Electricity Use Project Study Report

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

Students at ETH Zurich

1 Introduction

To reduce electricity consumption in households we are proposing a two-sided crowdfunding 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 others reduce electricity while getting a return on the 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 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.

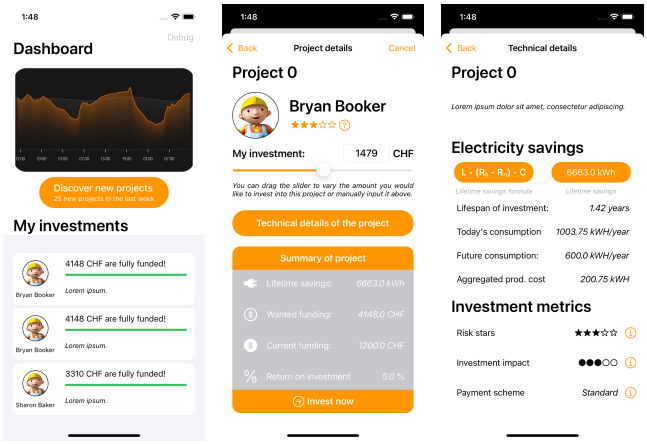
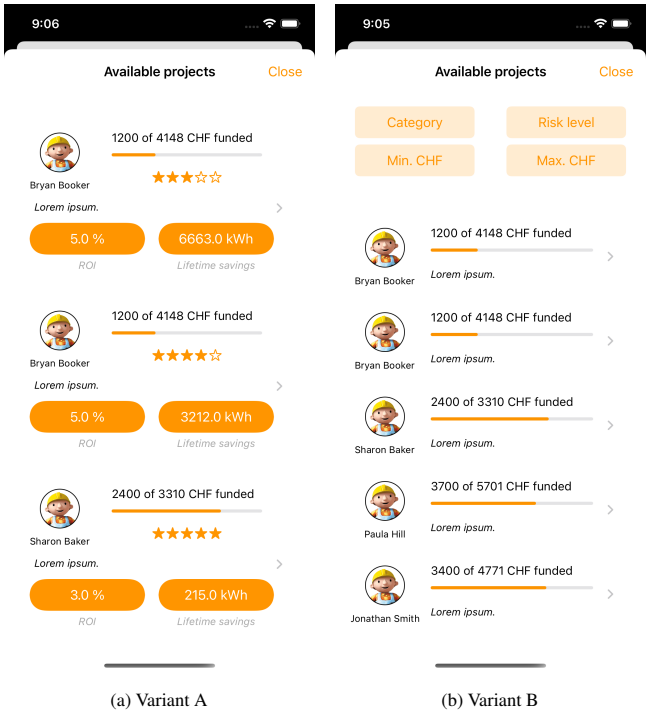


Figure 1: Screenshots of the UI of the App

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.



(a) Variant A (b) Variant B

Figure 2: Differences of the Variants

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 independent variable was the interface layout, specifically, that the A variant showed more details on the project selection screen, while the B variant 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.)

*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

  - The participants' lostness factor (The number of clicks and the number of (distinct) pages performed/visited 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:
  - There will be an effect of the project selection screen display on the task completion time.
  - These will be an effect of the project selection screen display on Usability (SUS score).
  - These will be an effect of the project selection screen display on lostness. (Visited Pages & Clicks)

- 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.

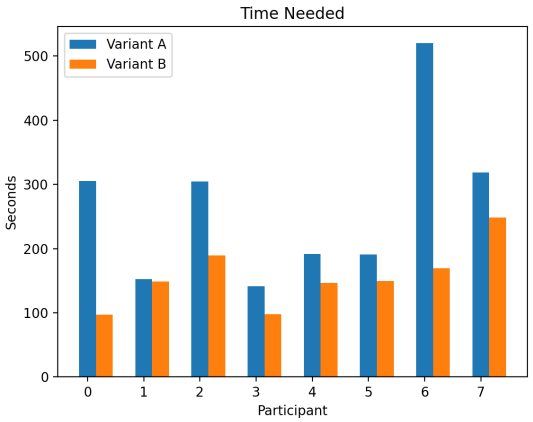
- Of our eight participants, half were in the age range 18-24, two were in the age range 25-39 and one was in each of the age ranges 40-55 and >55. 5 participants were male and 3 were female. 5 participants had a bachelors degree and the other 3 had a secondary education. None of our participants knew 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 chart below, you can see the time each participant needed to complete the task with each variant. Variant A is in blue and Variant B is in orange.



