions?]



Very unlikely	**1**	2%
Unlikely	**4**	8%
Neutral likelihood	**16**	32%
Likely	**23**	46%
Very Likely	**6**	12%

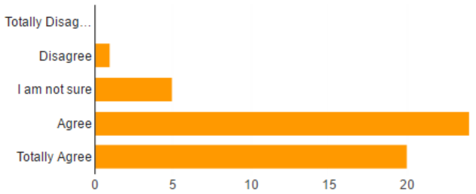
Appendix G: Post-game questionnaire answers

The tasks in the game were understandable and clearly described [Game experience: To what extent do you agree or disagree with the following statements?]



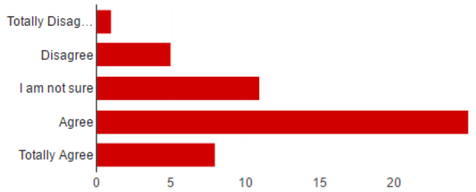
Totally Disagree	0	0%
Disagree	1	2%
I am not sure	3	6%
Agree	31	62%
Totally Agree	15	30%

The rules of the game were clear and straightforward [Game experience: To what extent do you agree or disagree with the following statements?]



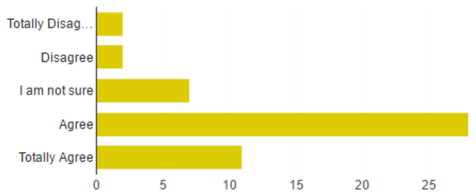
Totally Disagree	0	0%
Disagree	1	2%
I am not sure	5	10%
Agree	24	48%
Totally Agree	20	40%

It was easy to find the right options in the game to make my decisions [Game experience: To what extent do you agree or disagree with the following statements?]



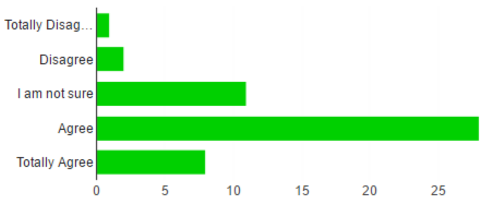
Totally Disagree	1	2%
Disagree	5	10%
I am not sure	11	22%
Agree	25	50%
Totally Agree	8	16%

The level of detail in the game is sufficient for the purpose of it [Game experience: To what extent do you agree or disagree with the following statements?]



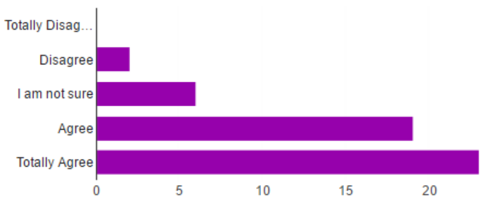
Totally Disagree	2	4%
Disagree	2	4%
I am not sure	7	14%
Agree	28	56%
Totally Agree	11	22%

The realism of the game is sufficient for the purpose of it [Game experience: To what extent do you agree or disagree with the following statements?]



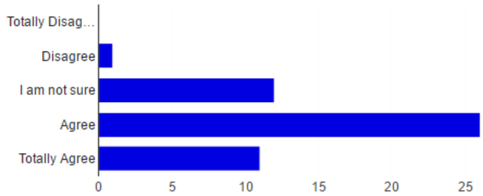
Totally Disagree	1	2%
Disagree	2	4%
I am not sure	11	22%
Agree	28	56%
Totally Agree	8	16%

The game was built up in an interesting and motivating way [Game experience: To what extent do you agree or disagree with the following statements?]



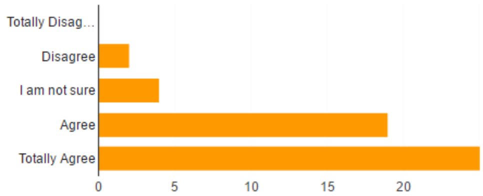
Totally Disagree	0	0%
Disagree	2	4%
I am not sure	6	12%
Agree	19	38%
Totally Agree	23	46%

The manual for players was helpful in order to understand the instructions better [Game experience: To what extent do you agree or disagree with the following statements?]



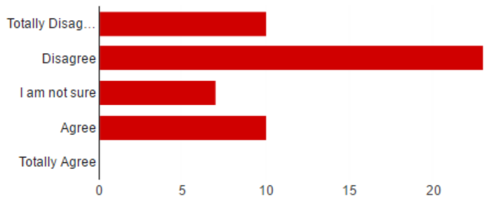
Totally Disagree	**0**	0%
Disagree	**1**	2%
I am not sure	**12**	24%
Agree	**26**	52%
Totally Agree	**11**	22%

I had fun playing the game [Game experience: To what extent do you agree or disagree with the following statements?]



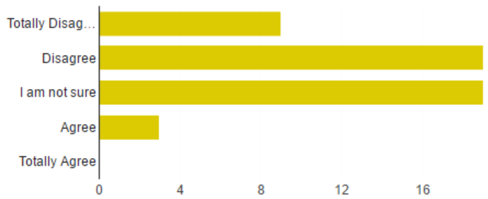
Totally Disagree	**0**	0%
Disagree	**2**	4%
I am not sure	**4**	8%
Agree	**19**	38%
Totally Agree	**25**	50%

I got bored in some parts of the game [Game experience: To what extent do you agree or disagree with the following statements?]



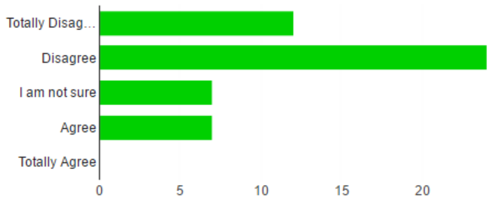
Totally Disagree	**10**	20%
Disagree	**23**	46%
I am not sure	**7**	14%
Agree	**10**	20%
Totally Agree	**0**	0%

The duration of the game was too long [Game experience: To what extent do you agree or disagree with the following statements?]



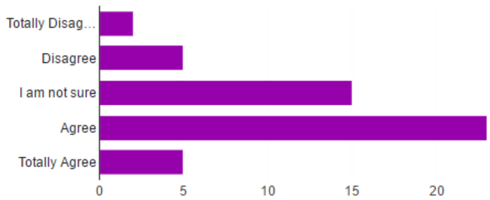
Totally Disagree	**9**	18%
Disagree	**19**	38%
I am not sure	**19**	38%
Agree	**3**	6%
Totally Agree	**0**	0%

The number of rounds of the game were too many [Game experience: To what extent do you agree or disagree with the following statements?]



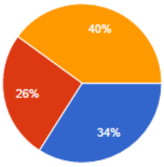
Totally Disagree	**12**	24%
Disagree	**24**	48%
I am not sure	**7**	14%
Agree	**7**	14%
Totally Agree	**0**	0%

I think in real life I would behave in a similar way as I did in the game [Game experience: To what extent do you agree or disagree with the following statements?]



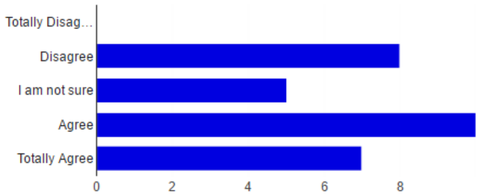
Totally Disagree	**2**	4%
Disagree	**5**	10%
I am not sure	**15**	30%
Agree	**23**	46%
Totally Agree	**5**	10%

What was the energy efficiency level of the house you play with?



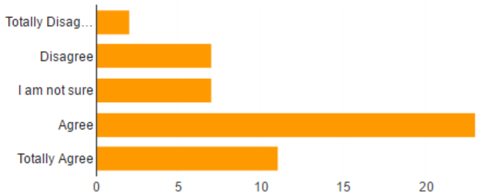
High efficiency	17	34%
Low efficiency	13	26%
I don't remember	20	40%

Knowing the energy efficiency level of the house I played with, influenced me in the way I made my decisions in the game [Do you agree or disagree with the statement:]



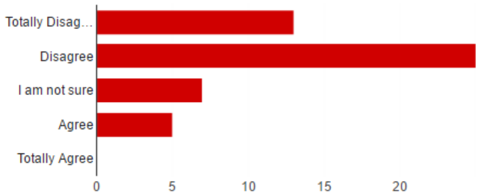
Totally Disagree	0	0%
Disagree	8	26.7%
I am not sure	5	16.7%
Agree	10	33.3%
Totally Agree	7	23.3%

I often checked my gas and electricity consumption [Game strategies and decisions: To what extent do you agree or disagree with the following statements?]



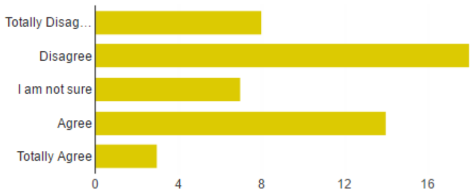
Totally Disagree	2	4%
Disagree	7	14%
I am not sure	7	14%
Agree	23	46%
Totally Agree	11	22%

I preferred to buy luxury rather than eco-friendly things [Game strategies and decisions: To what extent do you agree or disagree with the following statements?]



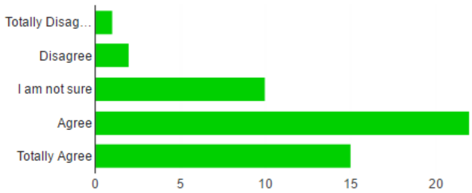
Totally Disagree	13	26%
Disagree	25	50%
I am not sure	7	14%
Agree	5	10%
Totally Agree	0	0%

I wish I had more money in the game to buy more luxurious things [Game strategies and decisions: To what extent do you agree or disagree with the following statements?]



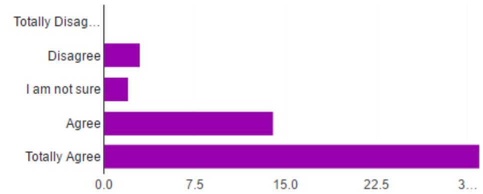
Totally Disagree	8	16%
Disagree	18	36%
I am not sure	7	14%
Agree	14	28%
Totally Agree	3	6%

I wish I had more money in the game to buy more eco-friendly things [Game strategies and decisions: To what extent do you agree or disagree with the following statements?]



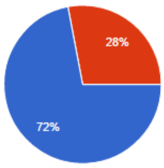
Totally Disagree	1	2%
Disagree	2	4%
I am not sure	10	20%
Agree	22	44%
Totally Agree	15	30%

I wish I had more money in the game to buy more production devices (like solar panels/wind turbines) [Game strategies and decisions: To what extent do you agree or disagree with the following statements?]



