Electricity Use Project Study Report

Alexandre Reol* Bernhard Maennel† Luca Kummer‡ Marco Wiedner§ Nikolaos Kaliorakis¶
Vivian Ziemke‖

Students at ETH Zurich

1 Introduction

Reducing electricity consumption can be a great challenge for households [3]. Even if households are thoroughly informed, aware, and value the concept of energy frugality, there are many reasons that prevent them from acting. These reasons can be traced back to cognitive biases (status quo, risk aversion, social mimicry, etc) [2] and other mechanisms that boundedly rational actors are subject to. Further reasons might be, lack of funding for energy efficient device upgrades, paralysis from many potential alternatives, and difficulties in judging better alternatives.

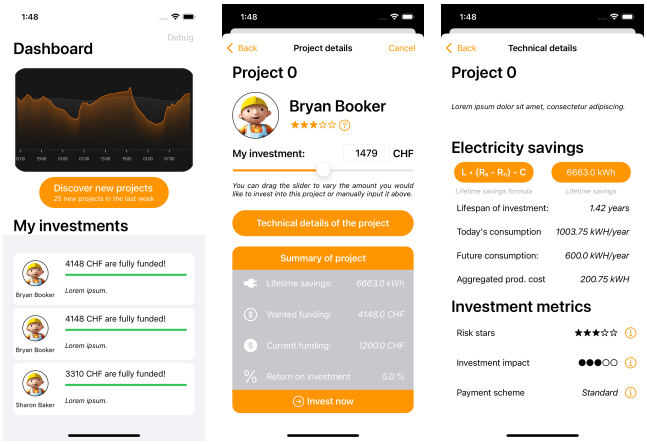
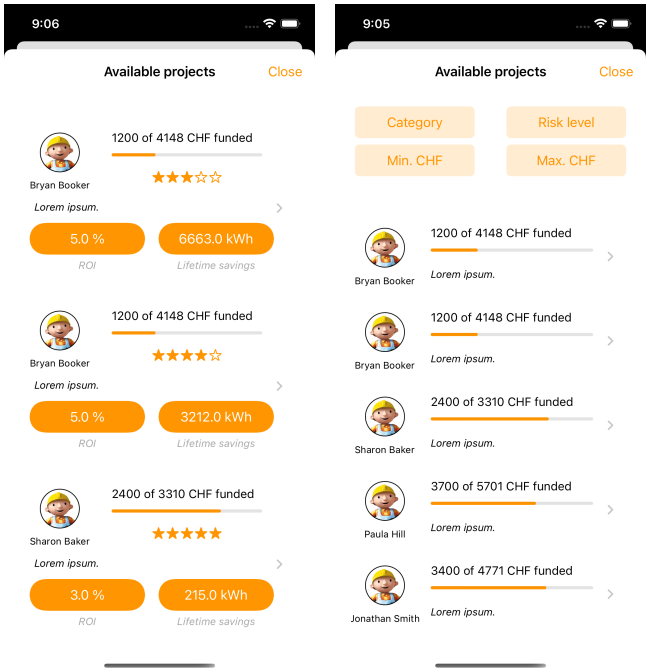


Figure 1: Screenshots of the UI of the App

To tackle some of these issues and with the goal reduce electricity consumption in households we are proposing a two-sided crowd-funding platform. On the one side, fundraisers can start projects were they can get funds to make upgrades (appliances, infrastructure) that will lead to reduction in electricity consumption. On the other side, investors will be able to fund such projects and help fundraisers make updates, and thus reduce electricity consumption, while getting a return on their investment. The app we are going to test in this user study focuses on getting investors to fund electricity saving projects. The main task of a user is to find a project that matches his expectations and fund it with a certain amount of money. This process has to be as user friendly as possible, since if the task of finding a appropriate project is to difficult or to cumbersome, the project which will save electricity will fall through. The main goal of our user study was to find out, which of the two proposed interfaces is the more efficient and convenient to fund projects which have the goal to reduce electricity consumption.