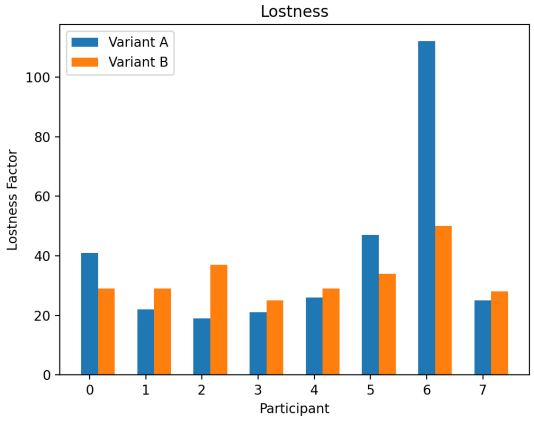
  The A variant has a mean of 266 seconds and a standard deviation of 125 seconds.

  The B variant 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 variant A and a p-value of 0.432 for variant B. 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 A takes the same time in expectation as B and the alternative hypothesis that A has a different expected completion time than B. Since we got a p-value of 0.0315, which is under 0.05, we rejected the null hypothesis and concluded, that variant A takes longer in expectation than B and thus variant B is superior by this metric.

- Lostness Factor

  In the chart below, you can see the lostness factor of each participant for the task with each variant. Variant A is in blue and Variant B is in orange.



  The A variant has a mean of 39.1 and a standard deviation of 31.1.

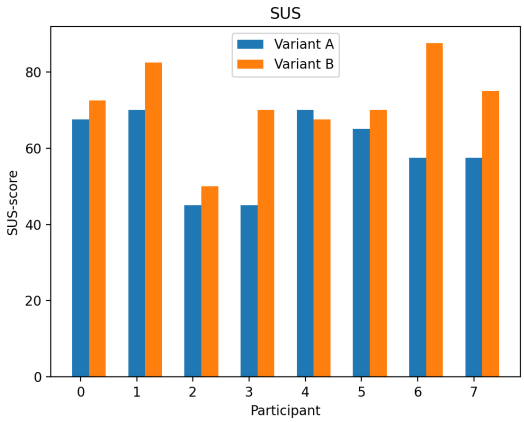
  The B variant has a mean of 32.6 and a standard deviation of 7.95.

  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 variant A and a p-value of 0.0316 for variant B. 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 A has the same expected lostness factor as B and the alternative hypothesis that A has a different expected lostness factor than 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 A and B in this metric.

- SUS-score

  In the chart below, you can see the SUS-score of each participant for the task with each variant. Variant A is in blue and Variant B is in orange.

  The A variant has a mean of 59.7 and a standard deviation of 10.3.



The B variant 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 variant A and a p-value of 0.502 for variant B. 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 A has the same expected SUS-score as B and the alternative hypothesis that A has a different expected SUS-score than B. Since we got a p-value of 0.0180, which is under 0.05, we rejected the null hypothesis and concluded, that variant A has a lower expected SUS-score than B and thus variant B is superior by this metric.

- Summary and Qualitative Feedback

  Since variant B turned out to be superior in time taken and SUS-score, and the result for the lostness factor was inconclusive, 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 variant A and that they really missed having a filtering function in A. It was also suggested that it might be good to have a combination of both approaches.

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 A, and the other half started with B. In general one can say that for really seeing a difference, we probaly would need simply more participants to get more significant data on our variants. One also has to remember that our testers were somewhat different. There were students but also parents which participated in the study, and those two groups weren't equally represented. Also most of them 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

As we saw in the measurements taken and in the feedback received by the participants of the study, 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