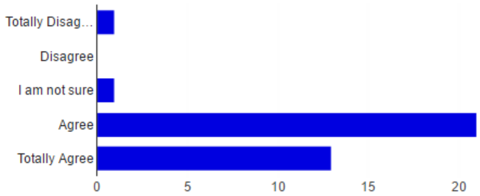
Totally Disagree	0	0%
Disagree	3	6%
I am not sure	2	4%
Agree	14	28%
Totally Agree	31	62%

Did you invest money in energy management devices (stand-by killer/movement sensors/smart meters) in the game? (energy management devices do NOT include energy production devices like solar panels or wind turbines. There will be a separate question about it)



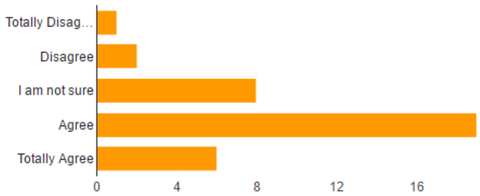
Yes	36	72%
No	14	28%

I wanted to reduce my costs in the game [The reason why I invested money in energy management devices (stand-by killer/movement sensors/smart meters) in the game is because:]



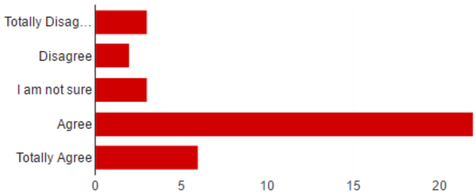
Totally Disagree	1	2.8%
Disagree	0	0%
I am not sure	1	2.8%
Agree	21	58.3%
Totally Agree	13	36.1%

My energy consumption was too high and I wanted to reduce it [The reason why I invested money in energy management devices (stand-by killer/movement sensors/smart meters) in the game is because:]



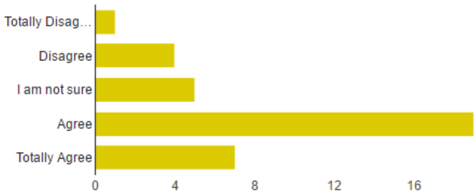
Totally Disagree	1	2.8%
Disagree	2	5.6%
I am not sure	8	22.2%
Agree	19	52.8%
Totally Agree	6	16.7%

Energy management devices were cheaper than other things to reduce consumption [The reason why I invested money in energy management devices (stand-by killer/movement sensors/smart meters) in the game is because:]



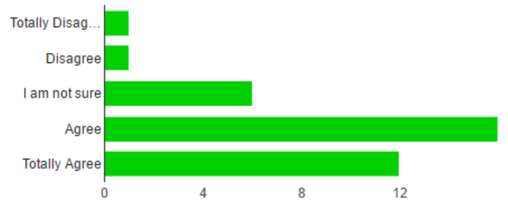
Totally Disagree	3	8.3%
Disagree	2	5.6%
I am not sure	3	8.3%
Agree	22	61.1%
Totally Agree	6	16.7%

I was curious about how much money I could save by buying energy management devices [The reason why I invested money in energy management devices (stand-by killer/movement sensors/smart meters) in the game is because:]



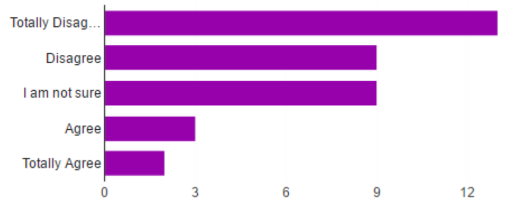
Totally Disagree	1	2.8%
Disagree	4	11.1%
I am not sure	5	13.9%
Agree	19	52.8%
Totally Agree	7	19.4%

I care about the environment and I wanted to reduce my environmental impact [The reason why I invested money in energy management devices (stand-by killer/movement sensors/smart meters) in the game is because:]



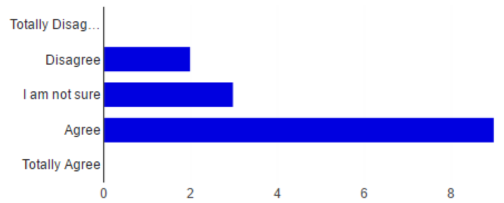
Totally Disagree	1	2.8%
Disagree	1	2.8%
I am not sure	6	16.7%
Agree	16	44.4%
Totally Agree	12	33.3%

No particular reason, I just felt like trying it out [The reason why I invested money in energy management devices (stand-by killer/movement sensors/smart meters) in the game is because:]



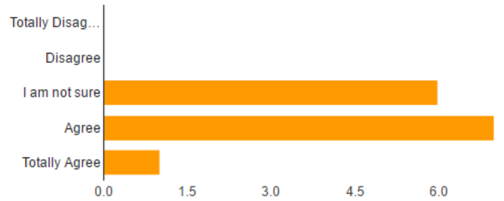
Totally Disagree	13	36.1%
Disagree	9	25%
I am not sure	9	25%
Agree	3	8.3%
Totally Agree	2	5.6%

I preferred to spend my money in more appliances for my house [The reason why I didn't invest money in energy management devices (stand-by killer/movement sensors/smart meters) in the game is because:]



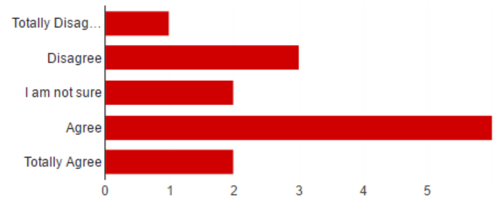
Totally Disagree	0	0%
Disagree	2	14.3%
I am not sure	3	21.4%
Agree	9	64.3%
Totally Agree	0	0%

The energy savings were too small compared with the investment (it wasn't economically profitable) [The reason why I didn't invest money in energy management devices (stand-by killer/movement sensors/smart meters) in the game is because:]



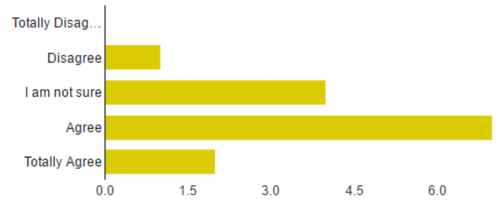
Totally Disagree	0	0%
Disagree	0	0%
I am not sure	6	42.9%
Agree	7	50%
Totally Agree	1	7.1%

I didn't have any money left [The reason why I didn't invest money in energy management devices (stand-by killer/movement sensors/smart meters) in the game is because:]



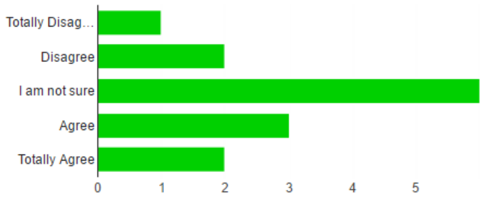
Totally Disagree	1	7.1%
Disagree	3	21.4%
I am not sure	2	14.3%
Agree	6	42.9%
Totally Agree	2	14.3%

The cost of energy management devices were too high for me to afford them [The reason why I didn't invest money in energy management devices (stand-by killer/movement sensors/smart meters) in the game is because:]



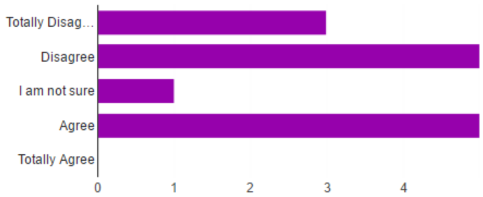
Totally Disagree	0	0%
Disagree	1	7.1%
I am not sure	4	28.6%
Agree	7	50%
Totally Agree	2	14.3%

I don't think energy management devices really reduce energy consumption [The reason why I didn't invest money in energy management devices (stand-by killer/movement sensors/smart meters) in the game is because:]



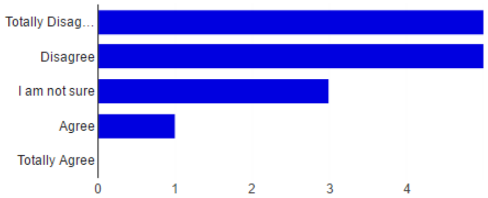
Totally Disagree	1	7.1%
Disagree	2	14.3%
I am not sure	6	42.9%
Agree	3	21.4%
Totally Agree	2	14.3%

I don't think my small contribution is worth to reduce global warming [The reason why I didn't invest money in energy management devices (stand-by killer/movement sensors/smart meters) in the game is because:]



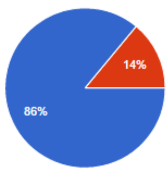
Totally Disagree	3	21.4%
Disagree	5	35.7%
I am not sure	1	7.1%
Agree	5	35.7%
Totally Agree	0	0%

I didn't know energy management devices were available to buy in the game [The reason why I didn't invest money in energy management devices (stand-by killer/movement sensors/smart meters) in the game is because:]



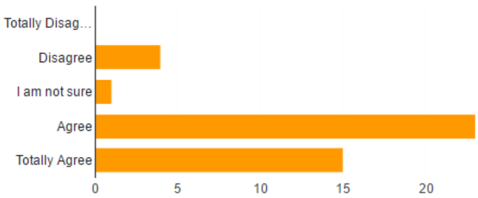
Totally Disagree	5	35.7%
Disagree	5	35.7%
I am not sure	3	21.4%
Agree	1	7.1%
Totally Agree	0	0%

Did you invest money in energy production devices (solar panels/wind turbines/heat pumps/micro co-generation) in the game? (energy production devices do NOT include energy management devices. You already answered a question about it)



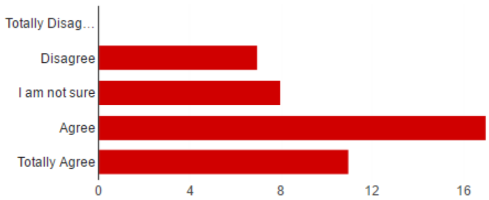
Yes	43	86%
No	7	14%

I wanted to reduce my costs in the game [The reason why I invested money in energy production devices (solar panels/wind turbines/heat pumps/micro co-generation) in the game is because:]



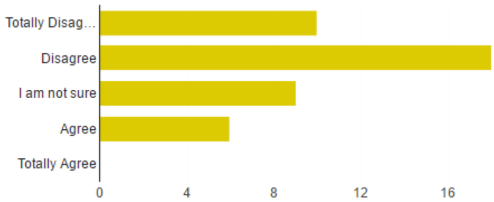
Totally Disagree	0	0%
Disagree	4	9.3%
I am not sure	1	2.3%
Agree	23	53.5%
Totally Agree	15	34.9%

My energy consumption was too high and I wanted to reduce it [The reason why I invested money in energy production devices (solar panels/wind turbines/heat pumps/micro co-generation) in the game is because:]



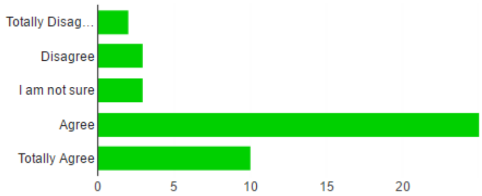
Totally Disagree	0	0%
Disagree	7	16.3%
I am not sure	8	18.6%
Agree	17	39.5%
Totally Agree	11	25.6%

Energy production devices were cheaper than other things to reduce consumption [The reason why I invested money in energy production devices (solar panels/wind turbines/heat pumps/micro co-generation) in the game is because:]



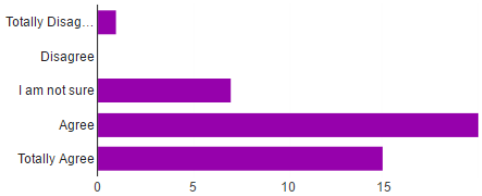
Totally Disagree	10	23.3%
Disagree	18	41.9%
I am not sure	9	20.9%
Agree	6	14%
Totally Agree	0	0%

I was curious about how much money I could save by buying energy production devices [The reason why I invested money in energy production devices (solar panels/wind turbines/heat pumps/micro co-generation) in the game is because:]



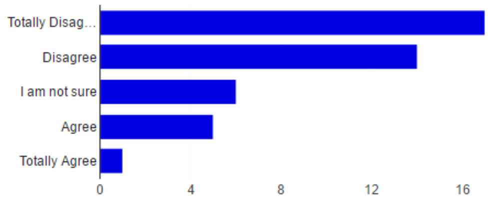
Totally Disagree	2	4.7%
Disagree	3	7%
I am not sure	3	7%
Agree	25	58.1%
Totally Agree	10	23.3%

I care about the environment and I wanted to reduce my environmental impact [The reason why I invested money in energy production devices (solar panels/wind turbines/heat pumps/micro co-generation) in the game is because:]



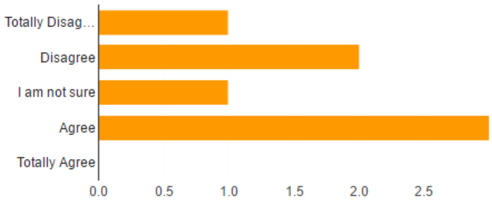
Totally Disagree	1	2.3%
Disagree	0	0%
I am not sure	7	16.3%
Agree	20	46.5%
Totally Agree	15	34.9%

No particular reason, I just felt like trying it out [The reason why I invested money in energy production devices (solar panels/wind turbines/heat pumps/micro co-generation) in the game is because:]



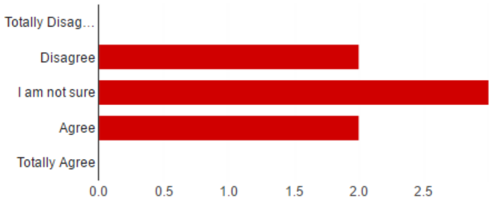
Totally Disagree	17	39.5%
Disagree	14	32.6%
I am not sure	6	14%
Agree	5	11.6%
Totally Agree	1	2.3%