*e-mail:alexandre@reol.ch
†e-mail:bmaennel@ethz.ch
‡e-mail:lkummer@student.ethz.ch
§e-mail:mwiedner@student.ethz.ch
¶e-mail:nkaliorakis@student.ethz.ch
‖e-mail:ziemkev@student.ethz.ch



(a) "No filter" variant (A) (b) "filter" variant (B)

Figure 2: Differences of the Variants

The first variant of the interface (Figure 2a) is more focused on getting the core information of the project directly to the user without needing the user to interact with the interface. So the idea behind this is to scroll through the available projects and look for some key-values that match the expectations of the investor (a.k.a. user), and then one can click on the cell of the project to see further details and the option to invest a certain amount in that project.
The second variant of the investment process (Figure 2b) is focused on filtering the available projects so that it matches the expectations of the investor in advance, while the user will receive the list of the filtered projects. The user then can, without having further information of the exact attributes of the project, click on it to see the exact details and the option to invest a certain amount in it.

2 Study Design

The study took place in two locations, which shared a similar layout. Participants were seated in a desk, in a relatively quit environment and were asked to complete a set of tasks using an iPhone we provided them.

The tests were completed alongside an accompanying survey, to ensure the same test structure was applied to all participants. The

survey was also used to record data (IDs, demographics, SUS score, qualitative, etc). The test followed a within-subjects approach and its duration was $\sim$ 15 minutes per participant. Each participant was asked to complete three trials for each version of the application. Each trial consists of successfully funding a project with specific characteristics, namely Project Duration, Risk, ROI, Lifetime Savings, and Category. The participants were asked to complete the task using either variant A or variant B. Each participant was randomly assigned to be exposed to variant A or B first, to mitigate learning effects. A minute was given to the participant to explore the variant they were given before starting the task. of the app before the trials begun.

The layout of the interface played is the independent variable. More specifically, the "No filter" variant (A) showed more details on the project selection screen, while the "filter" variant (B) had a filtering function on the project selection screen. The project data was also slightly altered between the two variants to make sure they couldn't just select the same projects for whichever variant they used second.

We had three different dependent variables in the study:

- The time it took for participants to finish the task (This was taken directly by the app.)
- The number of clicks performed (for switching pages or pressing buttons) by the user in the tasks. (This was also taken directly by the app.)
- The SUS score from the questionnaire the participants filled out after completing the task.

Our three alternative hypotheses were:

H1: There will be an effect of the project selection screen display on the task completion time.

H2: These will be an effect of the project selection screen display on Usability (SUS score).

H3: These will be an effect of the project selection screen display on lostness . (Visited Pages & Clicks)

Out of eight participants, half were in the age range 18-24, two were in the age range 25-39 and one was in each of the age ranges 40-55 and >55. 5 participants were male and 3 were female. 5 participants had a bachelors degree and the other 3 had a secondary education. None of our participants knew how many kWh of electricity they consume per month on average. One of our participants uses mobile applications/web-apps to invest money every week, another less than once a month, a third participant rarely does and the rest never use them.

3 RESULTS

(In this part, for the sake of readability we rounded the numbers to 3 significant figures.)

- Time Needed

  In the Figure 3, you can see the time each participant needed to complete the task with each variant. "No filter" variant (A)is in blue and "filter" variant (B)is in orange.

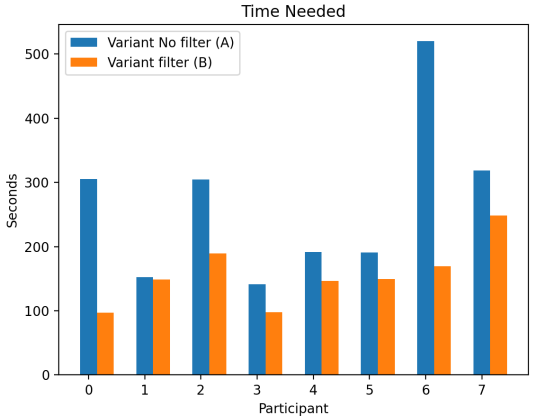


Figure 3: Time needed to complete the tasks

The "No filter" variant (A) has a mean of 266 seconds and a standard deviation of 125 seconds.

The "filter" variant (B) has a mean of 156 seconds and a standard deviation of 49.1 seconds.

As the time needed to complete the task is interval data, we first ran a Shapiro-Wilk test (with alpha = 0.05) on each variant, to see if they could reasonably be assumed to be the results of normal distributions. We got a p-value of 0.131 for "No filter" variant (A)and a p-value of 0.432 for "filter" variant (A). Since both values are greater than 0.05, we decided to keep the null-hypothesis that they were the results of normal distributions. We therefore ran a two-sided paired t-test (with alpha = 0.05) with the null hypothesis that "filter" variant (A) takes the same time in expectation as "filter" variant (B) and the alternative hypothesis that "filter" variant (A) has a different expected completion time than variant (B). Since we got a p-value of 0.0315, which is under 0.05, we rejected the null hypothesis and concluded, that "No filter" variant (A) takes longer in expectation than "no filter" variant (B) and thus "filter" variant (B)is superior by this metric.

- Lostness Factor

  In the chart displayed in Figure 4, you can see the lostness factor [1] of each participant for the task with each variant. We were inspired by the lostness factor as calculated in [1], but in the end we didn't use the lostness factor as we've seen it there. We rather just counted the number of clicks a user performed while completing the tasks. The clicks could be from pressing buttons or switching pages or choosing a project. "No filter" variant (A)is in blue and "filter" variant (B)is in orange.

  The "No filter" variant (A) has a mean of 39.1 and a standard deviation of 31.1.

  The "filter" variant (B) has a mean of 32.6 and a standard deviation of 7.95.

  As the lostness factor is interval data, we first ran a Shapiro-Wilk test (with alpha = 0.05) on each variant, to see if they could reasonably be assumed to be the results of normal distributions. We got a p-value of 0.00134 for "No filter" variant (A) and a p-value of 0.0316 for "filter" variant (A). Since both values are less than 0.05, we decided to reject the null-hypothesis that they were the results of normal distributions. We therefore ran Wilcoxon signed rank test (with alpha = 0.05) with the null hypothesis that " filter" variant (A) has the same expected lostness factor as "no filter" variant (B) and the alternative

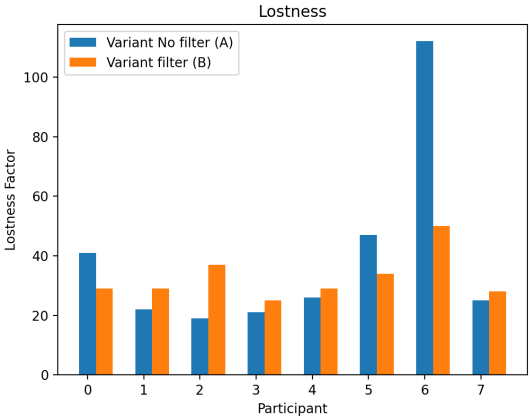


Figure 4: Clicks performed while completing the tasks

hypothesis that ” filter” variant (A) has a different expected lostness factor than ”no filter” variant (B). Since we got a p-value of 0.945, which is over 0.05, we failed to reject the null hypothesis and concluded, that our study didn't provide a statistically significant difference between ” filter” variant (A) and ”no filter” variant (B) in this metric.

- SUS-score

  In Figure 5 below, you can see the SUS-score of each participant for the task with each variant. ”No filter” variant (A)is in blue and ”filter” variant (B)is in orange.

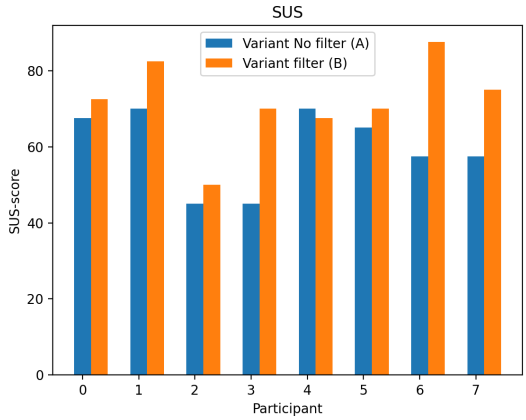


Figure 5: SUS score for each participant

The ”No filter” variant (A) has a mean of 59.7 and a standard deviation of 10.3.

The ”filter” variant (B) has a mean of 71.9 and a standard deviation of 11.2.