I preferred to spend my money in more appliances for my house [The reason why I didn't invest money in energy production devices (solar panels/wind turbines/heat pumps/micro co-generation) in the game is because:]



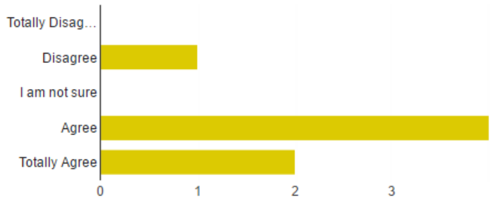
Totally Disagree	1	14.3%
Disagree	2	28.6%
I am not sure	1	14.3%
Agree	3	42.9%
Totally Agree	0	0%

The energy savings were too small compared with the investment (it wasn't economically profitable) [The reason why I didn't invest money in energy production devices (solar panels/wind turbines/heat pumps/micro co-generation) in the game is because:]



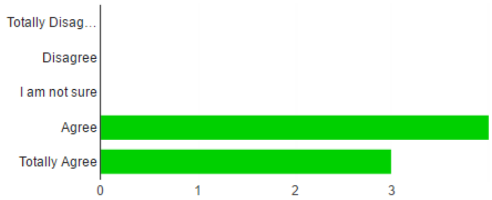
Totally Disagree	0	0%
Disagree	2	28.6%
I am not sure	3	42.9%
Agree	2	28.6%
Totally Agree	0	0%

I didn't have any money left [The reason why I didn't invest money in energy production devices (solar panels/wind turbines/heat pumps/micro co-generation) in the game is because:]



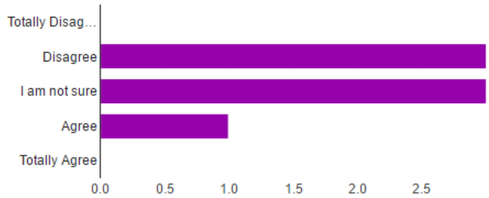
Totally Disagree	0	0%
Disagree	1	14.3%
I am not sure	0	0%
Agree	4	57.1%
Totally Agree	2	28.6%

The cost of energy production devices were too high for me to afford them [The reason why I didn't invest money in energy production devices (solar panels/wind turbines/heat pumps/micro co-generation) in the game is because:]



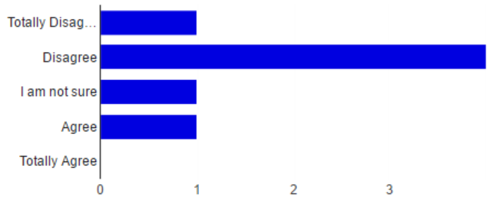
Totally Disagree	0	0%
Disagree	0	0%
I am not sure	0	0%
Agree	4	57.1%
Totally Agree	3	42.9%

I don't think energy production devices really reduce energy consumption [The reason why I didn't invest money in energy production devices (solar panels/wind turbines/heat pumps/micro co-generation) in the game is because:]



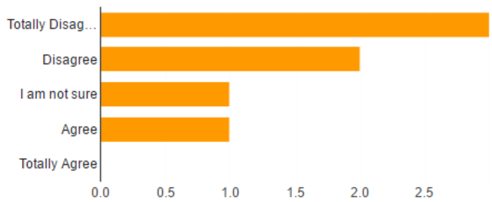
Totally Disagree	0	0%
Disagree	3	42.9%
I am not sure	3	42.9%
Agree	1	14.3%
Totally Agree	0	0%

I don't think my small contribution is worth to reduce global warming [The reason why I didn't invest money in energy production devices (solar panels/wind turbines/heat pumps/micro co-generation) in the game is because:]



Totally Disagree	1	14.3%
Disagree	4	57.1%
I am not sure	1	14.3%
Agree	1	14.3%
Totally Agree	0	0%

I didn't know energy production devices were available to buy in the game [The reason why I didn't invest money in energy production devices (solar panels/wind turbines/heat pumps/micro co-generation) in the game is because:]



Totally Disagree	3	42.9%
Disagree	2	28.6%
I am not sure	1	14.3%
Agree	1	14.3%
Totally Agree	0	0%

Appendix H: SPSS outcomes

H.1:

Kolmogorov-Smirnov test

H_0: The environmental friendliness variable follows the normal theoretical distribution in both groups.

H_1: The environmental friendliness variable does not follow the normal theoretical distribution in both groups.

One-Sample Kolmogorov-Smirnov Test

		Environmental _friendliness
N		25
Normal Parameters[a,b]	Mean	,7624
	Std. Deviation	,36676
Most Extreme Differences	Absolute	,148
	Positive	,108
	Negative	-,148
Test Statistic		,148
Asymp. Sig. (2-tailed)		,163[c]

a. Test distribution is Normal.
b. Calculated from data.
c. Lilliefors Significance Correction.

One-Sample Kolmogorov-Smirnov Test

		Environmental _friendliness
N		25
Normal Parameters[a,b]	Mean	,8918
	Std. Deviation	,48110
Most Extreme Differences	Absolute	,149
	Positive	,089
	Negative	-,149
Test Statistic		,149
Asymp. Sig. (2-tailed)		,157[c]

a. Test distribution is Normal.
b. Calculated from data.
c. Lilliefors Significance Correction.

On the left the tests results are shown for the high efficiency houses group and on the right the results of the low efficiency houses group are shown. The significance values are 16.3% and 15.7% respectively. Therefore, H_0 can't be rejected. The environmental friendliness variable in both groups is assumed to follow the normal theoretical distribution.

Student t-test for differences in averages

H_0: There is no difference between the two groups in terms of the environmental friendliness average.

H_1: There is a difference between the two groups in terms of the environmental friendliness average.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Environmental_friendliness	Equal variances assumed	2,031	,161	1,070	48	,290	,12941	,12099	-,11385	,37268
	Equal variances not assumed			1,070	44,853	,291	,12941	,12099	-,11430	,37312

The previous table shows the result of the test in SPSS. The significance value of the Levene's test is greater than 5%, therefore equal variances are assumed. The significance value of the student t-test is equal to 29%, leading to the conclusion that H_0 can't be rejected. As a result, we will assume that there is no difference between the two groups in terms of the environmental friendliness average.

H.2:

Chi-square test:

H_0: There is no difference between the two groups in terms of their abilities to play video games.

H_1: There is a difference between the two groups in terms of their abilities to play video games.

Abilities_to_play_video_games * Efficiency Crosstabulation

			Efficiency		
			Low	High	Total
Abilities_to_play_video_games	Bad	Count	9	10	19
		Expected Count	9,5	9,5	19,0
	Ok	Count	7	9	16
		Expected Count	8,0	8,0	16,0
	Good	Count	9	6	15
		Expected Count	7,5	7,5	15,0
Total		Count	25	25	50
		Expected Count	25,0	25,0	50,0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	,903[a]	2	,637
Likelihood Ratio	,907	2	,635
N of Valid Cases	50		

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 7,50.

The conditions for the chi-square test are satisfied. There are no expected counts less than 5. The significance value is equal to 63.7%. As a result, H_0 can't be rejected. We will assume that there is no difference between the two groups in terms of their abilities to play video games. Both groups can be compared in terms of their abilities to play video game.

H.3:

Chi-square test:

H_0: There is no difference between the two groups in terms of how often they play video games.

H_1: There is a difference between the two groups in terms of how often they play video games.

How_often_do_you_play_video_games_(such_as_pc/internet/mobile/consoles)? * Efficiency Crosstabulation

			Efficiency		Total
			Low	High	
How_often_do_you_play_video_games_(such_as_pc/internet/mobile/consoles)?	Never	Count	6	4	10
		Expected Count	5,0	5,0	10,0
	A couple of times per year	Count	8	10	18
		Expected Count	9,0	9,0	18,0
	Monthly	Count	7	4	11
		Expected Count	5,5	5,5	11,0
	Daily	Count	4	7	11
		Expected Count	5,5	5,5	11,0
Total		Count	25	25	50
		Expected Count	25,0	25,0	50,0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	2,259[a]	3	,521
Likelihood Ratio	2,283	3	,516
N of Valid Cases	50		

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 5,00.

The conditions for the chi-square test are satisfied. There are no expected counts less than 5. The significance value is equal to 52.1%. As a result, H_0 can't be rejected. We will assume that there is no difference between the two groups in terms of how often they play video games. Both groups can be compared in terms of how often they play video games.

H.4:

Chi-square test:

H_0: There is no difference between the two groups in terms of who pays for the rent, mortgage or bills in their households.

H_1: There is a difference between the two groups in terms of who pays for the rent, mortgage or bills in their households.

Who_pays_for_the_rent,_mortgage_or_bills_in_your_household? * Efficiency Crosstabulation

			Efficiency		Total
			Low	High	
Who_pays_for_the_rent,_mortgage_or_bills_in_your_household?	I pay, either part or the whole amount of the rent, mortgage or bills in my household	Count	23	19	42
		Expected Count	21,0	21,0	42,0
	Somebody else does it, I don't pay anything	Count	2	6	8
		Expected Count	4,0	4,0	8,0
Total		Count	25	25	50
		Expected Count	25,0	25,0	50,0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	2,381[a]	1	,123		
Continuity Correction[b]	1,339	1	,247		
Likelihood Ratio	2,475	1	,116		
Fisher's Exact Test				,247	,123
N of Valid Cases	50				

a. 2 cells (50,0%) have expected count less than 5. The minimum expected count is 4,00.

b. Computed only for a 2x2 table

The conditions for the chi-square test are not satisfied. There are 2 expected counts less than 5. Therefore, the Fisher's exact test must be analyzed instead. The significance value is equal to 24.7% (2-tailed). As a result, H_0 can't be rejected. We will assume that there is no difference between the two groups in terms of who pays for the rent, mortgage or bills in their households. Both groups can be compared in terms of who pays for the rent, mortgage or bills in their households.

H.5:

Kolmogorov-Smirnov test

H_0: The total energy consumption variable follows the normal theoretical distribution.

H_1: The total energy consumption variable does not follow the normal theoretical distribution.

One-Sample Kolmogorov-Smirnov Test

		TEC_Round_0	TEC_Round_1	TEC_Round_2	TEC_Round_3	TEC_Round_4	TEC_Round_5	TEC_Round_6	TEC_Round_7	TEC_Round_8	TEC_Round_9	TEC_Round_10
N		25	25	25	25	25	25	25	25	25	25	25
Normal Parameters^{a,b}	Mean	16347,50	16269,28	16007,23	15491,42	14933,58	14759,22	14208,31	13082,29	11535,44	11243,36	11129,75
	Std. Deviation	,000^{c}	868,623	1198,633	1854,821	2224,247	2109,448	2499,714	4359,891	5745,117	5472,960	5591,764
Most Extreme Differences	Absolute		,141	,133	,201	,212	,135	,178	,189	,236	,222	,232
	Positive		,141	,125	,152	,169	,109	,120	,111	,115	,114	,110
	Negative		-,129	-,133	-,201	-,212	-,135	-,178	-,189	-,236	-,222	-,232
Test Statistic			,141	,133	,201	,212	,135	,178	,189	,236	,222	,232
Asymp. Sig. (2-tailed)			,200^{d,e}	,200^{d,e}	,010^{d}	,005^{d}	,200^{d,e}	,040^{d}	,021^{d}	,001^{d}	,003^{d}	,001^{d}

a. Test distribution is Normal.

b. Calculated from data.

c. The distribution has no variance for this variable. One-Sample Kolmogorov-Smirnov Test cannot be performed.

d. Lilliefors Significance Correction.

e. This is a lower bound of the true significance.

One-Sample Kolmogorov-Smirnov Test

		TEC_Round_0	TEC_Round_1	TEC_Round_2	TEC_Round_3	TEC_Round_4	TEC_Round_5	TEC_Round_6	TEC_Round_7	TEC_Round_8	TEC_Round_9	TEC_Round_10
N		25	25	25	25	25	25	25	25	25	25	25
Normal Parameters^{a,b}	Mean	21212,72	20335,28	19259,59	18052,69	17210,40	16272,02	13543,81	12165,90	11284,31	10787,40	10628,84
	Std. Deviation	,000	1070,959	3423,604	5131,176	5520,603	5806,539	6857,722	6713,853	6917,655	6720,002	6830,983
Most Extreme Differences	Absolute	,500	,125	,198	,285	,263	,228	,172	,203	,142	,157	,164
	Positive	,500	,125	,198	,210	,183	,159	,125	,203	,128	,114	,137
	Negative	-,500	-,109	-,198	-,285	-,263	-,228	-,172	-,167	-,142	-,157	-,164
Test Statistic		,500	,125	,198	,285	,263	,228	,172	,203	,142	,157	,164
Asymp. Sig. (2-tailed)		,000^{c}	,200^{c,d}	,012^{c}	,000^{c}	,000^{c}	,002^{c}	,055^{c}	,009^{c}	,200^{c,d}	,115^{c}	,080^{c}

a. Test distribution is Normal.

b. Calculated from data.

c. Lilliefors Significance Correction.

d. This is a lower bound of the true significance.

On the upper table the tests results are shown for the high efficiency houses group and on the lower table the results for the low efficiency houses group are shown. The significance values greater than 5% mean that the variable follows the normal theoretical distribution (H_0 can't be rejected). As a result, total energy consumption in rounds 1, 2 and 5 follow the normal theoretical distribution in the high efficiency houses group. In the low efficiency houses group, rounds 1, 6, 8, 9 and 10 follow the normal theoretical distribution.

H.6:

Wilkoxon Rank-Sum test

H_0: There is no difference between the two groups in terms of the total energy consumption.

H_1: There is a difference between the two groups in terms of the total energy consumption.

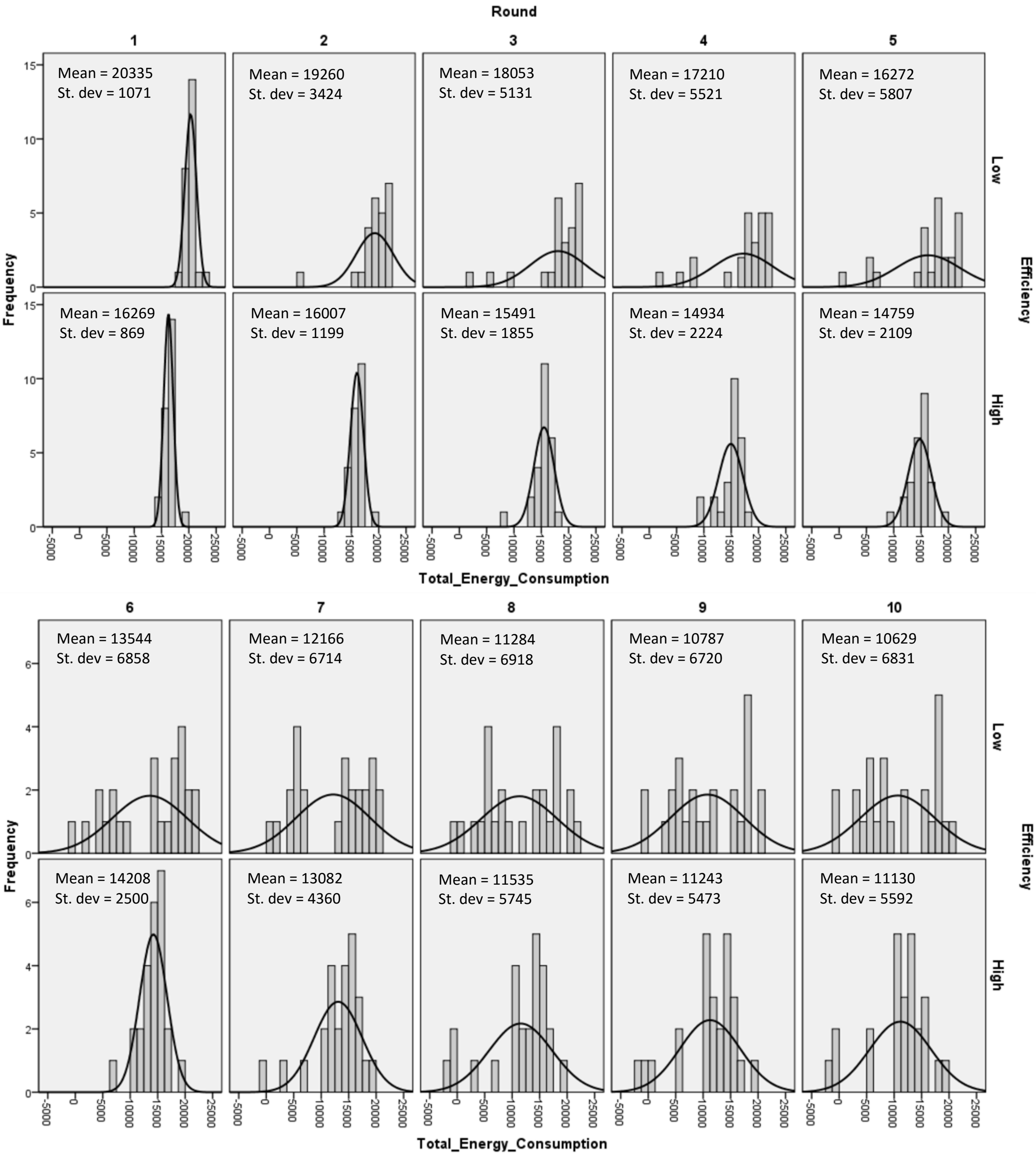
Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of TEC_Round_0 is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,000	Reject the null hypothesis.
2	The distribution of TEC_Round_1 is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,000	Reject the null hypothesis.
3	The distribution of TEC_Round_2 is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,000	Reject the null hypothesis.
4	The distribution of TEC_Round_3 is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,000	Reject the null hypothesis.
5	The distribution of TEC_Round_4 is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,000	Reject the null hypothesis.
6	The distribution of TEC_Round_5 is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,003	Reject the null hypothesis.
7	The distribution of TEC_Round_6 is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,554	Retain the null hypothesis.
8	The distribution of TEC_Round_7 is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,900	Retain the null hypothesis.
9	The distribution of TEC_Round_8 is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,823	Retain the null hypothesis.
10	The distribution of TEC_Round_9 is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,839	Retain the null hypothesis.
11	The distribution of TEC_Round_10 is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,808	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

The significance level from round 1 to round 5 is below 5%, resulting that H_0 is rejected. Therefore, from round 0 to round 5 significant differences between the two groups can be identified in terms of their total energy consumption. From round 6 to round 10 the significance level of the statistical test is greater than 5%, resulting that H_0 can't be rejected. Therefore, from round 6 to round 10 there are no significant differences between the two groups in terms of their total energy consumption.

In addition, in the following graphs can be seen different histograms of the total energy consumption per round per group. It can be visually seen how the two groups become similar throughout the rounds.


Round
1
2
3
4
5
Mean = 20335
St. dev = 1071
Mean = 19260
St. dev = 3424
Mean = 18053
St. dev = 5131
Mean = 17210
St. dev = 5521
Mean = 16272
St. dev = 5807
Mean = 16269
St. dev = 869
Mean = 16007
St. dev = 1199
Mean = 15491
St. dev = 1855
Mean = 14934
St. dev = 2224
Mean = 14759
St. dev = 2109
Low
High
Efficiency
Frequency
Total_Energy_Consumption
6
7
8
9
10
Mean = 13544
St. dev = 6858
Mean = 12166
St. dev = 6714
Mean = 11284
St. dev = 6918
Mean = 10787
St. dev = 6720
Mean = 10629
St. dev = 6831
Mean = 14208
St. dev = 2500
Mean = 13082
St. dev = 4360
Mean = 11535
St. dev = 5745
Mean = 11243
St. dev = 5473
Mean = 11130
St. dev = 5592
Low
High
Efficiency
Frequency
Total_Energy_Consumption

<u>H.7:</u>

<u>Kolmogorov-Smirnov test</u>

H_0: The comfort level variable follows the normal theoretical distribution.

H_1: The comfort level variable does not follow the normal theoretical distribution.

One-Sample Kolmogorov-Smirnov Test

		Comfort_Round_0	Comfort_Round_1	Comfort_Round_2	Comfort_Round_3	Comfort_Round_4	Comfort_Round_5	Comfort_Round_6	Comfort_Round_7	Comfort_Round_8	Comfort_Round_9	Comfort_Round_10
N		25	25	25	25	25	25	25	25	25	25	25
Normal Parameters[a,b]	Mean	5100,00	5166,43	5236,76	5349,87	5421,87	5444,88	5278,43	5283,61	5282,08	5273,13	5319,46
	Std. Deviation	,000[c]	275,974	351,666	507,306	565,412	594,517	648,416	787,138	807,094	787,148	844,397
Most Extreme Differences	Absolute		,125	,104	,163	,120	,087	,132	,159	,086	,078	,137
	Positive		,088	,104	,082	,088	,087	,084	,074	,065	,058	,137
	Negative		-,125	-,060	-,163	-,120	-,072	-,132	-,159	-,086	-,078	-,103
Test Statistic			,125	,104	,163	,120	,087	,132	,159	,086	,078	,137
Asymp. Sig. (2-tailed)			,200[d,e]	,200[d,e]	,086[d]	,200[d,e]	,200[d,e]	,200[d,e]	,103[d]	,200[d,e]	,200[d,e]	,200[d,e]

a. Test distribution is Normal.

b. Calculated from data.

c. The distribution has no variance for this variable. One-Sample Kolmogorov-Smirnov Test cannot be performed.

d. Lilliefors Significance Correction.

e. This is a lower bound of the true significance.

One-Sample Kolmogorov-Smirnov Test

		Comfort_Round_0	Comfort_Round_1	Comfort_Round_2	Comfort_Round_3	Comfort_Round_4	Comfort_Round_5	Comfort_Round_6	Comfort_Round_7	Comfort_Round_8	Comfort_Round_9	Comfort_Round_10
N		25	25	25	25	25	25	25	25	25	25	25
Normal Parameters[a,b]	Mean	5100,00	5281,61	5361,18	5483,01	5593,30	5713,47	5748,32	5750,03	5727,09	5760,37	5899,23
	Std. Deviation	,000[c]	297,977	380,726	425,106	583,938	608,962	721,338	793,946	786,293	820,680	771,569
Most Extreme Differences	Absolute		,128	,132	,100	,140	,138	,122	,133	,117	,111	,090
	Positive		,128	,109	,100	,078	,080	,064	,089	,081	,052	,072
	Negative		-,071	-,132	-,086	-,140	-,138	-,122	-,133	-,117	-,111	-,090
Test Statistic			,128	,132	,100	,140	,138	,122	,133	,117	,111	,090
Asymp. Sig. (2-tailed)			,200[d,e]	,200[d,e]	,200[d,e]	,200[d,e]	,200[d,e]	,200[d,e]	,200[d,e]	,200[d,e]	,200[d,e]	,200[d,e]

a. Test distribution is Normal.

b. Calculated from data.

c. The distribution has no variance for this variable. One-Sample Kolmogorov-Smirnov Test cannot be performed.

d. Lilliefors Significance Correction.

e. This is a lower bound of the true significance.

On the upper table the tests results are shown for the low efficiency houses group and on the lower table the results for the high efficiency houses group are shown. The significance values greater than 5% mean that the variable follows the normal theoretical distribution (H_0 can't be rejected). The test results in all the rounds in both groups show just values greater than 5%. Therefore, the comfort level in all the rounds in both groups follow the normal theoretical distribution.

H.8:

Student t-test for differences in averages

H_0: There is no difference between the two groups in terms of the comfort level.