While the SUS-score obtained from ordinal data, since it computes scores from several Likert scale reports, it can be treated as interval data. Therefore we first ran a Shapiro-Wilk test (with alpha = 0.05) on each variant, to see if they could reasonably be assumed to be the results of normal distributions. We got a p-value of 0.108 for ”No filter” variant (A) and a p-value of 0.502 for ”filter” variant (A). Since both values are greater than 0.05, we decided to keep the null-hypothesis that they were the results of normal distributions. We therefore ran a two-sided paired t-test (with alpha = 0.05) with the null hypothesis that ” filter” variant (A) has the same expected SUS-score as ”no filter” variant (B) and the alternative hypothesis that ” filter” variant (A) has a different expected SUS-score than ”no filter” variant (B). Since we got a p-value of 0.0180, which is under 0.05, we rejected the null hypothesis and concluded, that ”No filter” variant (A)has a lower expected SUS-score than ”no filter” variant (B) and thus ”filter” variant (B)is superior by this metric.

- Summary and Qualitative Feedback

  Since ”filter” variant (B) turned out to be superior in time taken and SUS-score, and the result for the lostness factor was inconclusive, ”filter” variant (B) seems to have preformed better overall. The qualitative feedback seems to confirm this, with participants saying they had to waste more time scrolling in ”No filter” variant (A) and that they really missed having a filtering function in ” filter” variant (A). It was also suggested that it might be good to have a combination of both approaches. For some participants it was difficult to know which area is clickable and which is just a label displaying information. In Figure 2a one can see that the ROI and the ”Lifetime savings” are display in a button like label, and thus the people were not sure if this is clickable. For example it could have display some additional information about the ROI if this label is clicked. Also a suggestion was, that for distinguishing the different projects in (Figure 2a), we could have used colors to emphasize the ROI or the ”Lifetime savings” amount. ” filter” variant (A) green one for a high ROI/saving and a red color for a low(er) ROI. Different hues for different ”levels” for example. Some users also wanted a ”confirm” button before the investment is actually performed after clicking the ”Invest now” button.

4 Limitations

Since we only conducted a user study with 8 participants, our study is somewhat limited. Every user, as mentioned before, used both interfaces. But only half of them started with ” filter” variant (A), and the other half started with ”no filter” variant (B). In general one can say that for really seeing a difference, we probably would need simply more participants to get more significant data on our variants. There are some concerns about the sample's demographic heterogeneity and how representative it was for our target users. Most of the participant probably aren't that interested in investing in general, thus its difficult to get them the feeling of being an investor. Also not all participants were speaking English equally well, and therefore some of the task descriptions, keywords in the app or the overall idea wasn't always as clear as we may have wished.

5 Future Work

The quantitative and qualitative results indicate that both versions have their flaws. But still, we managed to show, that a filtering option is critical when choosing in our case a project which they want to fund. Furthermore, from the qualitative questions we understand that we must be more selective and more mindful of the information that we display. Too many information points overwhelm and confuse the users whereas too little information can hinder the funding of projects. Careful deliberation of the points displayed followed by a subsequent A/B test can provide a solution to this problem. Additionally, all participants are unaware of their current electricity consumption, which implies they also do not understand the magnitude of electricity metrics. Making the communication of this magnitude more intuitive might be a necessity in order to base the crowdfunding platform.

6 Conclusion

When searching for solutions to reduce the electricity consumption among households, reducing the usage of certain appliances can have non-negligible effect. But upgrading or replacing some of these will exceed the budget of many households. We proposed a idea for households where they can start a project on an app, where they can collect money via crowdfunding and with that, realize the project. It not only saves electricity, but also saves money for the household, which then can be partially payed back to the investor as interest. We then compared two variants of an implementation on the investor side, such that the investing process is fast, clear and not cumbersome. We conducted a user study to measure metrics when performing the task of funding a project. Additionally we collected SUS and feedback from the participants. We concluded that a filtering option when searching for projects is essential, to narrow down the projects that match the criterias for the investor.

Acknowledgments

We want to thank all friends and family members who took time to participate in the user study.

References

[1] S. Doyle. Ux metric: Lostness. 2017.

[2] E. R. Frederiks, K. Stenner, and E. V. Hobman. Household energy use: Applying behavioural economics to understand consumer decision-making and behaviour. *Renewable and Sustainable Energy Reviews*, 41(C):1385–1394, 2015.

[3] S. Sorrell. Reducing energy demand: A review of issues, challenges and approaches. *Renewable and Sustainable Energy Reviews*, 47:74–82, 2015. doi: 10.1016/j.rser.2015.03.002