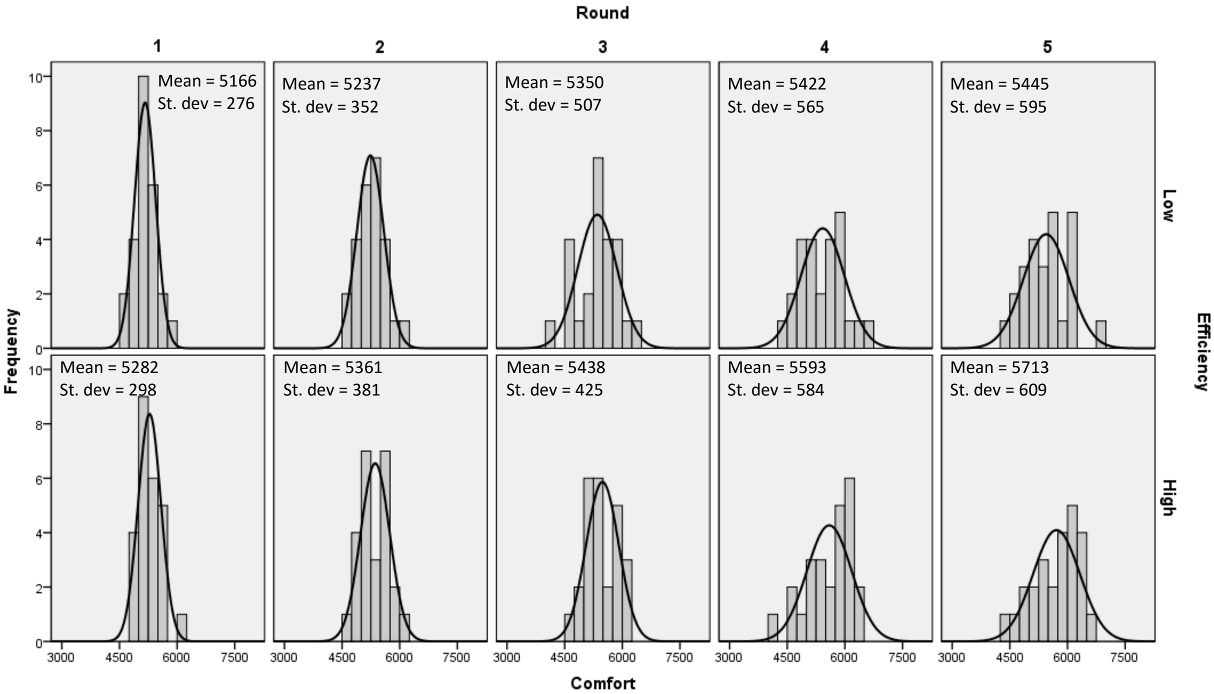
H_1: There is a difference between the two groups in terms of the comfort level.

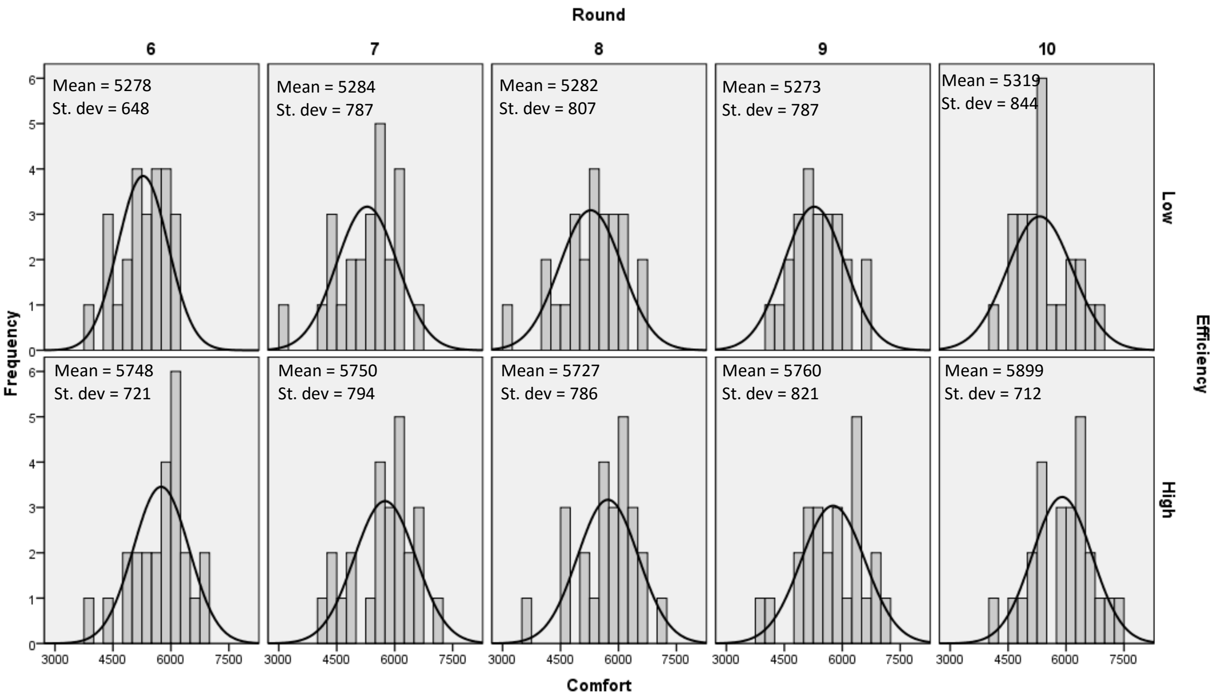
Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Comfort_Round_1	Equal variances assumed	,522	,473	-1,418	48	,163	-115,186	81,229	-278,507	48,135
	Equal variances not assumed			-1,418	47,720	,163	-115,186	81,229	-278,532	48,159
Comfort_Round_2	Equal variances assumed	,889	,350	-1,200	48	,236	-124,415	103,657	-332,833	84,002
	Equal variances not assumed			-1,200	47,701	,236	-124,415	103,657	-332,866	84,036
Comfort_Round_3	Equal variances assumed	,111	,741	-1,006	48	,320	-133,141	132,374	-399,298	133,015
	Equal variances not assumed			-1,006	46,574	,320	-133,141	132,374	-399,509	133,226
Comfort_Round_4	Equal variances assumed	,002	,966	-1,055	48	,297	-171,425	162,564	-498,281	155,431
	Equal variances not assumed			-1,055	47,950	,297	-171,425	162,564	-498,290	155,440
Comfort_Round_5	Equal variances assumed	,001	,980	-1,578	48	,121	-268,592	170,210	-610,822	73,637
	Equal variances not assumed			-1,578	47,972	,121	-268,592	170,210	-610,827	73,643
Comfort_Round_6	Equal variances assumed	,129	,721	-2,422	48	,019	-469,888	193,987	-859,924	-79,851
	Equal variances not assumed			-2,422	47,465	,019	-469,888	193,987	-860,038	-79,737
Comfort_Round_7	Equal variances assumed	,005	,943	-2,086	48	,042	-466,418	223,601	-915,999	-16,838
	Equal variances not assumed			-2,086	47,996	,042	-466,418	223,601	-915,999	-16,837
Comfort_Round_8	Equal variances assumed	,006	,939	-1,975	48	,054	-445,008	225,358	-898,121	8,105
	Equal variances not assumed			-1,975	47,967	,054	-445,008	225,358	-898,129	8,113
Comfort_Round_9	Equal variances assumed	,292	,592	-2,142	48	,037	-487,236	227,431	-944,516	-29,956
	Equal variances not assumed			-2,142	47,917	,037	-487,236	227,431	-944,536	-29,935
Comfort_Round_10	Equal variances assumed	,002	,967	-2,534	48	,015	-579,768	228,764	-1039,729	-119,807
	Equal variances not assumed			-2,534	47,615	,015	-579,768	228,764	-1039,825	-119,711

The significance value of the Levene's test is greater than 5% in all the rounds, meaning that equal variances can be assumed. From round 1 to round 5 the significance value of the student t-test is greater 5%, resulting that H_0 can't rejected. Therefore, from round 1 to round 5 no significant differences between the two groups can be identified in terms of their comfort level. From round 6 to round 10 the significance value of the student t-test is below 5%, resulting that H_0 is rejected. Therefore, from round 6 to round 10 significant differences between the two groups can be identified in terms of their comfort level.

In addition, in the following graphs can be seen different histograms of the comfort level per round per group. It can be visually seen how the two groups become different throughout the rounds.


Round
1
2
3
4
5
Mean = 5166
St. dev = 276
Mean = 5237
St. dev = 352
Mean = 5350
St. dev = 507
Mean = 5422
St. dev = 565
Mean = 5445
St. dev = 595
Low
Mean = 5282
St. dev = 298
Mean = 5361
St. dev = 381
Mean = 5438
St. dev = 425
Mean = 5593
St. dev = 584
Mean = 5713
St. dev = 609
High
Frequency
Efficiency
Comfort


Round
6
7
8
9
10
Mean = 5278
St. dev = 648
Mean = 5284
St. dev = 787
Mean = 5282
St. dev = 807
Mean = 5273
St. dev = 787
Mean = 5319
St. dev = 844
Low
Mean = 5748
St. dev = 721
Mean = 5750
St. dev = 794
Mean = 5727
St. dev = 786
Mean = 5760
St. dev = 821
Mean = 5899
St. dev = 712
High
Frequency
Efficiency
Comfort

H.9:

Kolmogorov-Smirnov test

H_0: The reduction of the total energy consumption by getting rid of initial appliances follows the normal theoretical distribution.

H_1: The reduction of the total energy consumption by getting rid of initial appliances does not follow the normal theoretical distribution.

One-Sample Kolmogorov-Smirnov Test

		Round 0 (1)	Round 1 (1)	Round 2 (1)	Round 3 (1)	Round 4 (1)	Round 5 (1)	Round 6 (1)	Round 7 (1)	Round 8 (1)	Round 9 (1)	Round 10 (1)
N		25	25	25	25	25	25	25	25	25	25	25
Normal Parameters[a,b]	Mean	,00	-208,20	-943,99	-1555,31	-2139,94	-2311,14	-4484,64	-5203,11	-5186,06	-5348,06	-5342,41
	Std. Deviation	,000[c]	357,395	2597,864	3581,816	4314,995	4221,746	5953,249	6100,924	6052,029	5957,872	6030,039
Most Extreme Differences	Absolute		,440	,396	,441	,425	,349	,342	,341	,341	,311	,311
	Positive		,280	,358	,332	,310	,292	,226	,208	,211	,209	,207
	Negative		-,440	-,396	-,441	-,425	-,349	-,342	-,341	-,341	-,311	-,311
Test Statistic			,440	,396	,441	,425	,349	,342	,341	,341	,311	,311
Asymp. Sig. (2-tailed)			,000[d]	,000[d]	,000[d]	,000[d]	,000[d]	,000[d]	,000[d]	,000[d]	,000[d]	,000[d]

a. Test distribution is Normal.

b. Calculated from data.

c. The distribution has no variance for this variable. One-Sample Kolmogorov-Smirnov Test cannot be performed.

d. Lilliefors Significance Correction.

One-Sample Kolmogorov-Smirnov Test

		Round 0 (1)	Round 1 (1)	Round 2 (1)	Round 3 (1)	Round 4 (1)	Round 5 (1)	Round 6 (1)	Round 7 (1)	Round 8 (1)	Round 9 (1)	Round 10 (1)
N		25	25	25	25	25	25	25	25	25	25	25
Normal Parameters[a,b]	Mean	,00	-246,00	-416,00	-501,59	-539,19	-554,78	-606,98	-1409,03	-2306,67	-2337,07	-2372,87
	Std. Deviation	,000[c]	395,389	435,429	450,680	470,947	478,513	466,601	2800,877	3847,166	3800,454	3845,852
Most Extreme Differences	Absolute		,373	,230	,187	,194	,197	,143	,425	,403	,399	,402
	Positive		,267	,170	,133	,126	,139	,139	,307	,274	,269	,269
	Negative		-,373	-,230	-,187	-,194	-,197	-,143	-,425	-,403	-,399	-,402
Test Statistic			,373	,230	,187	,194	,197	,143	,425	,403	,399	,402
Asymp. Sig. (2-tailed)			,000[d]	,001[d]	,024[d]	,016[d]	,014[d]	,198[d]	,000[d]	,000[d]	,000[d]	,000[d]

a. Test distribution is Normal.

b. Calculated from data.

c. The distribution has no variance for this variable. One-Sample Kolmogorov-Smirnov Test cannot be performed.

d. Lilliefors Significance Correction.

On the upper table the tests results are shown for the low efficiency houses group and on the lower table the results for the high efficiency houses group are shown. The significance values greater than 5% mean that the variable follows the normal theoretical distribution (H_0 can't be rejected). Just round 6 in the high efficiency houses group the significance value is greater than 5%, meaning that only in that round the distribution follows the normal theoretical distribution.

H.10:

Wilkoxon Rank-Sum test

H_0: There is no difference between the two groups in terms of the reduction of the total energy consumption by getting rid of initial appliances.

H_1: There is a difference between the two groups in terms of the reduction of the total energy consumption by getting rid of initial appliances.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Round 0 (1) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	1,000	Retain the null hypothesis.
2	The distribution of Round 1 (1) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,760	Retain the null hypothesis.
3	The distribution of Round 2 (1) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,567	Retain the null hypothesis.
4	The distribution of Round 3 (1) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,372	Retain the null hypothesis.
5	The distribution of Round 4 (1) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,393	Retain the null hypothesis.
6	The distribution of Round 5 (1) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,150	Retain the null hypothesis.
7	The distribution of Round 6 (1) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,019	Reject the null hypothesis.
8	The distribution of Round 7 (1) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,003	Reject the null hypothesis.
9	The distribution of Round 8 (1) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,020	Reject the null hypothesis.
10	The distribution of Round 9 (1) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,013	Reject the null hypothesis.
11	The distribution of Round 10 (1) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,013	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

From round 1 to round 5 the significance value of the test is greater 5%, resulting that H_0 can't rejected. Therefore, from round 1 to round 5 no significant differences between the two groups can be identified in terms of the reduction of the total energy consumption by getting rid of initial appliances. From round 6 to round 10 the significance value of the test is below 5%, resulting that H_0 is rejected. Therefore, from round 6 to round 10 significant differences between the two groups can be identified in terms of the reduction of the total energy consumption by getting rid of initial appliances.

H.11:

Kolmogorov-Smirnov test

H_0: The reduction of the total energy consumption by improving efficiency of the initial appliances follows the normal theoretical distribution.

H_1: The reduction of the total energy consumption by improving efficiency of the initial appliances follows the normal theoretical distribution.

One-Sample Kolmogorov-Smirnov Test

		Round 0 (2)	Round 1 (2)	Round 2 (2)	Round 3 (2)	Round 4 (2)	Round 5 (2)	Round 6 (2)	Round 7 (2)	Round 8 (2)	Round 9 (2)	Round 10 (2)
N		25	25	25	25	25	25	25	25	25	25	25
Normal Parameters[a,b]	Mean	,00	-63,96	-165,11	-622,89	-803,48	-1141,52	-1478,32	-1720,00	-2635,26	-2742,98	-2773,36
	Std. Deviation	,000[c]	682,996	809,657	2162,053	2125,750	2736,522	3460,641	3507,999	3792,872	3762,488	3769,055
Most Extreme Differences	Absolute		,423	,339	,307	,291	,299	,356	,333	,244	,242	,245
	Positive		,423	,339	,307	,273	,258	,255	,288	,244	,233	,231
	Negative		-,377	-,301	-,307	-,291	-,299	-,356	-,333	-,230	-,242	-,245
Test Statistic			,423	,339	,307	,291	,299	,356	,333	,244	,242	,245
Asymp. Sig. (2-tailed)			,000[d]	,000[d]	,000[d]	,000[d]	,000[d]	,000[d]	,000[d]	,001[d]	,001[d]	,000[d]

a. Test distribution is Normal.

b. Calculated from data.

c. The distribution has no variance for this variable. One-Sample Kolmogorov-Smirnov Test cannot be performed.

d. Lilliefors Significance Correction.

One-Sample Kolmogorov-Smirnov Test

		Round 0 (2)	Round 1 (2)	Round 2 (2)	Round 3 (2)	Round 4 (2)	Round 5 (2)	Round 6 (2)	Round 7 (2)	Round 8 (2)	Round 9 (2)	Round 10 (2)
N		25	25	25	25	25	25	25	25	25	25	25
Normal Parameters[a,b]	Mean	,00	-111,26	-151,39	-341,88	-782,55	-694,80	-986,27	-909,88	-943,26	-1159,37	-1194,93
	Std. Deviation	,000[c]	378,841	444,170	1171,283	1890,524	1647,362	1949,076	1958,749	1970,296	1978,295	1985,513
Most Extreme Differences	Absolute		,409	,388	,373	,365	,354	,262	,234	,265	,207	,214
	Positive		,344	,327	,345	,299	,292	,262	,226	,198	,185	,180
	Negative		-,409	-,388	-,373	-,365	-,354	-,249	-,234	-,265	-,207	-,214
Test Statistic			,409	,388	,373	,365	,354	,262	,234	,265	,207	,214
Asymp. Sig. (2-tailed)			,000[d]	,000[d]	,000[d]	,000[d]	,000[d]	,000[d]	,001[d]	,000[d]	,007[d]	,005[d]

a. Test distribution is Normal.

b. Calculated from data.

c. The distribution has no variance for this variable. One-Sample Kolmogorov-Smirnov Test cannot be performed.

d. Lilliefors Significance Correction.

On the upper table the tests results are shown for the low efficiency houses group and on the lower table the results for the high efficiency houses group are shown. The significance values greater than 5% mean that the variable follows the normal theoretical distribution (H_0 can’t be rejected). In all the rounds in both groups the significance value of the statistical test is below 5%, meaning that no normal distributions can be identified.

H.12:

Wilkoxon Rank-Sum test

H_0: There is no difference between the two groups in terms of the reduction of the total energy consumption by improving efficiency of the initial appliances.

H_1: There is a difference between the two groups in terms of the reduction of the total energy consumption by improving efficiency of the initial appliances.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Round 0 (2) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	1,000	Retain the null hypothesis.
2	The distribution of Round 1 (2) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,002	Reject the null hypothesis.
3	The distribution of Round 2 (2) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,008	Reject the null hypothesis.
4	The distribution of Round 3 (2) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,005	Reject the null hypothesis.
5	The distribution of Round 4 (2) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,032	Reject the null hypothesis.
6	The distribution of Round 5 (2) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,056	Retain the null hypothesis.
7	The distribution of Round 6 (2) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,808	Retain the null hypothesis.
8	The distribution of Round 7 (2) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,218	Retain the null hypothesis.
9	The distribution of Round 8 (2) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,024	Reject the null hypothesis.
10	The distribution of Round 9 (2) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,035	Reject the null hypothesis.
11	The distribution of Round 10 (2) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,051	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Rounds 5, 6, 7 and 10 show a significance value of the test greater 5%, resulting that H_0 can't rejected. Therefore, in rounds 5, 6, 7 and 10 no significant differences between the two groups can be identified in terms of the reduction of the total energy consumption by improving efficiency of the initial appliances. Rounds 1, 2, 3, 4, 8 and 9 show a significance value of the test below 5%, resulting that H_0 is rejected. Therefore, in rounds 1, 2, 3, 4, 8 and 9 significant differences between the two groups can be identified in terms of the reduction of the total energy consumption by improving efficiency of the initial appliances.

H.13:

Kolmogorov-Smirnov test

H_0: The increase of the total energy consumption by including new appliances follows the normal theoretical distribution.

H_1: The increase of the total energy consumption by including new appliances does not follow the normal theoretical distribution.

One-Sample Kolmogorov-Smirnov Test

		Round 0 (3)	Round 1 (3)	Round 2 (3)	Round 3 (3)	Round 4 (3)	Round 5 (3)	Round 6 (3)	Round 7 (3)	Round 8 (3)	Round 9 (3)	Round 10 (3)
N		25	25	25	25	25	25	25	25	25	25	25
Normal Parameters[a,b]	Mean	,00	52,44	280,56	499,79	609,97	611,88	681,40	734,30	839,46	806,92	889,06
	Std. Deviation	,000[c]	117,701	493,925	604,180	626,909	669,182	819,753	817,417	1192,062	871,269	962,440
Most Extreme Differences	Absolute		,460	,285	,230	,217	,286	,308	,260	,242	,257	,278
	Positive		,460	,276	,230	,217	,286	,308	,260	,242	,257	,278
	Negative		-,328	-,285	-,204	-,165	-,180	-,203	-,185	-,241	-,177	-,178
Test Statistic			,460	,285	,230	,217	,286	,308	,260	,242	,257	,278
Asymp. Sig. (2-tailed)			,000[d]	,000[d]	,001[d]	,004[d]	,000[d]	,000[d]	,000[d]	,001[d]	,000[d]	,000[d]

a. Test distribution is Normal.
b. Calculated from data.
c. The distribution has no variance for this variable. One-Sample Kolmogorov-Smirnov Test cannot be performed.
d. Lilliefors Significance Correction.

One-Sample Kolmogorov-Smirnov Test

		Round 0 (3)	Round 1 (3)	Round 2 (3)	Round 3 (3)	Round 4 (3)	Round 5 (3)	Round 6 (3)	Round 7 (3)	Round 8 (3)	Round 9 (3)	Round 10 (3)
N		25	25	25	25	25	25	25	25	25	25	25
Normal Parameters[a,b]	Mean	,00	297,67	391,71	452,72	599,37	674,31	709,14	721,25	709,43	826,34	925,06
	Std. Deviation	,000[c]	544,308	534,258	527,160	627,078	623,170	620,747	626,320	637,089	833,562	887,965
Most Extreme Differences	Absolute		,292	,240	,281	,210	,175	,156	,162	,232	,221	,168
	Positive		,284	,240	,281	,210	,175	,156	,162	,232	,221	,168
	Negative		-,292	-,232	-,195	-,170	-,140	-,127	-,125	-,133	-,161	-,149
Test Statistic			,292	,240	,281	,210	,175	,156	,162	,232	,221	,168
Asymp. Sig. (2-tailed)			,000[d]	,001[d]	,000[d]	,006[d]	,046[d]	,117[d]	,090[d]	,001[d]	,003[d]	,068[d]

a. Test distribution is Normal.
b. Calculated from data.
c. The distribution has no variance for this variable. One-Sample Kolmogorov-Smirnov Test cannot be performed.
d. Lilliefors Significance Correction.

On the upper table the tests results are shown for the low efficiency houses group and on the lower table the results for the high efficiency houses group are shown. The significance values greater than 5% mean that the variable follows the normal theoretical distribution (H_0 can't be rejected). Just in rounds 6, 7 and 10 in the high efficiency houses group the significance value of the statistical test is greater 5%. Therefore, just in rounds 6, 7 and 10 in the high efficiency houses group we can identify normal distributions.

H.14:

Wilkoxon Rank-Sum test

H_0: There is no difference between the two groups in terms of the increase of the total energy consumption by including new appliances.

H_1: There is a difference between the two groups in terms of the increase of the total energy consumption by including new appliances.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Round 0 (3) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	1,000	Retain the null hypothesis.
2	The distribution of Round 1 (3) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,042	Reject the null hypothesis.
3	The distribution of Round 2 (3) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,149	Retain the null hypothesis.
4	The distribution of Round 3 (3) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,718	Retain the null hypothesis.
5	The distribution of Round 4 (3) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,907	Retain the null hypothesis.
6	The distribution of Round 5 (3) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,541	Retain the null hypothesis.
7	The distribution of Round 6 (3) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,496	Retain the null hypothesis.
8	The distribution of Round 7 (3) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,648	Retain the null hypothesis.
9	The distribution of Round 8 (3) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,712	Retain the null hypothesis.
10	The distribution of Round 9 (3) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,861	Retain the null hypothesis.
11	The distribution of Round 10 (3) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,691	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Just round 1 shows a significance value of the test below 5%, resulting that H_0 is rejected. Therefore, in round 1 significant differences between the two groups can be identified in terms of the increase of the total energy consumption by including new appliances. In all the other rounds the significance value of the test is greater 5%, resulting that H_0 can't rejected. Therefore, in all the other rounds no significant differences between the two groups can be identified in terms of the increase of the total energy consumption by including new appliances.

H.15:

Kolmogorov-Smirnov test

H_0: The reduction of the total energy consumption by including energy management devices follows the normal theoretical distribution.

H_1: The reduction of the total energy consumption by including energy management devices does not follow the normal theoretical distribution.

One-Sample Kolmogorov-Smirnov Test

		Round 0 (4)	Round 1 (4)	Round 2 (4)	Round 3 (4)	Round 4 (4)	Round 5 (4)	Round 6 (4)	Round 7 (4)	Round 8 (4)	Round 9 (4)	Round 10 (4)
N		25	25	25	25	25	25	25	25	25	25	25
Normal Parameters[a,b]	Mean	,00	-466,22	-763,89	-763,47	-889,65	-983,95	-1091,89	-1128,02	-1058,07	-1192,71	-1376,77
	Std. Deviation	,000[c]	599,721	1037,088	855,456	889,697	1005,316	1154,547	1143,529	1032,177	1201,847	1569,413
Most Extreme Differences	Absolute		,302	,223	,181	,161	,245	,216	,229	,186	,179	,190
	Positive		,206	,223	,176	,150	,157	,166	,162	,152	,160	,190
	Negative		-,302	-,169	-,181	-,161	-,245	-,216	-,229	-,186	-,179	-,190
Test Statistic			,302	,223	,181	,161	,245	,216	,229	,186	,179	,190
Asymp. Sig. (2-tailed)			,000[d]	,002[d]	,034[d]	,092[d]	,000[d]	,004[d]	,002[d]	,026[d]	,038[d]	,020[d]

a. Test distribution is Normal.

b. Calculated from data.

c. The distribution has no variance for this variable. One-Sample Kolmogorov-Smirnov Test cannot be performed.

d. Lilliefors Significance Correction.

One-Sample Kolmogorov-Smirnov Test

		Round 0 (4)	Round 1 (4)	Round 2 (4)	Round 3 (4)	Round 4 (4)	Round 5 (4)	Round 6 (4)	Round 7 (4)	Round 8 (4)	Round 9 (4)	Round 10 (4)
N		25	25	25	25	25	25	25	25	25	25	25
Normal Parameters[a,b]	Mean	,00	48,62	134,70	63,63	-29,00	-9,81	5,42	-174,96	-245,72	-379,69	-457,56
	Std. Deviation	,000[c]	493,980	552,297	569,536	679,595	717,787	762,979	1204,779	1174,339	1462,167	1467,467
Most Extreme Differences	Absolute		,379	,276	,228	,194	,137	,153	,181	,202	,225	,179
	Positive		,379	,276	,228	,194	,137	,153	,092	,108	,111	,083
	Negative		-,205	-,230	-,135	-,117	-,111	-,082	-,181	-,202	-,225	-,179
Test Statistic			,379	,276	,228	,194	,137	,153	,181	,202	,225	,179
Asymp. Sig. (2-tailed)			,000[d]	,000[d]	,002[d]	,016[d]	,200[d,e]	,135[d]	,033[d]	,010[d]	,002[d]	,037[d]

a. Test distribution is Normal.

b. Calculated from data.

c. The distribution has no variance for this variable. One-Sample Kolmogorov-Smirnov Test cannot be performed.

d. Lilliefors Significance Correction.

e. This is a lower bound of the true significance.

On the upper table the tests results are shown for the low efficiency houses group and on the lower table the results for the high efficiency houses group are shown. The significance values greater than 5% mean that the variable follows the normal theoretical distribution (H_0 can't be rejected). In the low efficiency houses group just round 4 shows a significance value greater than 5%, meaning that only in round 4 we can identify a normal distribution in the low efficiency houses group. In the high efficiency houses group just rounds 5 and 6 show a significance value greater than 5%, meaning that only in round 5 and 6 we can identify a normal distribution in the high efficiency houses group.

H.16:

Wilkoxon Rank-Sum test

H_0: There is no difference between the two groups in terms of the reduction of the total energy consumption by including energy management devices.

H_1: There is a difference between the two groups in terms of the reduction of the total energy consumption by including energy management devices.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Round 0 (4) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	1,000	Retain the null hypothesis.
2	The distribution of Round 1 (4) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,042	Reject the null hypothesis.
3	The distribution of Round 2 (4) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,000	Reject the null hypothesis.
4	The distribution of Round 3 (4) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,002	Reject the null hypothesis.
5	The distribution of Round 4 (4) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,002	Reject the null hypothesis.
6	The distribution of Round 5 (4) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,001	Reject the null hypothesis.
7	The distribution of Round 6 (4) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,001	Reject the null hypothesis.
8	The distribution of Round 7 (4) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,002	Reject the null hypothesis.
9	The distribution of Round 8 (4) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,005	Reject the null hypothesis.
10	The distribution of Round 9 (4) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,013	Reject the null hypothesis.
11	The distribution of Round 10 (4) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,019	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

All the rounds show a significance value of the test below 5%, resulting that H_0 is rejected. Therefore, in all the rounds significant differences between the two groups can be identified in terms of the reduction of the total energy consumption by including energy management devices.

H.17:

Kolmogorov-Smirnov test

H_0: The reduction of the total energy consumption by including energy production devices follows the normal theoretical distribution.

H_1: The reduction of the total energy consumption by including energy production devices does not follow the normal theoretical distribution.

One-Sample Kolmogorov-Smirnov Test

		Round 0 (5)	Round 1 (5)	Round 2 (5)	Round 3 (5)	Round 4 (5)	Round 5 (5)	Round 6 (5)	Round 7 (5)	Round 8 (5)	Round 9 (5)	Round 10 (5)
N		25	25	25	25	25	25	25	25	25	25	25
Normal Parameters[a,b]	Mean	,00	-191,50	-360,70	-718,16	-779,22	-1115,96	-1295,45	-1729,99	-1888,49	-1948,49	-1980,39
	Std. Deviation	,000[c]	548,987	779,972	1199,188	1243,143	1364,793	1457,500	1453,461	1554,863	1507,162	1532,619
Most Extreme Differences	Absolute		,516	,478	,325	,335	,207	,187	,163	,119	,137	,143
	Positive		,364	,322	,275	,265	,207	,187	,117	,112	,098	,098
	Negative		-,516	-,478	-,325	-,335	-,193	-,173	-,163	-,119	-,137	-,143
Test Statistic			,516	,478	,325	,335	,207	,187	,163	,119	,137	,143
Asymp. Sig. (2-tailed)			,000[d]	,000[d]	,000[d]	,000[d]	,007[d]	,024[d]	,086[d]	,200[d,e]	,200[d,e]	,200[d]

a. Test distribution is Normal.

b. Calculated from data.

c. The distribution has no variance for this variable. One-Sample Kolmogorov-Smirnov Test cannot be performed.

d. Lilliefors Significance Correction.

e. This is a lower bound of the true significance.

One-Sample Kolmogorov-Smirnov Test

		Round 0 (5)	Round 1 (5)	Round 2 (5)	Round 3 (5)	Round 4 (5)	Round 5 (5)	Round 6 (5)	Round 7 (5)	Round 8 (5)	Round 9 (5)	Round 10 (5)
N		25	25	25	25	25	25	25	25	25	25	25
Normal Parameters[a,b]	Mean	,00	-67,25	-299,29	-528,96	-662,56	-1003,20	-1260,50	-1492,60	-2025,84	-2054,34	-2117,45
	Std. Deviation	,000[c]	237,266	500,309	635,535	830,526	994,575	1108,803	1340,847	1376,478	1377,904	1417,227
Most Extreme Differences	Absolute		,532	,445	,317	,267	,203	,172	,194	,173	,180	,188
	Positive		,388	,275	,203	,213	,157	,128	,133	,108	,102	,095
	Negative		-,532	-,445	-,317	-,267	-,203	-,172	-,194	-,173	-,180	-,188
Test Statistic			,532	,445	,317	,267	,203	,172	,194	,173	,180	,188
Asymp. Sig. (2-tailed)			,000[d]	,000[d]	,000[d]	,000[d]	,009[d]	,056[d]	,016[d]	,053[d]	,035[d]	,022[d]

a. Test distribution is Normal.

b. Calculated from data.

c. The distribution has no variance for this variable. One-Sample Kolmogorov-Smirnov Test cannot be performed.

d. Lilliefors Significance Correction.

On the upper table the tests results are shown for the low efficiency houses group and on the lower table the results for the high efficiency houses group are shown. The significance values greater than 5% mean that the variable follows the normal theoretical distribution (H_0 can't be rejected). In the low efficiency houses group just round 7, 8, 9 and 10 show a significance value greater than 5%, meaning that only in those rounds we can identify a normal distribution in the low efficiency houses group. In the high efficiency houses group just rounds 6 and 8 show a significance value greater than 5%, meaning that only in those rounds we can identify a normal distribution in the high efficiency houses group.

H.18:

Wilkoxon Rank-Sum test

H_0: There is no difference between the two groups in terms of the reduction of the total energy consumption by including energy production devices.

H_1: There is a difference between the two groups in terms of the reduction of the total energy consumption by including energy production devices.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Round 0 (5) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	1,000	Retain the null hypothesis.
2	The distribution of Round 1 (5) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,564	Retain the null hypothesis.
3	The distribution of Round 2 (5) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,786	Retain the null hypothesis.
4	The distribution of Round 3 (5) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,856	Retain the null hypothesis.
5	The distribution of Round 4 (5) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,751	Retain the null hypothesis.
6	The distribution of Round 5 (5) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,928	Retain the null hypothesis.
7	The distribution of Round 6 (5) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,890	Retain the null hypothesis.
8	The distribution of Round 7 (5) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,418	Retain the null hypothesis.
9	The distribution of Round 8 (5) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,741	Retain the null hypothesis.
10	The distribution of Round 9 (5) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,853	Retain the null hypothesis.
11	The distribution of Round 10 (5) is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,793	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

All the rounds show a significance value of the test greater 5%, resulting that H_0 can't rejected. Therefore, in all the rounds no significant differences between the two groups can be identified in terms of the reduction of the total energy consumption by including energy production devices.

H.19:

Chi-square test:

H_0: There is no difference between the number of people who invested in energy management devices and the number of people who invested in energy production devices.

H_1: The number of people who invested in energy management devices is less than the number of people who invested in energy production devices.

Did_you_invest_money_in_energy_management_devices? * Did_you_invest_money_in_energy_production_devices? Crosstabulation

			Did_you_invest_money_in_energy_production_devices?		
			No	Yes	Total
Did_you_invest_money_in_energy_management_devices?	No	Count	5	9	14
		Expected Count	2,0	12,0	14,0
	Yes	Count	2	34	36
		Expected Count	5,0	31,0	36,0
Total		Count	7	43	50
		Expected Count	7,0	43,0	50,0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	7,615[a]	1	,006		
Continuity Correction[b]	5,316	1	,021		
Likelihood Ratio	6,799	1	,009		
Fisher's Exact Test				,014	,014
Linear-by-Linear Association	7,463	1	,006		
N of Valid Cases	50				

a. 1 cells (25,0%) have expected count less than 5. The minimum expected count is 1,96.

b. Computed only for a 2x2 table

The conditions for the chi-square test are not satisfied. There is 1 expected count less than 5. Therefore, the Fisher's exact test must be analyzed instead. The significance value is equal to 1.4% (2-tailed). As a result, H_0 is rejected. The number of people who invested in energy management devices is, indeed, less than the number of people who invested in energy production devices.

H.20:

Kolmogorov-Smirnov test

H_0: The variables follow the normal theoretical distribution.

H_1: The variables do not follow the normal theoretical distribution.

One-Sample Kolmogorov-Smirnov Test

		[I_wanted_to_reduce_my_costs_in_the_game]	[My_energy_consumption_was_too_high_and_I_wanted_to_reduce_it]
N		36	36
Normal Parameters[a,b]	Mean	1,25	,75
	Std. Deviation	,770	,906
Most Extreme Differences	Absolute	,317	,303
	Positive	,266	,225
	Negative	-,317	-,303
Test Statistic		,317	,303
Asymp. Sig. (2-tailed)		,000[c]	,000[c]

a. Test distribution is Normal.

b. Calculated from data.

c. Lilliefors Significance Correction.

One-Sample Kolmogorov-Smirnov Test

		[I_wanted_to_reduce_my_costs_in_the_game]	[My_energy_consumption_was_too_high_and_I_wanted_to_reduce_it]
N		43	43
Normal Parameters[a,b]	Mean	1,14	,74
	Std. Deviation	,861	1,026
Most Extreme Differences	Absolute	,319	,250
	Positive	,216	,146
	Negative	-,319	-,250
Test Statistic		,319	,250
Asymp. Sig. (2-tailed)		,000[c]	,000[c]

a. Test distribution is Normal.

b. Calculated from data.

c. Lilliefors Significance Correction.

On the upper table the tests results are shown for the low efficiency houses group and on the lower table the results for the high efficiency houses group are shown. The significance values are 0% in all cases, meaning that the variables do not follow the normal theoretical distribution (H_0 is rejected).

H.21:

Wilkoxon Rank-Sum test

H_0: There is no difference between reducing costs and reducing energy consumption into the reasons to invest in energy management/production devices.

H_1: Reducing costs is more important than reducing energy consumption into the reasons to invest in energy management devices.

Ranks

		N	Mean Rank	Sum of Ranks
[My_energy_consumption_was_too_high_and_I_wanted_to_reduce_it] - [I_wanted_to_reduce_my_costs_in_the_game]	Negative Ranks	14[a]	8,18	114,50
	Positive Ranks	1[b]	5,50	5,50
	Ties	21[c]		
	Total	36		

a. [My_energy_consumption_was_too_high_and_I_wanted_to_reduce_it] < [I_wanted_to_reduce_my_costs_in_the_game]

b. [My_energy_consumption_was_too_high_and_I_wanted_to_reduce_it] > [I_wanted_to_reduce_my_costs_in_the_game]

c. [My_energy_consumption_was_too_high_and_I_wanted_to_reduce_it] = [I_wanted_to_reduce_my_costs_in_the_game]

Test Statistics[a]

	[My_energy_consumption_was_too_high_and_I_wanted_to_reduce_it] - [I_wanted_to_reduce_my_costs_in_the_game]
Z	-3,218[b]
Asymp. Sig. (2-tailed)	,001

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

Ranks

		N	Mean Rank	Sum of Ranks
[My_energy_consumption_was_too_high_and_I_wanted_to_reduce_it] - [I_wanted_to_reduce_my_costs_in_the_game]	Negative Ranks	15[a]	11,07	166,00
	Positive Ranks	5[b]	8,80	44,00
	Ties	23[c]		
	Total	43		

a. [My_energy_consumption_was_too_high_and_I_wanted_to_reduce_it] < [I_wanted_to_reduce_my_costs_in_the_game]

b. [My_energy_consumption_was_too_high_and_I_wanted_to_reduce_it] > [I_wanted_to_reduce_my_costs_in_the_game]

c. [My_energy_consumption_was_too_high_and_I_wanted_to_reduce_it] = [I_wanted_to_reduce_my_costs_in_the_game]

Test Statistics[a]

	[My_energy_consumption_was_too_high_and_I_wanted_to_reduce_it] - [I_wanted_to_reduce_my_costs_in_the_game]
Z	-2,358[b]
Asymp. Sig. (2-tailed)	,018

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

The upper tables show the results of the statistical test for the energy management devices, while the lower tables show the results of the statistical test for the energy production devices. For energy management devices the significance value is 0.1% and for energy production devices the significance value is 1.8%. Therefore, H_0 is rejected in both cases. Reducing costs is more important than reducing energy consumption as reasons for investing in energy management and energy production devices.

H.22:

Kolmogorov-Smirnov test

H_0: The variables follow the normal theoretical distribution.

H_1: The variables do not follow the normal theoretical distribution.

One-Sample Kolmogorov-Smirnov Test

		[I_often_checked_my_gas_and_electricity_consumption]	[I_preferred_to_buy_luxury_rather_than_eco-friendly_things]	[I_wish_I_had_more_money_in_the_game_to_buy_more_luxurious_things]	[I_wish_I_had_more_money_in_the_game_to_buy_more_eco-friendly_things]	[I_wish_I_had_more_money_in_the_game_to_buy_more_production_devices_(like_solar_panels/wind_turbines)]
N		25	25	25	25	25
Normal Parameters[a,b]	Mean	,84	-,92	-,52	1,12	1,48
	Std. Deviation	,850	,862	1,194	,781	,823
Most Extreme Differences	Absolute	,375	,297	,296	,279	,376
	Positive	,265	,297	,296	,241	,264
	Negative	-,375	-,223	-,178	-,279	-,376
Test Statistic		,375	,297	,296	,279	,376
Asymp. Sig. (2-tailed)		,000[c]	,000[c]	,000[c]	,000[c]	,000[c]

a. Test distribution is Normal.

b. Calculated from data.

c. Lilliefors Significance Correction.

One-Sample Kolmogorov-Smirnov Test

		[I_often_checked_my_gas_and_electricity_consumption]	[I_preferred_to_buy_luxury_rather_than_eco-friendly_things]	[I_wish_I_had_more_money_in_the_game_to_buy_more_luxurious_things]	[I_wish_I_had_more_money_in_the_game_to_buy_more_eco-friendly_things]	[I_wish_I_had_more_money_in_the_game_to_buy_more_production_devices_(like_solar_panels/wind_turbines)]
N		25	25	25	25	25
Normal Parameters[a,b]	Mean	,52	-,92	-,04	,80	1,44
	Std. Deviation	1,295	,954	1,207	1,041	,870
Most Extreme Differences	Absolute	,205	,293	,206	,216	,340
	Positive	,127	,293	,187	,144	,260
	Negative	-,205	-,187	-,206	-,216	-,340
Test Statistic		,205	,293	,206	,216	,340
Asymp. Sig. (2-tailed)		,008[c]	,000[c]	,008[c]	,004[c]	,000[c]

a. Test distribution is Normal.

b. Calculated from data.

c. Lilliefors Significance Correction.

On the upper table the tests results are shown for the low efficiency houses group and on the lower table the results for the high efficiency houses group are shown. The significance values are 0% in all cases, meaning that the variables do not follow the normal theoretical distribution (H_0 is rejected).

H.23:

Wilkoxon Rank-Sum test

H_0: There is no difference between the two groups in terms of each of the questions analyzed.

H_1: There is a difference between the two groups in terms of each of the questions analyzed.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of [I_often_checked_my_gas_and_electricity_consumption] is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,477	Retain the null hypothesis.
2	The distribution of [I_preferred_to_buy_luxury_rather_than_eco-friendly_things] is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,891	Retain the null hypothesis.
3	The distribution of [I_wish_I_had_more_money_in_the_game_to_buy_more_luxurious_things] is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,155	Retain the null hypothesis.
4	The distribution of [I_wish_I_had_more_money_in_the_game_to_buy_more_eco-friendly_things] is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,273	Retain the null hypothesis.
5	The distribution of [I_wish_I_had_more_money_in_the_game_to_buy_more_production_devices_(like_solar_panels/wind_turbines)] is the same across categories of Efficiency.	Independent-Samples Mann-Whitney U Test	,848	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

In all the cases the significance value is greater than 5%, therefore H_0 can´t be rejected. There are no significant differences in each of the 5 analyzed questions between the two groups.

H.24:

Chi-square test:

H_0: v.

H_1: There is a difference between the two groups in terms of the number of people who remembered their initial energy efficiency and the people who did not.

Efficiency * Do_you_remember_the_efficiency_of_the_house_you_play_with Crosstabulation

			Do_you_remember_the_efficiency_of_the_house_you_play_with		
			No	Yes	Total
Efficiency	Low	Count	14	11	25
		Expected Count	13,0	12,0	25,0
	High	Count	12	13	25
		Expected Count	13,0	12,0	25,0
Total		Count	26	24	50
		Expected Count	26,0	24,0	50,0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	,321[a]	1	,571		
Continuity Correction[b]	,080	1	,777		
Likelihood Ratio	,321	1	,571		
Fisher's Exact Test				,778	,389
Linear-by-Linear Association	,314	1	,575		
N of Valid Cases	50				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 12,00.

b. Computed only for a 2x2 table

The conditions for the chi-square test are satisfied. There is no expected count less than 5. The significance value of the chi-square tests is equal to 57.1% (2-tailed). As a result, H_0 can't be rejected. There is no difference between the two groups in terms of the number of people who remembered their initial energy efficiency and the people who did not.

H.25:

Wilkoxon Rank-Sum test

H_0: There is no difference in the total energy consumption between the people who remembered their initial energy efficiency and the people who did not.

H_1: There is a difference in the total energy consumption between the people who remembered their initial energy efficiency and the people who did not.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of TEC_Round_0 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,575	Retain the null hypothesis.
2	The distribution of TEC_Round_1 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,140	Retain the null hypothesis.
3	The distribution of TEC_Round_2 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,404	Retain the null hypothesis.
4	The distribution of TEC_Round_3 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,286	Retain the null hypothesis.
5	The distribution of TEC_Round_4 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,560	Retain the null hypothesis.
6	The distribution of TEC_Round_5 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,313	Retain the null hypothesis.
7	The distribution of TEC_Round_6 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,712	Retain the null hypothesis.
8	The distribution of TEC_Round_7 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,756	Retain the null hypothesis.
9	The distribution of TEC_Round_8 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,801	Retain the null hypothesis.
10	The distribution of TEC_Round_9 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,756	Retain the null hypothesis.
11	The distribution of TEC_Round_10 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,877	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

The significance value in all the rounds is greater than 5%. Therefore, H_0 can't be rejected. The total energy consumption is not influenced by the fact of remembering the initial energy efficiency or not.

Wilkoxon Rank-Sum test

H_0: There is no difference in the total energy production between the people who remembered their initial energy efficiency and the people who did not.

H_1: There is a difference in the total energy production between the people who remembered their initial energy efficiency and the people who did not.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of EP_Round_0 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	1,000	Retain the null hypothesis.
2	The distribution of EP_Round_1 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,146	Retain the null hypothesis.
3	The distribution of EP_Round_2 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,560	Retain the null hypothesis.
4	The distribution of EP_Round_3 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,252	Retain the null hypothesis.
5	The distribution of EP_Round_4 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,300	Retain the null hypothesis.
6	The distribution of EP_Round_5 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,549	Retain the null hypothesis.
7	The distribution of EP_Round_6 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,790	Retain the null hypothesis.
8	The distribution of EP_Round_7 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,777	Retain the null hypothesis.
9	The distribution of EP_Round_8 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,830	Retain the null hypothesis.
10	The distribution of EP_Round_9 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,741	Retain the null hypothesis.
11	The distribution of EP_Round_10 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,785	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

The significance value in all the rounds is greater than 5%. Therefore, H_0 can't be rejected. The total energy production is not influenced by the fact of remembering the initial energy efficiency or not.

Wilkoxon Rank-Sum test

H_0: There is no difference in the comfort level between the people who remembered their initial energy efficiency and the people who did not.

H_1: There is a difference in the comfort level between the people who remembered their initial energy efficiency and the people who did not.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Comfort_Round_0 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	1,000	Retain the null hypothesis.
2	The distribution of Comfort_Round_1 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,808	Retain the null hypothesis.
3	The distribution of Comfort_Round_2 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,341	Retain the null hypothesis.
4	The distribution of Comfort_Round_3 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,135	Retain the null hypothesis.
5	The distribution of Comfort_Round_4 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,831	Retain the null hypothesis.
6	The distribution of Comfort_Round_5 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,846	Retain the null hypothesis.
7	The distribution of Comfort_Round_6 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,404	Retain the null hypothesis.
8	The distribution of Comfort_Round_7 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,437	Retain the null hypothesis.
9	The distribution of Comfort_Round_8 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,969	Retain the null hypothesis.
10	The distribution of Comfort_Round_9 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,877	Retain the null hypothesis.
11	The distribution of Comfort_Round_10 is the same across categories of Do_you_remember_the_efficiency_of_the_house_you_play_with.	Independent-Samples Mann-Whitney U Test	,907	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

The significance value in all the rounds is greater than 5%. Therefore, H_0 can't be rejected. The comfort level is not influenced by the fact of remembering the initial energy efficiency or not.